2528-9705

Örgütsel Davranış Araştırmaları Dergisi
Journal Of Organizational Behavior Research
Cilt / Vol.: 3, Sayı / Is.: S2, Yıl/Year: 2018, Kod/ID: 81S2105

CONSTRUCTING AND NORMALIZING MATHEMATICS SCHOLASTIC ACHIEVEMENT TEST IN CHABAHAR COUNTY

Reza RASHEDI[1*], Valiollah FARZAD[2]

[1]MSc in Educational Research, Department of Methods and Curriculum, Faculty of Psychology and Educational Sciences, University of Tehran, Tehran, Iran,
[2]Associate Professor, Department of General Psychology, Faculty of Psychology, Central Branch, Islamic Azad University, Tehran, Iran.

***Corresponding Author:**
E-mail: rashedi.r@yahoo.com

ABSTRACT

The primary objective of the present study is constructing and normalizing mathematics scholastic achievement test for the assessment of the first grade high school students' learnt materials in Chabahar County. The study was implemented in three stages. In the preliminary stage, six students were presented with 70 questions in two turns, 35 questions each. The second stage was the experimental implementation of the test in which 70 questions were administered to 56 students and the test's statistical specifications, including the difficulty level and the question recognition power, were analyzed based on a classic model. The remaining questions were arranged based on difficulty coefficient and distributed in two forms (each containing 30 questions) to be later on given to the whole study sample reaching in volume to 600 students. The two test forms' reliability rates were measured based on Cronbach's Alpha method equal to 0.82 and 0.80, respectively. Moreover, factor analysis method made it clear that the mathematic test in the first exam form is comprised of three factors: "algebra; mathematical concepts' understanding; and, calculations". As for the second exam form, it was found comprised of three components: "geometry; algebraic; and, solving equations". The study results were also indicative of the idea that there is no significant difference between the girls and the boys' performance in first grade high school mathematics.

Keywords: *Scholastic Achievement Test, First Grade High School Mathematics, Classic Test Theory, Difficulty Index, Determination Coefficient, Reliability, Validity.*

INTRODUCTION

Undoubtedly, measurement and evaluation are inseparable parts of the education system. Various measurement and evaluation methods and techniques can be used to assess the learning output of the students at schools and universities and determine the extent to which they have succeeded in accomplishing the educational objectives. Generally, the results obtained from educational evaluation provide for making judgment regarding the educational factors in line with the improvement of the education system's activities. Academic or scholastic achievement evaluation is one type of educational assessment that can be undertaken for different lessons and in various grades. The results of academic achievement evaluation return feedbacks that can be employed for the improvement of learning quality and modification of teaching methods and instructional plans (Allen & Wendy, 1979). Academic achievement evaluation can be defined as follows: "assessing the learners' performance and comparing the obtained results with the predetermined educational objectives in line with making decisions in this regard that whether the teacher's instructional activities and the learning efforts of the students have led to an optimal output or not and that if yes, to what extent?" (Saif, 2008, p.27).

Geliş tarihi/Recieved: 11.12.2017 – Kabul tarihi/Accepted: 05.03.2018 – Yayın tarihi/Published: 21.08.2018

Psychometry principles and theories have to be applied to design and codify academic achievement tests and psychological-educational tests, in general. There are various theories proposed up to now for the analysis of the educational-psychological tests in such scientific fields as psychometry and academic achievement assessment and measurement and these theories form the basis of all the analyses (Naderi, 2007). The classic test theory is one of these theories posited in 1904 by Spearman and it is still most frequently used by many of the test-developers for its understandable underlying concepts as well as its high efficiency (Baghi, 1995).
According to the extant grounds in mathematics and the shortages of standard mathematical tests, it is necessary to prepare and codify mathematical achievement tests based on the whole array of the instructional materials as well as the first grade high school mathematics' educational objectives. The substantial value of these achievement tests lies in their improvement of the learning quality, assisting the adoption of educational decisions, providing the students with feedbacks and evaluation of various educational aspects in respect to mathematics (Sharifi, 2009).

