Eurasian Journal of Educational Research, Issue 62, 2016, 57-76

A Study on the Development of a Mathematics Creativity Scale[1]

Savas AKGUL[*]
Nihat Gurel KAHVECI[**]

Suggested Citation:

Akgul. S., & Kahveci, N. G. (2016). A study on the development of a mathematics creativity scale. *Eurasian Journal of Educational Research*, 62, 57- 76 http://dx.doi.org/10.14689/ejer.2016.62.5

Abstract

Statement of Problem: Although there is a rapidly increasing interest in creativity in specific domains such as science and math, there are not enough valid and reliable scales to assess students' creativity in these domains. If mathematical creativity potential can be measured, then the curriculum can be adapted to address the needs of creative students. In addition to measuring mathematical creativity potential, it is crucial to identify students at as early an age as possible and develop a mathematics curriculum that would complement and supplement their potential (Balka, 1974; Mann, 2005). For these reasons, valid and reliable scales of measurement are necessary for determining the middle school students' (5th, 6th, 7th and 8th grade: 10-15 years old) mathematical creativity.

Purpose of the Study: The purpose of this study is to develop a valid and reliable Mathematical Creativity Scale for middle school students (5th, 6th, 7th and 8th grade: 10-15 years old).

Method: The main aim of the study is to develop a valid and reliable scale of math creativity for use with middle school students. The pilot study's sample consists of 50 middle school students who attend 5th, 6th, 7th and 8th grades. The field study's sample consists of 297 students who attend 5th, 6th, 7th and 8th grades at 4 middle schools in Istanbul. Exploratory factor

[1] This study was presented at the 1st International Eurasian Educational Research Congress in Istanbul, 24-26 April, 2014

[*] Corresponding author: Dr., Biruni University, Faculty of Education, lugkasavas@yahoo.com.
[*] Dr., Istanbul University, Hasan Ali Yücel Faculty of Education, nihatgurel@yahoo.com.

analysis (EFA) was applied to evaluate the factor structure of the scale. Item analyses were conducted to check on item discrimination, internal consistency, and agreement between scorers, construct-related validity, and face validity.

Findings: Item analysis including the calculation of item discrimination, item total and item remainder values showed that each item was consistent with the entire scale and that distinctive powers of the items were at an acceptable level. To test internal consistency, Cronbach's Alpha coefficient value was found to be .80, which supports scale reliability. Since it has a naturally subjective scoring process, the scale was rated by two scorers, and the same scorer re-rated it at a different time. The correlation showed that the scale has an interrater reliability and intrarater reliability. Test-retest coefficient values showed that the scale measurements are consistent. Content, construct and face validity results are presented as a part of validation works of the scale. Acceptability to students' results are also presented.

Results: The study results showed that this scale was an appropriate instrument to evaluate middle school (5th, 6th, 7th and 8th grade) students' mathematical creativity. This scale can be used for measurement of the students' creativity in mathematics for the purpose of educational interventions and also for the purpose of determining gifted and talented students in mathematics.

Keywords: Mathematical Creativity, Gifted and Talented Students, Mathematical Creativity Scale.

Introduction

When we consider the contemporary attraction of gifted and talented education, the Marland Report (1972) can be seen as an important cornerstone in the field. Another peak point of interest in creativity might be dated to 1950 as Treffinger (2004) described Guilford's (1950) presidential address to the American Psychological Association, "his extensive work on the Structure of Intellect model also served as a catalyst for new and expanding conceptions of intelligence and giftedness" (Treffinger, 2004, xxiii). Therefore, it was literally a juxtaposition of giftedness and creativity. As creativity is an invaluable skill for the new century, gifted education programs must take it into consideration. Examination of literature on giftedness showed us some patterns such as: studies mostly focuses on creativity's relationship to intelligence, identification and development of general and domain specific creativities.

Although there is a rapidly increasing interest in creativity in specific domains such as science and math, there are not enough valid and reliable scales to assess students' creativity in these domains. If mathematical creativity potential can be measured, then the curriculum can be adapted to address the needs of the creative students. In addition to measuring mathematical creativity potential, it is crucial to identify students at as early an age as possible and develop a mathematics curriculum that would complement and supplement their potential (Balka, 1974; Mann, 2005). For of these reasons, valid and reliable scales of measurement are needed for determining the middle school (also called secondary school in some educational contexts) students' (5^{th}, 6^{th}, 7^{th} and 8^{th} graders: 10-15 years olds) mathematical creativity.

