Journal of Science Education
Volume: III, Issue I, 2021, 1-13

Developing a Test to Measure Mathematical Reasoning Among High School Students

Saeedullah*
Rafaqat Ali Akbar**

Abstract

Mathematical reasoning has become an essential tool for secondary school students to excel in every walk of life. A valid measuring tool is a basic need to measure mathematical reasoning. The objective of the present study was to develop a valid and reliable instrument to measure mathematical reasoning among high school students. Five constructs of mathematical reasoning were selected for the development of a mathematical reasoning test. The five constructs of mathematical reasoning were mathematical inductive reasoning, mathematical deductive reasoning, mathematical generalization, mathematical adaptive reasoning, and mathematical problem-solving. The process of developing an instrument comprised of preparation of table of specifications, preparation of the instrument items, and assessing validity and reliability. The content and construct validity and the reliability of the instrument were found in an acceptable range. Principal Component Analysis for extraction of Components was also carried out. The final tool consists of 30 items and Cronbach's Alpha was obtained 0.91. In the end, a valid and reliable instrument was developed.

Keywords: Mathematical reasoning, high school students, validity, reliability

* Corresponding Author: Institute of Education and Research. University of the Punjab, Lahore. Email: sipk43@gmail.com

** Director, Institute of Education and Research University of the Punjab, Lahore. Email: rafaqat@ier.pu.edu.pk

Introduction

Mathematics is necessary for all to progress in every walk of life. The essence of mathematics is mathematical reasoning. Mathematical reasoning helps the students to cope with the challenges of the modern world. In the modern world, machines e.g. calculators are available to perform mathematical operations, but the machine has not been able to replace humans in reasoning. The intellectual life of individuals demands the culture of prosperous reasoning. The reasoning skill enables students to apply mathematics content in a correct and logical way. Mathematical reasoning also enables the students to solve problems within the classroom and in daily life activities (Saragih & Napitupulu, 2015; Wiles, 2013).

Secondary school education is a gateway to professional education and the job market in Pakistan. Therefore, secondary education is considered as an important level of education in Pakistan. Being a developing country, Pakistan needs skilled manpower to cope with poor economic and educational conditions. The developed countries have progressed with the help of their skilled youth. Unfortunately, in Pakistan, attention has not been given to preparing high school students to meet the challenges of the modern world. The literature reveals that the studies for the enhancement of mathematical reasoning have not been conducted in Pakistan (Mohammad, 2004; Sarwar, Yousaf, Hussain, & Noreen, 2009).

There is a dearth of studies to assess the effect of effective teaching strategies on mathematical reasoning. Even after the inclusion of reasoning in the standards of the mathematics curriculum at the secondary school level, assessment for the development of mathematical reasoning could not get the attention of researchers in Pakistan (Akhter, Akhtar, & Abaidullah, 2011). Due to the non-availability of literature, the researcher could not find a measuring instrument of mathematical reasoning in Pakistan. Therefore, the researchers conducted the study to develop a valid and reliable instrument to measure mathematical reasoning.

Review of Related Literature

Mathematical reasoning is the ability to make an argument on the basis of mathematical premises (Mueller, Yankelewitz, & Maher, 2014). Brousseo and Gibel (2005) defined it as the connection between “a condition or observed fact and a consequence”. Duval (2007) defined mathematical reasoning as a logical linking of propositions that may change the epistemic value of a claim.

Bieda, Ji. Drwencke, and Picard (2013) defined it as the procedure of understanding the ideas and concepts of mathematics. Boesen, Lithner, and Palm (2010) defined mathematical reasoning as the line of thoughts adopted to produce assertions and reach conclusions. It can be concluded from the discussion that mathematical reasoning is the process of drawing conclusions based on true mathematical premises.

Mathematical inductive reasoning and mathematical deductive reasoning are the structural aspects of mathematical reasoning. In addition to structural aspects; researchers have identified some more aspects of mathematical reasoning. Mathematical generalization, adaptive reasoning in mathematics, proof, and mathematical problem solving are the common indicators of mathematical reasoning in different studies (Rizqi & Surya, 2017; Sari, 2018). Napitupulu (2017) developed an instrument for the measurement of mathematical reasoning. The level and concepts of the test items were according to the curriculum of Indonesia.