STUDY METHOD:

The present study sample volume included 600 first grade high school students (300 girls and 300 boys) from Chabahar County. The sample volume was selected based on a multistep randomized sampling. The behavioral goals of first grade high school mathematics book were determined in the first stage after a thorough study thereof following which a test specification table was prepared. Based thereon, 70 multiple choice objective questions were constructed in two forms (A and B), each containing 35 questions. Next, the questions were presented to a number of experienced first grade high school teachers to investigate them in terms of their match with the intended educational objectives as well as their theoretical significance. Then, six students were administered with the prepared 70 questions to recognize the terms, expressions and ambiguous questions and this was followed by a partial change in some expressions or, in other words, the questions were subjected to an editorial revision.
The second stage was the empirical implementation of the test in which time 70 questions were given in two forms to 56 randomly selected students. Then, the test's statistical specifications, including the difficulty level and the question recognition ability, were analyzed based on classic model and this led to the omission of ten questions. In a third stage, the remaining questions were arranged based on difficulty coefficient and they were again distributed to the foresaid students in two forms (each containing 30 questions) to be analyzed and prepared for the final distribution to the entire study sample volume. SPSS software was employed to analyze the questions based on classic test theory.

STUDY FINDINGS:

A) Descriptive Analysis:

As it can be understood from the frequency distribution table and diagram, the distribution ranges from 3 to 30 for the form A and it was found out that one student (0.3%) out of all the participants to whom Form A had been administered acquired a mark equal to 3 and three students (1%) acquired a mark equal to 30. The highest frequency belongs to the mark 16 in such a way that 30 students (10%) were found belonging to this class. Also, it was made clear

based on Kolmogorov-Smirnov test results' normality that there is no significant difference between the students' mark distribution and the normal distribution in an alpha=0.05 level. Moreover, the raw mark range for the Form B was found between 3 and 30. Out of all the participants to whom Form B of the questions had been administered, two students (0.7%) acquired marks equal to 3 and two students (0.7%) acquired marks equal to 30. The highest frequency went to marks 16 and 17 and 20 students (19.4% in total) were found equally belonging to this class. Furthermore, based on the table, the average score and median of the mark distributions are 15.736 and 16, respectively. Also, the mark scattering's standard deviation and variance were found equal to 5.2535 and 27.599, respectively. In terms of the distribution form, the marks were found with low skewness, -0.002, and low kurtosis, -0.004. It was also figured out based on Kolmogorov-Smirnov test results' normality that there is no significant difference between the students' mark distribution and the normal distribution in an α=0.05 level.

Descriptive statistics table and mark distribution diagram for Form (A) questions

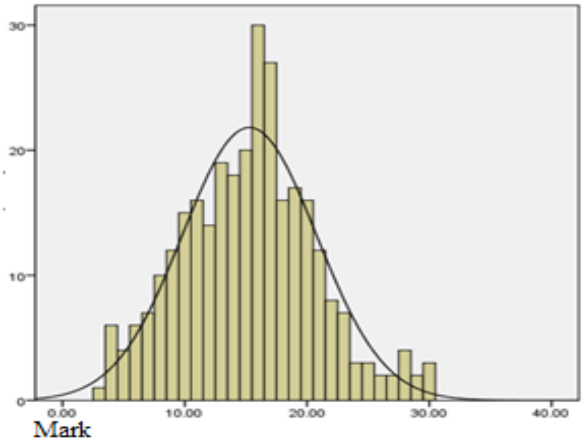

Figure 1: mark distribution diagram Form (A)

Table 1: Descriptive statistics table Form (A)

Central tendency scores	
Mean	15.2833
Median	16
Exponent	16
Sum	4585
Scattering scores	
Maximum	30
Minimum	3
Standard deviation	5.4848
Variance	30.083
Standard error of the means	0.3166
Distribution form scores	
Skewness	0.196
Kurtosis	-0.018
Standard error of skewness	0.141
Standard error of kurtosis	0.281
Kolmogorov-Smirnov test	1.048