Theoretical Framework

As a result of the fact that mathematics, once known as the science of numbers and formulas, started to focus on skills such as revealing patterns in life, producing solutions to problems encountered, and critical and analytical thinking, mathematical creativity emerged as an important cognitive and affective factor.

While first observations on creativity tend to evaluate creativity in the realms of art and aesthetics, it is possible to see that, today, it has expanded to a broader field including social, natural and mathematical sciences. Also, approaches associating creativity with the greats, people with superior intelligence (Charles Darwin, Wolfgang Amadeus Mozart, Isaac Newton, William Shakespeare, etc.), gave way to the view that creativity is a skill or a way of thinking that might be possessed by anyone (Sternberg, Kaufman & Grigorenko, 2008). Beghetto and Kaufman (2007) state that even though creativity is common in children, their creative potential regresses with aging. They also suggest that this results from the need to fit in to society and that this situation suppresses creative potential.

The concept of creativity is comprehended in various disciplines in different ways, and is perceived differently even within the same discipline, therefore no commonly accepted definition of creativity is available. In addition to its complexity in definition, some experts think that mathematical creativity is not achievable because of the unfavorable approaches of conventional education to teaching creativity (Meissner, 2005).

When it comes to math creativity, the Russian psychologist Krutetskii characterized mathematical creativity in the context of problem formation (problem finding), invention, independence, and originality (Krutetskii, 1976). Others included fluency, flexibility and originality in the context of mathematical creativity (Mann, 2005; Haylock, 1987; Jensen, 1973; Kim et al., 2003). In addition to these notions, Holland also included elaboration and sensibility (Mann, 2005). Singh (1988) characterized mathematical creativity as "the process of formulating a hypothesis concerning cause and effect in a mathematical situation, testing and retesting these hypotheses and making modifications and finally communicating the results" (p. 15).

Studies on mathematical creativity investigated either situations of flexibility, fluency and originality of students' answers to a given problem, or mathematical progress obtained from situational data (Mann, 2005; Balka, 1974; Evans, 1964; Getzels and Jackson, 1962; Haylock, 1984; Jensen, 1973; Singh, 1988). These three factors were measured as follows: Fluency was used alone (Dunn, 1976, cited Lee, Hwang and Seo, 2003; Krutetski, 1976), fluency and flexibility were used together (Turkan, 2010), fluency and originality were used together (Mainville, 1972), and finally fluency, flexibility and originality were used together (Evans, 1964; Balka, 1974; Lee, Hwang and Seo, 2003). Singh (1987) developed a mathematical creativity test by means of considering three counterparts of creativity which he indicated as fluency, flexibility and originality and elaboration. While this study reverberates with some traces of the studies mentioned, this study takes into consideration fluency, flexibility, and originality to develop a mathematical creativity test as Singh (1987) contended, "to develop better insight in the field of mathematics" in a Turkish context (p. 181).

In light of the information previously given, a mathematical creativity scale was determined by grading fluency, flexibility and originality of answers given by students. As seen from the determination of mathematical creativity based on these three contexts, mathematical creativity includes divergent thinking in mathematics. Even though Balka (1974) predicated convergent thinking upon measurement of mathematical creativity, Mann (2005) and Erbas and Bas (2015) used this scale in their research and concluded that this method would not lead to valid measurement for students participating in their study. Therefore, they evaluated mathematical creativity by grading the criteria that require divergent thinking, listed in Balka's scale. Furthermore, in his three-dimensional structure of intellect, Guilford (1956) associated divergent thinking with creativity, appointing to it several characteristics such as fluency, flexibility, originality and elaboration.

Whereas mathematical creativity is attributed to the field of mathematics, competence determined by these three subjects served as the basis: logical thinking, spatial thinking and problem formation. Carlton's (1959) analysis distinguished between two types of creative mathematical minds: logical and intuitive minds. "Intuitionalists are described as those who use geometrical intuition, are capable of "seeing in space," and "have the faculty of seeing the end from afar," whereas the logicians work from strict definitions, reason by analogy and work step-by–step through a very great number of elementary operations" (Carlton, p. 234-236). Moreover, Benbow, Lubinski and Kell (2013) noted that spatial ability had a unique role in the development of creativity. In addition, there are also opinions claiming that problem formation is related to creativity. For example, Jensen (1973) viewed posing mathematical questions related to a given scenario as a measure of creativity. Jensen (1973) considers that for being creative in mathematics, students should be able to pose mathematical questions that extend and deepen the original problem as well as solve the problem in a variety of ways.