However, the instrument helped the researcher to prepare the test items in the extended form of multiple-choice questions. Another instrument was developed by Sari (2018) to measure mathematical reasoning in the context of Indonesia. Sari included transactions as an indicator of mathematical reasoning. The transaction was a unique indicator of mathematical reasoning in the instrument. It is in the subject of Biology to solve problems. The instrument was developed according to the needs of Indonesian students. The concept of the transaction is not used in the mathematics curriculum of Punjab Textbook and Curriculum Authority. The concept of the transaction was not included in the mathematical reasoning test.

These studies helped the researcher to prepare instruments in the extended form of multiple-choice questions (MCQs). The extended form of the MCQ format is considered an appropriate instrument format to assess students' skills in mathematics (Black & William, 2009; Blum & Borromeo, 2009; Grobe, 2017).

Objective of the Study

The objective of the study was to develop an instrument for the measurement of mathematical reasoning among secondary school students.

Research Question

What kind of an instrument can effectively measure mathematical reasoning among secondary school students?

Population

The population of the study consisted of secondary school students of science groups (Biology and Computer Science).

Sample

The sample of the study was 600 students of the 9th class.

Methodology

The research and development (R & D) method was adopted in the present study. In the light of related literature, five constructs of mathematical reasoning were selected for the development of a mathematical reasoning test. The content of mathematics is generally distributed into 3 portions; arithmetic, algebra, and geometry (Mustafa, 2008; Sidhu, 2019).

Therefore, the researcher also distributed the content of mathematics into these portions. The summary of test items has been provided in the table1.

Table 1.
Table of Specifications

Content	M.I.R	M.D.R	M.G	M.A.R	M.P.S	Total
Portion 1	2	2	2	2	2	10
Portion 2	2	2	2	2	2	10
Portion 3	2	2	2	2	2	10
Total	6	6	6	6	6	30

Note: M.I.R= Mathematical Inductive Reasoning
M.D.R= Mathematical deductive Reasoning
M.G= Mathematical Generalization
M.A.R= Mathematical Adaptive Reasoning
M.P.S= Mathematical Problem solving

Table 1 shows that 2 questions had been selected from each area of mathematics for each aspect of mathematical reasoning. The researcher developed the mathematical reasoning test (MRT) in the extended form of multiple-choice questions to assess mathematical reasoning among the participants of the study. The extended form of MCQs format is considered an appropriate instrument format to assess students' skills in mathematics (Bolarinva, 2015; Grobe, 2017). For each stem of the instrument, there were four options. One option was the most appropriate and correct answer. The other three options were used as distractors. The plausible (wrong-response options) distractors were used in the instrument. The Plausible distractors are good characteristics of an instrument (Blum & Borromeo, 2009; Grobe, 2017). The process of test development has been illustrated in the following figure.

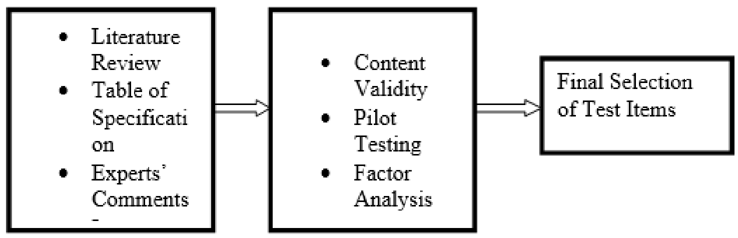

Fig 1. Process of Test Development

Validation of Mathematical Reasoning Test

After developing a mathematical reasoning test, the researcher found its content and constructs validity.

Content Validity of MRT

The MRT was sent to the 10 experts for content validation. For content validation, 10 experts are considered as sufficient numbers (Davis, 1992; Polit & Beck, 2006; Polit, Beck, & Owen, 2007). The definitions and terminologies used in the present study were provided to the subject experts along with a validation sheet. They were requested to rate the instrument items and deliver their written comments on the validation sheet to improve the relevancy of instrument items. The experts were requested to rate the instrument according to the degree of relevance as; the item is not relevant, the item is relevant to some extent, the item is quite relevant, and the item is highly relevant.