Descriptive statistics table and mark distribution diagram for Form (B) questions

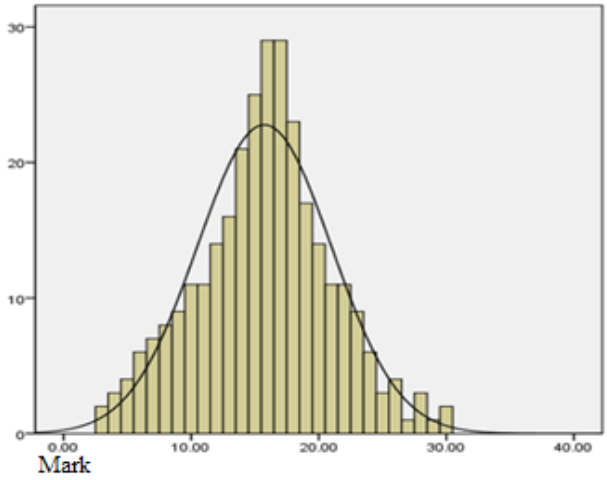

Figure 2: mark distribution diagram Form (B)

Table 2: Descriptive statistics table Form (B)

Central tendency scores	
Mean	15.7367
Median	16
Exponent	16*
Sum	4721
Scattering scores	
Maximum	30
Minimum	3
Standard deviation	5.2535
Variance	27.599
Standard error of the means	0.3033
Distribution form scores	
Skewness	-0.002
Kurtosis	-0.004
Standard error of skewness	0.141
Standard error of kurtosis	0.281
Kolmogorov-Smirnov test	1.228

B) Comparing Female and Male Students' Performance in Test:

Based on the study findings, the performance of both the female and male students from Chabahar County was found identical in regard of both of the forms. The results of the present study are consistent with the findings of the researches undertaken by the psychologists. Of course, it has to be pointed out that the present study results do not conform to the mean marks acquired by the students for June mathematic exams (girls' mean mathematic mark was about 11 and boys' mean mathematic mark was 9).

Descriptive data and t-test for the independent (male and female) groups in regard of Form (A)

Table 3: Descriptive data Form (A)

Gender	Number	Mean	Standard deviation	Mean standard error
Boy	150	15.3667	5.40124	0.44101
Girl	150	15.2000	5.58401	0.45593

Table 4: t-test Form (A)

Variance	t-value	Degree of freedom	Two-way significance level	Difference standard error	95% confidence level	
Equal	0.263	298	0.793	0.63432	-1.08165	-1.41498
Unequal	0.263	297.671	0.793	0.63432	-1.08166	1.41499

Descriptive data and t-test for the independent (male and female) groups in regard of Form (B)

Table 5: Descriptive data Form (B)

Gender	Number	Mean	Standard deviation	Mean standard error
Boy	150	15.4533	5.14173	0.41982
Girl	150	16.0200	5.36515	0.43806

Table 6: t-test Form (B)

Variance	t-value	Degree of freedom	Two-way significance level	Difference standard error	95% confidence level	
Equal	-0.934	298	0.351	0.60675	-1.76073	0.62740
Unequal	-0.934	297.462	0.351	0.60675	-1.76074	0.62740

According to the above tables, it can be concluded that there is no significant difference between the mean marks of the female and male students in terms of Form (A) and Form (B) questions addressing the mathematic achievement in a significance level of alpha=0.05. In other words, the null hypothesis indicating the lack of inter-group differences is accepted in a 95% confidence level.