Method

Research Sample

Convenience sampling was used for this study. The pilot study's sample consists of 50 middle school students (23 male and 27 female) who attend 5th, 6th, 7th and 8th grades. The field study's sample consists of 297 middle school students (144 male and 153 female) who attend 5th, 6th, 7th and 8th grades at 4 schools which in Istanbul. Those 4 schools were selected because of the administrators' and teachers' interest in and support for the study. The age of participants ranged from 10 to 15 years. For determining test re-test reliability, the mathematical creativity scale was applied to a group of 40 students (19 male and 21 female), twice, two months apart. All students voluntarily participated in the study.

Research Instrument and Procedure

In this study, a Mathematical Creativity Scale (MCS) was developed by the researchers that can be used to measure mathematical creativity of students. Firstly, the literature on mathematical creativity was reviewed. Methods used in measurement of mathematical creativity were examined. Opinions were taken from teachers who had studied mathematical creativity. These opinions and relevant literature were analyzed, then a scale of 10 items was developed (Balka, 1974; Olkun and Akkurt, 2012; A. Baykal, personal communication, November 14, 2012).

As Haylock (1987, p. 68) emphasized, "Performance on such tasks of divergent production in mathematical situations appears to be unrelated to performance on general divergent production tests, suggesting that divergent production in mathematics might be a specific ability and not just a combination of some sort of general creative ability and mathematics attainment." So, mathematical creativity can be considered as a different cognitive factor apart from either general creativity or mathematical attainment. That is to say, in light of all of these theoretical considerations in literature, in this study the math creativity scale's items were developed based on divergent production in these three important subjects: logical thinking, spatial thinking and problem formation.

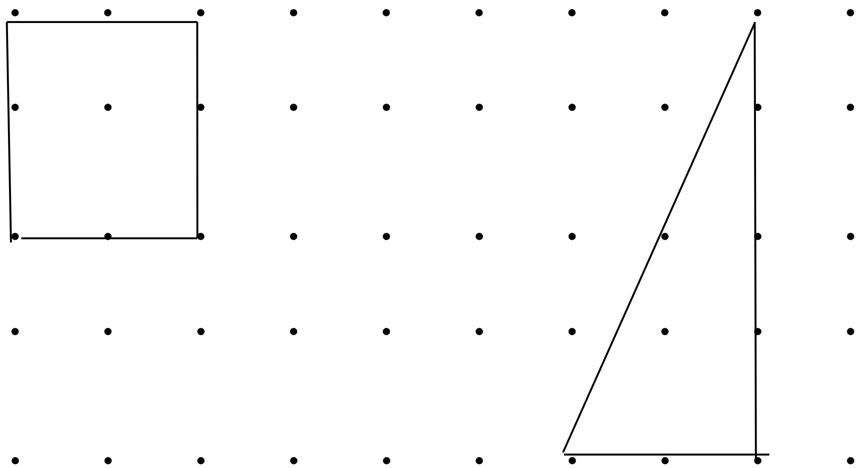
Secondly, seven experts were interviewed about their opinions on whether this scale of ten items was able to measure mathematical creativity. For this purpose, an Expert's Questionnaire was prepared. After the questionnaires were collected, items accepted by at least 5 of 7 experts were retained. After having taken the experts' opinions, some modifications were made to the initial 10 items to develop an outline of the math creativity scale.

A pilot study was conducted, and considering problems encountered during this application, necessary measures were taken. For example, the initial 40 minutes' duration was raised to 50 minutes upon review of the pilot study's implications. The main reason for this increase was the fact that students had not encountered mathematical questions requiring divergent thinking before, and the prescribed time was insufficient for them to complete the given task. In accordance with another implication of the pilot study application, and suggestions of experts both in the areas of mathematics education and creativity, the number of questions was decreased to five, with intent to minimize the challenges caused by limited time. After the pilot study, the scale took its final form of five items.

Brief Descriptions and Illustrations of the Scale

The scale was designed for group administration. The time limit is 50 minutes as mentioned above. The examiner seeks to make the students feel at ease but also work hard to complete the items. There is a general instruction printed on the first page of the scale.

ITEM 1: The area of the below polygons equals 4 units square.



Draw polygons whose areas equal 4 units square in the given space. Make sure the polygons you draw are not the same.