The experts rated the items of the instrument and provided their written comments about the test items. Before calculating the content validity index, the experts' responses were recoded for relevance on a scale of 1-4. Recoding of relevance rating, facilitates the calculation of the content validity index (Metin & Korkman, 2021; Ozair et al., 2017). Item–

level content validity index (I-CVI) was calculated using MS-Excel software. Summary of analysis for content validity index has been given in the following table.

Table 2.
Item-Level Content Validity Index of Mathematical Reasoning Test

Item	I-CVI	Item	I-CVI	Item	I-CVI
1	1	11	1	21	1
2	1	12	1	22	1
3	1	13	1	23	1
4	1	14	1	24	1
5	1	15	1	25	1
6	1	16	1	26	1
7	1	17	0.9	27	1
8	1	18	1	28	1
9	1	19	1	29	1
10	0.9	20	1	30	1

Table 2 indicates that the values of I-CVI for all items had been found greater than .75. The values show good content validity (Devellis, 2012; Hadie et al., 2017; Lan, 2018; Yusoff, 2019). Therefore, all the items were retained in the MRT.

Construct Validity

Factor analysis is used to find construct validity statistically (Gaskin & Happell, 2014; Webb, Lubinski, & Benboe, 2014). Therefore, the researcher performed factor analysis using Statistical Package for the Social Sciences (SPSS) version 26 to measure the construct validity. The sample size for factor analysis was 600, which is considered to be an acceptable sample size for factor analysis (Drost, 2015; Everitt, 1975; Gorsuch, 1983; Watkins, 2018). The summary of the analysis has been given in table 3 & 4.

Table 3.
Principal Component Analysis for Extraction of Components

Component a	Component b				
	M.I.R	M.D.R	M.G	M.A.R	M.P.S
M.I.R 1	.579				
M.I.R 2	.719				
M.I.R 3	.638				
M.I.R 4	.673				
M.I.R 5	.801				
M.I.R 6	.765				
M.D.R 1		.710			
M.D.R 2		.649			
M.D.R 3		.675			
M.D.R 4		.633			
M.D.R 5		.666			
M.D.R 6		.732			
M.G 1			.640		
M.G 2			.717		
M.G 3			.655		
M.G 4			.716		
M.G 5			.629		
M.G 6			..808		
M.A.R 1				.542	
M.A.R 2				.518	
M.A.R 3				.542	
M.A.R 4				.520	
M.A.R 5				.542	
M.A.R 6				.518	
M.P.S 1					.692
M.P.S 2					.813
M.P.S 3					.786
M.P.S 4					.824
M.P.S 5					.780
M.P.S 6					.517

Extraction Method: Principal Component analysis
5 Components Extracted
Note: M= Mathematical, I.R= Inductive Reasoning, D.R= Deductive Reasoning
G= Generalization, A.R= Adaptive Reasoning, P.S= Problem Solving

Table 3 illustrates the summary of the component matrix. The Principal Component Analysis was done to extract the components. The results indicated the selection of each test item for 1 component.

Table 4.
Principal Component Analysis for Communalities of MRT

Item	Initial	Extraction	Item	Initial	Extraction
M.I.R 1	1	.666	M.I.R 4	1	.859
M.I.R 2	1	.775	M.I.R 5	1	.908
M.I.R 3	1	.865	M.I.R 6	1	.875
M.D.R 1	1	.937	M.D.R 4	1	.924
M.D.R 2	1	.955	M.D.R 5	1	.732
M.D.R 3	1	.921	M.D.R 6	1	.608
M. G 1	1	.665	M. G4	1	.520
M. G2	1	.660	M. G 5	1	.533
M. G 3	1	.612	M. G 6	1	.769
M. A R1	1	.541	M. A R 4	1	.878
M. A R 2	1	.748	M. A R 5	1	.852
M. A R 3	1	.844	M. A R 6	1	.935
MPS1	1	.668	MPS4	1	.880
MPS2	1	.885	MPS5	1	.844
MPS3	1	.878	MPS6	1	.923

Extraction: Principal Component Analysis
Note: M= Mathematical, I.R= Inductive Reasoning, D.R= Deductive Reasoning
G= Generalization, A.R= Adaptive Reasoning, P.S= Problem Solving

Table 4 indicates that the values of communalities are greater than 0.5. If values of communalities are greater than .5, the items are generally accepted in the test (Costello & Osborne, 2005; Field, 2009; Guadagnoli & Velicer, 1998; Tabachnick & Fidell, 2014).

