Test Development Study on the "Probability" Subject in the Middle School Grade

Mehmet Dönmez and İbrahim Yıldırım

Abstract—The aim of this study is to develop a reliable and valid achievement test for the probability learning domain of middle school grade mathematics. For this purpose, the learning outcomes (LO) related to the probability of the Ministry of National Education Mathematics Curriculum (MoNE, 2018) were examined. The questions were prepared in line with these LOs. The test questions were presented to the experts and the final version of the test consisting of 23 questions was created. The draft form of test was conducted to 8th grade students in Turkey and the sample statistics of the test and item statistics were calculated. The mean item difficulty of the test was 0.49 and the KR20 value was 0.83. In line with these results, the developed test is a valid and reliable.

Keywords—Probability, test development, revised Bloom taxonomy.

I. INTRODUCTION

This study was conducted to develop a valid and reliable multiple-choice test in the 8th grade mathematics course for the learning domain of "Possibility". It is aimed at measuring cognitive domains and the items are written according to Revised Bloom Taxonomy.

II. METHOD

This is a test development study and the draft form of the test consists of 23 multiple-choice questions. The test questions were conducted to Middle School Students in Turkey. The duration of the test is 40 minutes. First step of the test development process, four field experts examined the test items and the items were re-arranged in line with the expert opinions. In this study, Tap and Factor package programs were used in statistical analysis.

III. RESULTS

The following findings were obtained about validity, reliability. Test and item statistics were calculated using the TAP program. The KR20 reliability coefficient value is 0.83. The mean of item difficulty of the test was 0.51 and Mean Item Discrimination Index value was 0.49. The test statistics are presented in Table 1.

Table I. Test statistics

	Correct Answer	Item Difficulty Index	Item Discrimination Index	Point-biserial Correlation
Item1	(D)	0.75	0.75	0.32
Item 2	(A)	0.90	0.90	0.23
Item3	(B)	0.86	0.86	0.37
Item4	(C)	0.62	0.62	0.31
Item5	(A)	0.81	0.81	0.29
Item6	(A)	0.70	0.70	0.44
Item7	(D)	0.59	0.59	0.58
Item8	(B)	0.60	0.60	0.62
Item9	(C)	0.29	0.31	0.00
Item10	(B)	0.43	0.43	0.51
Item11	(A)	0.29	0.29	0.56
Item12	(D)	0.71	0.71	0.37
Item13	(A)	0.41	0.41	0.70
Item14	(D)	0.24	0.24	0.74
Item15	(B)	0.54	0.54	0.47
Item16	(B)	0.29	0.29	0.62
Item17	(A)	0.60	0.60	0.48
Item18	(B)	0.46	0.46	0.51
Item19	(D)	0.32	0.32	0.29
Item20	(B)	0.44	0.44	0.60
Item21	(C)	0.16	0.21	0.44
Item22	(D)	0.22	0.22	0.26
Item23	(C)	0.44	0.44	0.62

Also, the distribution of questions is presented according to the Revised Bloom Taxonomy in Table 2.

Mehmet Dönmez
Gaziantep University
Turkey

Ibrahim Yıldırım
Gaziantep University
Turkey

Table II. Distribution of questions according to the Revised Bloom Taxonomy

	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge						
Conceptual Knowledge		i.8, i.11	i.23, i.17	i.2, i.4	i.12, i.14, i.6	
Procedural Knowledge		i.7	i.1, i.9	i.10, i.16	i.3, i.5, i.15, i.18, i.19, i.20	i.13
Meta Cognitive Knowledge				i.21	i.22	

Factor analysis was performed to control the construct validity of the test. The test items were subjected to factor analysis and the analysis was performed with Factor 11.5 program. According to the results of the analysis, the number of factors was recommended as one. Therefore, it is seen that the developed achievement test measure a single structure.

Table III. Factor Loadings

	Factor-1 Loadings	Unrotated Factor Loadings
i1	0.355	0.143
i2	0.453	0.185
i3	0.339	0.130
i4	0.313	0.171
i5	0.403	0.245
i6	0.336	0.209
i7	0.492	0.242
i8	0.585	0.344
i9	0.311	0.013
i10	0.439	0.281
i11	0.587	0.362
i12	0.352	0.082
i13	0.698	0.494
i14	0.812	0.663
i15	0.563	0.835
i16	0.675	0.455
i17	0.385	0.275
i18	0.458	0.215
i19	0.379	0.238
i20	0.590	0.394
i21	0.412	0.271
i22*	0.216	0.112
i23	0.663	0.523

The items factor loadings fell below 0.3 (i22) were eliminated (Büyüköztürk, 2008; Kline, 2011). It was verified that other 22 items had a high factor loading.

IV. CONCLUSION

In this study, a reliable and valid achievement test for the "probability" learning domain of middle school grade mathematics was developed. The developed test contains 22 questions and can be conducted for 40 minutes.

REFERENCES

[1] Büyüköztürk, Ş. (2008). Sosyal Bilimler İçin Veri Analizi El Kitabı (9th ed.). Ankara: Pegema Publishing.
[2] Kline, R. B. (2011). Principles and practice of structural equation modeling. New York: Guilford publications.
[3] MoNE (2018). Mathematics Course Curriculum. Retrieved from http://mufredat.meb.gov.tr/Dosyalar/201813017165445-MATEMAT%C4%B0K%20%C3%96%C4%9ERET%C4%B0M%20PROGRAMI%202018v.pdf

Mehmet Dönmez works as a Mathematics Teacher in a public school in Turkey and also he is a Mechatronics Engineer.
He is a student at Gaziantep University, Department of Measurement and Evaluation.

Assoc. Prof. Dr. Ibrahim Yıldırım is working at Gaziantep University, Turkey. His research interests focus on assessment in education and educational technologies.