The first item is about finding areas (geometry). This item was adapted from Haylock's (1987) article, namely a framework for assessing mathematical creativity in school children. This item lets students think spatially and make manipulations of shapes, as well as think analogically to demonstrate their problem-solving ability. In addition to these skills, intuitive minds, which are categorized as mathematically creative minds by Carlton, were in force to perform in this item. Results were categorized in 20 different categories. The most frequent category was finding areas by "square units" and the least frequent one was finding "surface areas of three-dimensional figures."

ITEM 2: Example Problem: Ali is three times older than Ahmet. The sum of the two is 48. So what are their ages?

Answer: Ahmet: 1 fold, Ali: 3 fold.

Their total ages are 4 fold and this equals 48. 48:4=12.

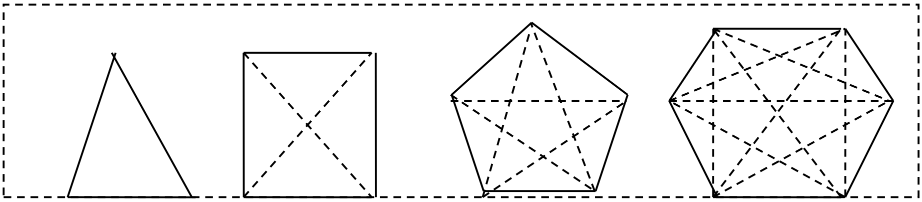
If one fold is 12, 3 fold is 12x3=36.

Ahmet is 12 and Ali is 36.

Make up questions as in the example which requires you to find two unknowns. Make sure the information provided in your question is adequate to find the desired unknowns. Make sure the problem is correct. You *do not have to* write the answer for your questions. If the space given below is not sufficient, you can use the back side of the paper.

The second item was originally developed by researchers. This item primarily aims to determine students' ability to generate mathematical problems according to number of unknowns. Since this item requires them to write problems that have two unknowns, they have to write problems that provide at least two equations. Results showed that "number" word problems are the most frequent and "physics and chemistry" problems were the least frequent problems. This item allows students to think analytically and holistically, as well as analogically. This item also requires the use of problem-solving abilities.

ITEM 3: In the below figure, there are some polygons and diagonals belonging to them (dotted lines). Find out what properties change as the number of edges of a polygon increases. Write down all the changes you find.



Examples: 1. The number of diagonals increases. 2. The number of intersections of diagonals increases.

The third item was adapted from Balka's Scale (1974) because of its novelty and strong relevance to mathematical creativity. This item was used to define mathematical creativity alone, so it is adapted to the scale without any change. This item allows students to use geometrical intuition ("seeing in space," and "have the faculty of seeing the end from afar"). The least frequent category was "the equation of lines changes," and the most frequent category was "different angles occur."

ITEM 4: Write down problems the solutions of which follow the same arithmetic operations. Write as many problems as you can. First, decide on the solution and then set up questions suitable for the solution.

Example: Solution: 10-5 = 5

1. Cemil gave 5 walnuts out of his 10 walnuts to his brother. How many walnuts does he have at the end?
2. Erhan has 10 liras and pays 5 liras for a meal in the restaurant. How much is left?
3. If $\Delta + 5 = 10$, $\Delta = ?$ And so on…

The fourth item was constructed by researchers through inference from Olkun and Akkurt's (2012) study. Since the ability to pose problems in mathematics is linked to mathematical creativity (Jensen, 1973), this item was constructed to evaluate students' creativity in mathematics based on fluency, flexibility and originality of their answers. Problems were written according to their own determined solutions. So, students found an opportunity to create problems in any mathematical area that came to mind. In this item, students used their problem-solving, analytic and also analogical thinking skills. Many categories were found in students responses. The most frequent one was "number" word problems that can be solved with one arithmetic operation, and the least frequent was geometry problems that can be solved with three or more arithmetic operations.

ITEM 5: There are two examples below.

Please provide examples that express a similar relationship to the below examples. If the space given below is not sufficient, you can use the back side of the paper.

Examples:

1. All natural numbers are integers, but all integers are not natural numbers.

2. If A= {1, 2, 3, 4} and B = {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}, every element of A is also an element of B. Yet the reverse (every element of B is also an element of A) is not correct.