Reliability of MRT

The internal consistency method was used to find the reliability of MRT. The method is associated with Cronbach's Alpha. The summary reliability statistics have been given in table 5.

Table 5.
Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha based on Standardized items	No. of Items
.91	.96	30

Table 5 indicates that the value of Cronbach's Alpha is 0.91. The value lies in an acceptable range. Moreover, the item analysis was also performed to find difficulty index and discrimination index of test items. The values of difficulty index and discrimination index were in acceptable range. Therefore, all test items were retained in the MRT.

The results showed empirical evidence of its validity and reliability. The values of the item-level content validity index (I-CVI) also shows the high content validity as the value of I-CVI is greater than 0.8. The findings of factor analysis indicate the single factor loading.

Conclusion

Considering the findings of this study, it is concluded that the developed instrument is a valid and reliable instrument. Therefore, the instrument can be used as a standardized test to measure mathematical reasoning among secondary school students.

References

Akhter, N., Akhtar, M., & Abaidullah, M. (2015). The Perceptions of high school mathematics problem solving teaching methods in mathematics education. *Bulletin of Education and Research*, 37(1), 55-77.

Bieda, K. N., Ji, X., Drwencke, J. & Picard, A. (2013). Reasoning and proving opportunities in elementary mathematical textbook. International Journal of Educational Research, 7, 43-57.

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability, *21(*1), 5–31.

Blum, W., & Borromeo, F. R. (2009). Mathematical modeling: Can it be taught and learnt? Journal of Mathematical Modelling and Application, *1*, 45-58.

Boesen, J., Lithner, J. & Palm, T. (2010). The relation between types of assessment tasks and the mathematical reasoning students use. Educ Stud Math, 75, 89-105.

Bolarinwa, O. A. (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Nigerian Postgraduate Medical Journa*, 22(4), 195-201.

Brousseau, G., & Gibel, P. (2005). Didactical handling of students' reasoning processes in problem solving situations. In C. Laborde, M.-J. Perrin-Glorian, & A. Sierpinska (Eds.), *Beyond the apparent banality of the mathematics classroom* (pp. 13–58). Springer US.

Budaev, S.V. (2010). Using principal components and factor analysis in animal behavior research: Caveats and guidelines. *Ethology*, *116*, 472–480.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10(*7), 1-9.

Davis, L. L. (1992). Instrument review: getting the most from a panel of experts. *Applied Nursing Research*, *5(*4), 194–7. https://doi.org/10.1016/s0897-1897(05)80008-4

DeVellis, R. F.(2012). *Scale Development: Theory and Applications*. 3rd ed. SAGE.

Drost, E. A. (2015). Validity and reliability in social science research. *Education Research and Perspectives, 38(*1), 105-123.

Duval, R. (2007). Cognitive functioning and the understanding of mathematical processes of proof. In P. Boero (Ed.), *Theorems in school* (pp. 137–161). Sense.

Everitt, B. S., (1975), Multivariate analysis: the need for data and other problems, *British Journal of Psychiatry, 126*, 237-240.

Field, A. (2009). Discovering statistics using SPSS: Introducing statistical method (3rd ed.). Sage Publications.

Field, A. (2013). Discovering statistics using IBM SPSS statistics. SAGE.

Gaskin, C. J., & Happell, B. (2014). On exploratory factor analysis: A review of recent evidence, an assessment of current practice, and recommendations for future use. *International Journal of Nursing Studies, 51*, 511-521.