C) Test Validity:

The content validity of the test has been taken into account during the test preparation stages in such a manner that 70 questions in proportion to the constituent elements of the above table were prepared considering the relative importance of each topic after coming to a clear-cut and detailed description of the assessed field and preparing the characteristics' table. Following the implementation of the empirical stage and elimination of the inappropriate questions, 60 questions were realized appropriate in the end and they were used as the final test form. The investigation of the questions' distribution based on the objectives and the field of study showed that the mathematic achievement test enjoys a relatively satisfactory content validity. Besides, the ideas of the teachers and corresponding experts as well as the head of the math teachers' group were suggestive of the test content's match with the contents of their instructional materials.

D) Test Reliability:

To calculate the test reliability, there was made use of Cronbach's alpha method and the results have been summarized in the table below:

Table 7: Reliability Estimation

Test Form	Number of questions	Alpha	Standardized alpha	Measurement standard error
A	45	0.712	0.814	2.365
B	45	0.706	0.794	2.384

As it can be seen from the table, the math achievement testability rates were found equal to 0.712 and 0.706 for forms (A) and (B), respectively. In the meanwhile, if the scores obtained for each question are transformed to their z-scale, tangible changes will be observed in the obtained coefficients in such a way that the forms A and B's reliability questions will be respectively 0.814 and 0.794 for the standardized questions.

E) Statistical Characteristics of the Test:

The present part offers the findings pertinent to difficulty index and correlation coefficients of the questions inserted into the math test forms based on classic test theory.

The difficulty index of the Form A questions range are scattered from 0.36 (for question 28) to 0.67 (for question 6). In other words, question no.6 is the easiest question and question no.28 is the most difficult question. It is worth mentioning that the difficulty index range of the questions is very narrow, meaning that many questions feature difficulty levels or easiness levels of equal size.

The correlation coefficients of the questions and test were computed based on bi-serial and Pierson correlation tests. In terms of bi-serial correlation, the coefficients obtained for all of the questions were higher than 0.201. In terms of Pierson correlation test, as well, the coefficients of each question and achievement test were found ranging from 0.244 (for question 29) to 0.533 (for question 22). As for the Form B, the questions' difficulty index was found ranging from 0.37 (for question 2) to 0.61 (for question 5). The easiest question of the test form was question five and the most difficult question was question two.

In regard of the bi-serial correlation coefficient, lowest rate, 0.196, was scored for question 19 the highest rate, 0.502, was documented for question 25. Pierson correlation coefficients for each question and the total test score were found in a range from 0.242 (question 19) to 0.553 (question 25).

F) Results of the Test's Factor Analysis:

After subjecting the math achievement test to factor analysis using principal component analysis (PCA), it can be stated that the math test presented in Form A is composed of three factors with eigenvalues equal to 4.255, 3.449 and 1.965 in such a manner that factors one to three, respectively, account for 14.18%, 11.49% and 6.55% of the total variance making for a total variance elaboration of 32.23%.

It was made clear in an investigation of the questions with the highest factor load for their first factor that the phrase "algebra: solving the equation and parities" can possibly be the best title for the first factor.

It was also found out for the second factor through examining the questions related thereto that the phrase "comprehension of the mathematical concepts" can be suggested as its title. As for the third factor, the phrase "calculations" can be proposed as the title based on the nature of the questions demonstrating relative correlation with the factor.

The math test presented in Form B as well was found comprised of three factors with eigenvalues equal to 4.137, 3.855 and 1.876 in such a way that the factors one to three were found each accounting respectively for 13.790%, 12.851% and 6.224% of the variance making for a total of 32.86% of the total variance explanation.

It can be stated in an investigation of the questions with the highest factor load for the first factor that the title "geometry and trigonometry" is the best name for it. For the second factor, the phrase "algebraic calculations" was suggested in an examination of the correlated questions. As for the third factor, the phrase "solving the equations and integrations" is suggested as the title. Scree-Kettle diagram has been given below for inferring the existence of all three primary factors in the two test forms.