The fifth item was constructed by the researchers. This item was about logical thinking in mathematics, since Carlton emphasized two different mathematical creative minds as the logical and intuitive minds. This item provided students the opportunity to work with strict definitions and reason by analogy which can demonstrate students' logical abilities. Categories found in this item were fewer than in others. The most frequent category was about the relationship of number sets, and the least frequent was about the relationship of three-dimensional figures.

A scoring method to evaluate the mathematical creativity scale was established based on fluency, flexibility and originality. Different methods were used to grade these three dimensions in the literature. In this study, fluency was scored by giving one point for every idea produced by the student, flexibility was scored by categorizing ideas produced and awardingone point for every category, and finally, the originality score was determined by giving the highest mark to the rarest mark. The originality sample score was determined by formulating an exponential function based on giving the highest score to the rarest exceptional ideas. The table given below was used for scoring originality (Baykal, 2009). Sampling consists of 297 students. Accordingly, originality was determined based on following Table 1.

Table 1.
Grading Key Based on Sparsity for MCS Originality Score

FREQUENCY	SCORE
1	9
2	8
3-4	7
5-8	6
9-16	5
17-32	4
33-64	3
65-128	2
129-256	1
257-445	0

Results

The detailed results of the validity and reliability studies on the MCS were developed and conducted with 297 participants as given below.

Correlation values between fluency, flexibility, originality and total scores vary from .72, which is the lowest value (between fluency and originality), to .96, which is the highest value (between originality and creativity). Correlation between items and total scores vary from .72 to .96.

Table 2.
Correlation Coefficients of Fluency, Flexibility, Originality and Total Scores (N=297)

	Fluency	*Flexibility*	*Originality*	*Total*
Fluency	1.00			
Flexibility	.78	1.00		
Originality	.72	.93	1.00	
Total	.78	.95	.96	1.00

All the correlations are significant for confidence level of .01 (2-tailed)

Item Analyses

Item analyses consist of item discrimination, item-total, item-remainder and internal consistency of test.

Item Discrimination

Item discrimination was calculated according to the ratio "*t*" involving upper and lower groups of 27 percent. Items were considered acceptable for the final form of the test only if their "t" values were significant for confidence level of .01 or less. As seen in Table 3, the mathematical creativity scale was discriminating for every question and total score.

Table 3.
Results of MCS Item Discrimination Analysis

	group	N	Ort	SS	SHx	t	Sd	P
Q1	Upper	80	23.18	7.89	0.89	11.03	158	.00
	Lower	80	11.11	5.74	0.64	11.01	142.38	
Q2	Upper	80	17.01	6.22	0.70	13.09	158	.00
	Lower	80	6.51	3.55	0.40	13.05	123.62	
Q3	Upper	80	29.63	11.94	1.34	13.63	158	.00
	Lower	80	9.21	6.05	0.68	13.58	115.23	.00
Q4	Upper	80	12.92	8.39	0.94	7.61	158	.00
	Lower	80	5.31	3.07	0.34	7.58	98.32	.00
Q5	Upper	80	15.30	8.38	0.94	10.55	158	.00
	Lower	80	4.58	3.53	0.40	10.50	104.60	.00
Total	Upper	80	98.05	19.42	2.19	25.16	158	.00
	Lower	80	36.73	9.86	1.10	25.06	115.37	.00

Internal Consistency

Internal consistency was examined using two methods. Firstly, coefficients between every item and total score, along with Pearson product-moment coefficients, were calculated. Data is given in Table 4. Correlation values between items vary from .33 mean value (between second and fifth item) to .54 which is the highest value (between first and second item). Correlation between items and total score varies from .59 to .82 and is relatively high.

Table 4.
Correlation Coefficients of Item Scores Along With Item and Total Scores (N=297)

Item	One	Two	Three	Four	Five	Total
One	1.00					
Two	.54	1.00				
Three	.51	.48	1.00			
Four	.58	.52	.53	1.00		
Five	.43	.33	.44	.39	1.00	
Total	.72	.71	.70	.71	.49	1.00

All the correlations are significant for confidence level of .01 (2-tailed)

Secondly, by calculating Cronbach's Alpha internal consistency coefficient, it was determined how consistent the evaluation of skills defined in the test was, according to the obtained results. Alpha value adapted from the scores made by 297 middle school students was found to be .80.

Item-total correlation, which is corrected by removing every item one by one, and Cronbach's Alpha coefficients were calculated. Details are shown in Table 5. Corrected item-total correlation and all of the Alpha values were lower than they were without removal of items. All of these results demonstrate that the test satisfies the predicted internal consistency and that all the items measure the same entity, both individually and together.