Gorsuch, R. L. (1983), Factor Analysis (2nd. Ed). Erlbaum.

Grobe C. S. (2017). Effects of multiple choice options in mathematics learning. *European Journal of Science and Mathematics Education, 5(*2), 165-177.

Guadagnoli, E., & Velicer, W. F. (1998). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103(*2), 265-275.

Hadie, S. N. H., Hassan, A., Ismail, Z. I. M., Asari, M. A., Khan, A. A., & Kasim, F. (2017). Anatomy education environment measurement inventory: a valid tool to measure the anatomy learning environment. *Anatomical Sciences Education, 10*(5), 423–432.

Jung, S. (2013). Exploratory factor analysis with small sample sizes: A comparison of three approaches. *Behavioural Processes, 97*, 90– 95.

Lan, N.T.H. (2018). A perspective of metacognition in solving math problems in Vietnam secondary schools. *Open Access Journal of Science*, 2(6), 339–343.

Metin, M., & Korkman, N. (2021). A valid and reliable scale development study to determine the problems encountered by teachers in the distance education process. *Journal of Educational Technology & Online Learning, 4 (*2), 215-235.

Mohammed, R. F. (2004). Practical constraints upon teacher development in Pakistani schools. *Journal of In-service Education, 1(*30), 101-111.

Mueller, M., Yankelewitz, D., & Maher, C. (2014). Teachers promoting student mathematical reasoning. *Investigations in Mathematics Learning*, *7*(2), 1-20.

Mustafa, M. (2008). *Teaching of mathematics: New trends and innovations* (3rd ed.). Deep & Deep Publications.

Napitupulu, E. L. (2017). Analyzing the teaching and learning of mathematical reasoning skills in secondary school. *Asian Social Science, 13(*12), 32-48.

Ozair, M. M., Baharuddin, K. A., Mohamed, S. A., Esa, W., & Yusoff, M. S. B. (2017). Development and validation of the knowledge and clinical reasoning of acute asthma management in emergency department (K-CRAMED). *Education in Medicine Journal*, *9(*2), 1–17.

Polit, D. F, Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? *Appraisal and recommendations. Research in Nursing & Health, 30(*4), 459–467.

Polit, D.F, Beck, C. T. (2006).The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health. 29(*5), 489–497.

Rizqi, N. R., & Surya, E. (2017). An analysis of students' mathematical reasoning in VIII grade of Sabilina Tembung junior high school. *International Journal of Advance Research and Innovative Ideas in Education, 3(*2), 2395-4396.

Saragih, S., & Napitupulu, E. E. (2015). Developing student-centered learning model to improve high order mathematical thinking ability. *International Education Studies*, *8*(06), 104-112.

Sari, D. P. (2018). Developing instrument to measure mathematical reasoning ability. *Advances in Social Science, Education and Humanities Research, 57,* 34-47.

Sarwar, M., Yousaf, M. I.,Hussain, S., & Noreen, S. (2009). Relationship between achievement goals, meta-cognition and academic success in Pakistan. *Journal of College Teaching & Learning, 6(*5), 19-30.

Sidhu, K. S. (2019). The teaching of Mathematics, (4th ed). Sterling Publisher.

Tabachnick, B. G., & Fidell, L. S. (2014). *Using Multivariate Statistics (6th ed.). Pearson.*

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology, 44*(3), 219–246.

Webb, L., Lubinski, E., & Benboe, K. (2014). Development of an instrument to measure student use of academic success skills: An exploratory factor analysis. *Measurement and Evaluation in Counseling and Development, 47(*3), 171- 180.

Wiles, P. (2013). Folding corners of the habits of mind. *Mathematics Teaching in the Middle School, 19*(4), 208-213.

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation*. Education in Medicine Journal, 11(*2), 49–54.

Zaman, A. (2011). *Relationship between mathematical thinking and achievement in mathematics among secondary school students of North West Frontier province Pakistan* (PhD thesis). International Islamic University, Islamabad.

Citation of this Article:
Saeedullah & Akbar, R.A. (2021). Developing a test to measure mathematical reasoning among high school students. *Journal of Science Education, 3*(1), 1-13.