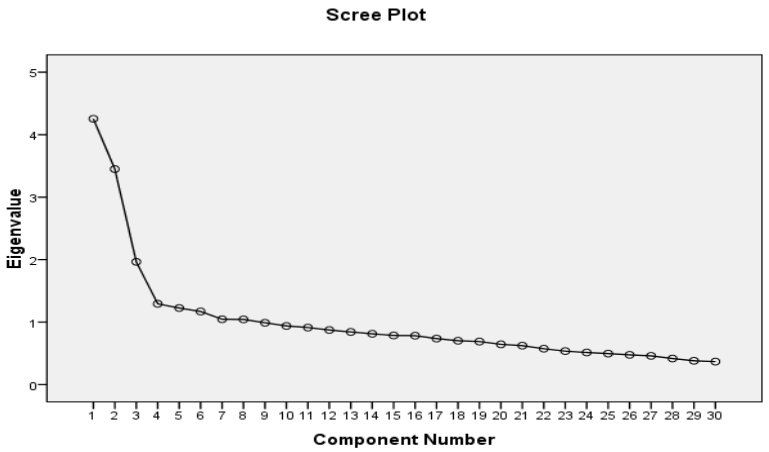

Figure 3: Scree diagram of the form A math test factors

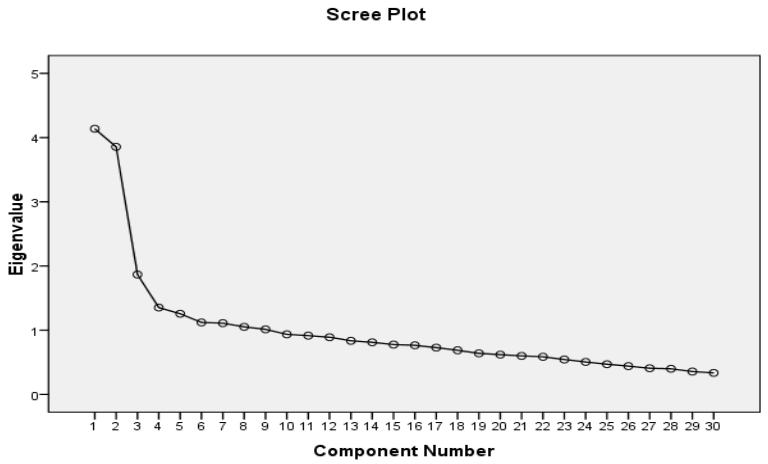

Figure 4: Scree diagram of the form B math test factors

DISCUSSION AND CONCLUSION:

Considering the study findings, it can be stated in general that both of the academic achievement test forms codified herein for the first grade high school math lesson enjoy an appropriate validity and reliability and that the difficulty index was indicative of an intermediate difficulty level of the test forms as well as their high coefficient of determination. Therefore, the two test forms of the first grade high school math can be employed as a suitable scale for the

measurement of the learning levels, academic achievement and comparison of the cognitive aspects of the students in mathematics.

Also, it can be concluded based on the lack of a significant difference between the girls' and boys' scores that the test forms do not have any bias in respect to gender.

The constructed mathematical test possesses relative superiorities some of which are mentioned in the following statements: it is amongst the few constructed standard math tests for the high school grades. This becomes a lot more sensitive when there is felt a need for standard tests for measuring math achievement; various analyses were utilized in the test evaluations so that a precise and standard estimation of the math test could be attained.

References

Allen, Marry J., and Wendy, M., (1979), "an introduction to measurement theories", tr. Ali Delavar, (1995), Tehran, Smat.

Baghi, Heibatiollah, (1995), "improving the quality of educational testing with computer adaptive testing", Tehran: conference on improving public education.

Naderi, Ezzatullah and Saif Naraghi, Maryam, (2007), "assessing and measuring as well as the analytical foundations of their instruments in educational and psychological sciences", Tehran, Arasbaran.

Saif, Ali Akbar, (2008), "educational measurement, assessment and evaluation", Tehran, Dawran.

Sharifi, Hassan Pasha, (2009), "principles of psychometry and psychoanalysis", Tehran, Roshd.