Table 5.
Effects of Removing Items Individually from Test (N=297)

Item	*Corrected item-total correlation*	*Alpha value in case of removal (Total alpha value of test = .8)*
One	.51	.73
Two	.58	.73
Three	.62	.69
Four	.46	.74
Five	.56	.73

Scoring Reliability

Because of the contingent subjectivity in interpretation of grading rules, a test of the grading system by someone who had not involved in the testing process was required. Students' grades were determined by two raters, independently. One of the raters **was** the researcher, and the other one was not connected to this research. Pearson product-moment correlation coefficients between the two score sets are given in Table 6. Correlation between scores has a median of .87 and varies from .81 to .91.

Table 6.
Correlation between Two Raters (Intra-scorer Reliability)

Item	*Correlation*
One	.81
Two	.87
Three	.86
Four	.91
Five	.88

Consistency of grading made by the same rater at different times was examined to satisfy the scoring reliability. For this purpose, the researchers regraded the creativity scores that they had graded a month earlier. Pearson product-moment correlation coefficients between scores graded by the same person at different times are shown in Table 7. Correlation between the scores has a median of .91 and varies from .88 to .96.

Table 7.
Correlation between MCS Grades Given by the Same Rater at Different Times (Inter-scorer Reliability)

Item	Correlation between grades given by the same rater at different times
One	.91
Two	.91
Three	.88
Four	.92
Five	.96

In order to evaluate the test-retest reliability, mathematical creativity scale was applied to a group of 40 people, twice, two months apart. Pearson product-moment correlation coefficients of the results obtained from these two applications are shown in Table 8.

Table 8.
MCS Test-retest Reliability

Item	Correlation of data obtained at different times
One	.74
Two	.62
Three	.69
Four	.71
Five	.58

Validity

Validity is analyzed in two different terms. One of them is construct validity. Guilford (1950) suggested factorial validity as the first step in checking validity of creativity tests. Factorial validity mentioned here is a type of construct validity, and it is determined by factor analysis of test scores. Correlation of the information acquired from this test and factor analysis made with the main components results in just one factor. The solution cannot be transformed. As seen in Table 9, all of the items are loaded largely (.60 to .71) on a single factor which corresponds to 42% of total variance.

Table 9.
MCS Factor Load of Every Item

Item	Component (n = 297)
One	.60
Two	.71
Three	.67
Four	.60
Five	.69

The other validity type determined in this study is the face validity, which questions whether the items appear to measure mathematical creativity. In order to obtain a measure related to the appearance of the creativity test, 40 mathematical education researchers and mathematics teachers were asked to answer the following question: "Which of the items in the test are able to measure the mathematical creativity of a middle school student?" Results are given in Table 10.

Table 10.
MCS Face Validity, Teachers and Mathematics Researchers

Item	Able to measure creativity	Unable to measure creativity
One	35	5
Two	31	9
Three	33	7
Four	32	8
Five	28	12

Acceptability by Students

Finally, to be able to measure the mathematical creativity of middle school students, a test should be more or less acceptable to students. To test this, 80 students were asked which of the questions in the test they found interesting. Results are given in Table 11.

Table 11.
Attitude of Students Towards MCS test

Item	Interesting	Not interesting
One	62	18
Two	65	15
Three	64	16
Four	63	17
Five	57	23

Discussion and Conclusion

First of all, as mathematical creativity scores were determined by students' responses to items in the means of fluency, flexibility and originality, researchers expected that these three would have high correlation with each other and also with the total score that is called the creativity score. Results showed that correlations are between .78 and .96, as expected. On the other hand, originality and flexibility had a correlation value of almost 1 with creativity, which means that they can be used to assess creativity alone. So, if it is difficult to determine originality or flexibility on any creativity scale, it is possible to determine creativity by flexibility or originality individually. As it is seen that fluency has a minimum correlation with total score, fluency should not be used alone. Runco and Acar (2012, p. 67) stated that "fluency is not as closely tied to creativity as is originality and flexibility".

In order to determine content validity, items accepted by at least five of seven experts were kept in the scale. Construct validity is determined by factor analysis of test scores. Correlation of the information acquired from this test and factor analysis of the main components results in just one factor. The solution cannot be transformed. All of the items are loaded largely (.60 to .71) on a single factor which corresponds to 42% of total variance. Kline (1993) states that items chosen for a single factor test should be loaded on a single factor, and the loading should be greater than 0.3 in most instances. Balka (1974) noted in his study that answers given to situations requiring divergent thinking gather under a single factor. According to these researchers, creativity of middle school students should be single factoral. Thus, it is significant for us to obtain only one factor in our analyses. For determining face validity, 40 mathematical education researchers and mathematics teachers were asked to answer the question of face validity. The answers given by teachers and educational researchers demonstrate that the test has a high level of face validity. The measurement construct should be acceptable by students (Hu and Adey, 2002). To test students' acceptability, 80 students were asked which of the questions in the test they found interesting. Responses demonstrated that acceptance by students was highly satisfying.

In the item analysis, the corrected item-total correlation coefficients were found to be between .46 and .62; and, item remainder values were found to be between .69 and .73. To determine the item discrimination, the ratio "*t*" involving upper and lower groups of 27 percent was calculated and results showed that the mathematical creativity scale was discriminating for every question and total score. Internal consistency was examined with two methods: a) coefficients between every item and total score, along with Pearson product-moment coefficients. Correlation coefficients were found to be between .39 and .72 b) Cronbach's Alpha internal consistency coefficients were calculated and were found to be .80. This consistency value is satisfying for a test containing only five questions.

As a result of the subjectivity in interpretation of grading of the scale, three methods were used to determine scoring reliability. First of all, intrascorer reliability was determined by calculating The Pearson product-moment correlation coefficients for each items grade that were determined by two raters, independently. Correlations were found to be .81 and above. This scale can therefore be regarded as having intrascorer reliability. Secondly, interscorer reliability was determined by calculating the Pearson product-moment correlation coefficients for each item grade that was rated by the same scorer at different times. Correlation between the scores has a median of .91 and varies from .88 to .96. These results showed that the scale also has a high degree of interscorer reliability. Eventually, test-retest reliability was determined by the obtained data from a group of 40 students who took a part in this study twice, two months apart. Pearson product-moment correlation coefficients of the results obtained from these two applications were between .58 and .74. As the former grading reliabilities demonstrate, the scale has also test re-test reliability.

Creative Ability in Mathematics (CAMT) developped by Balka (1974) can be identified as the first scale particular to the middle school grades, with completed validity and reliability analyses. Recently, models measuring mathematical creativity and new measuring tools are being developed; however, there is no sufficient analysis of their validity and reliability. This study also reports reliability and validity analyses of the developed mathematical creativity scale as appropriate for the middle school students.

Since validity of a scale is not a dichotomy, like valid or not, more studies are necessary to further validate the test. Relationships between this test and other mathematics creativity tests, as well as other general and scientific creativity tests, might be studied. More importantly, predictive validity should be determined. Confirmatory Factor Analysis should be applied to find an additional evidence for validity. Neverthless, it can be said that this scale was an appropriate instrument to evaluate middle school (5^{th}, 6^{th}, 7^{th} and 8^{th} grade) students' mathematical creativity.

Four types of creativity tests are identified, namely divergent thinking tests, attitude and interest inventories, personality inventories, and biographical inventories (Hocevar, 1981). Nevertheless, divergent thinking tests have dominated the field of creativity assessment, while other inventories also offer important and useful information about creativity (Runco and Acar, 2012). For example, Erbas and Bas (2015) found significant correlations between creative ability in mathematics and openness to experience and conciousness in their study among Turkish students. Openness to experience and instrinc goal orientation was also found as a significant predictor of mathematical creative ability. For further studies, variables tested by Erbas and Bas (2015) and others mentioned above, namely: interest, motivation and attitude toward mathematics, mathematical self-efficacy, personality traits, biographical information, might be reconsidered to evaluate for different grade levels. The Mathemetics Creativity Scale might be used as an intsrument for comparing and analyzing the variables tested.

As constructivist learning approaches have become popular, skills that are expected from individuals have changed. As a result of the fact that mathematics, once known as the science of numbers and formulas, started to focus on skills such as revealing patterns in life, producing solutions to problems encountered, and critical and analytical thinking, mathematical creativity emerged as an important cognitive/affective factor. So, this scale can be used for measurement of students' creativity in mathematics for the purpose of new assessment applications and educational interventions, as well as for the purpose of determining gifted and talented students in mathematics.

Acknowledgements

This study was supported by the Istanbul University Scientific Research Projects (project number: 30220). We would like to thank Professor Ali Baykal from Bahcesehir University, Professor Sinan Olkun from TED University, Research Assistant Sule Gucyeter from Istanbul University and Research Assistant Seyma Sengil Akar from Hacettepe University, Turkey and Assistant Professor Eric L. Mann from Hope College, USA, for their kind help.

References

Balka, D. S. (1974). The development of an instrument to measure creative ability in mathematics. *Dissertation Abstracts International, 36*(01), 98. (UMI No. AAT 7515965).

Baykal, A. (2009). *Yaratıcılık: Eğilimi ve Eğitimi[Creativity: its tendency and education].* Öğrenme ve Öğrenme Bozuklukları. Ataköy: İstanbul Kültür Üniversitesi Yayınları.

Beghetto, R. A., & Kaufman, J. C. (2007). Toward a broader conception of creativity: A case for mini-c creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 1, 73-79.

Carlton, V. L. (1959). An analysis of the educational concepts of fourteen outstanding mathematicians, 1790–1940, in the areas of mental growth and development, creative thinking, and symbolism and meaning, *Dissertation Abstracts*, *20*(06), 2131. (UMI No. AAT 5904782).

Erbas, A. K. & Bas, S. (2015). The contribution of personality traits, motivation, academic risk-taking and metacognition to the creative ability in mathematics, *Creativity Research Journal*, 27:4, 299-307, DOI: 10.1080/10400419.2015.1087235

Evans, E. W. (1964). Measuring the ability of students to respond in creative mathematical situations at the late elementary and early junior high school level. *Dissertation Abstracts*, 25 (12), 7107. (UMI No. AAT 6505302).

Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence: Explorations with gifted students*. New York: Wiley.

Guilford, J. P. (1950). Creativity. *American Psychologist*, 5, 444–454.

Guilford, J. P. (1956). The structure of intellect. *Psychological bulletin*, 53(4), 267.

Haylock, D. W. (1984). Aspects of mathematical creativity in children aged 11–12. Unpublished Doctoral dissertation, Chelsea College, University of London.

Haylock, D. W. (1987). A framework for assessing mathematical creativity in school children. *Education Studies in Mathematics, 18*(1), 59–74.

Hocevar, D. (1981). Measurement of creativity: Review and critique. *Journal of Personality Assessment*, 45, 450-464.

Hu, W. & Adey, P. (2002). A scientific creativity test for middle school students. *International Journal of Science Education*, *24* (4): 389-403 DOI:10.1080/09500690110098912

Jensen, L. R. (1973). The relationships among mathematical creativity, numerical aptitude and mathematical achievement. *Dissertation Abstracts International, 34* (5), 2168.

Kell, H. J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2013). Creativity and technical innovation: Spatial ability's unique role. *Psychological Science, 24,* 1831–1836.

Kim, H., Cho, S., Ahn, D. (2003). Development of mathematical creative problem solving ability test for identification of gifted in math. *Gifted Education International, 18*, 184-193.

Kline, P. (1993). *The Handbook of Psychological Testing.* London: Routledge.

Krutetskii, V. A. (1976). *The psychology of mathematical abilities in school children.* Chicago: University of Chicago Press.

Lee, K. S., Hwang, D., & Seo, J. J. (2003). A development of the test for mathematical creative problem solving ability. *Journal of the Korea Society of Mathematical Education Series D: Research in Mathematical Education*, *7*, 163–189.

Mainville, W.E. (1972) A study of the effects of mathematics activity material upon certain aspects of creative thinking ability of prospective elementary school teachers. Doctoral dissertation, Michigan State University.

Mann, E. L. (2005). Mathematical Creativity and School Mathematics: Indicators of Mathematical Creativity in Middle School Students. PhD thesis. University of Connecticut.

Marland, S. P. (1972). Education of the gifted and talented: Report to the Congress of the United States by the U.S. Commissioner of Education. Washington, D.C: Department of Health, Education and Welfare.

Meissner, H. (2005). Creativity and Mathematics Education. Proceeding of The Third East Asia Regional Conference on Mathematics Education. Shanghai, Nanjing, Hangzhou. Retrieved, 11 Dec. 2012, from http://math.ecnu.edu.cn/earcome3/SYM1.htm

Olkun, S. & Akkurt Z. (2012). Matematiksel Problemlerin çözümünde yaratıcı eylem örnekleri [Creative action examples in mathematical problem solving]. Retrieved, 21 March 2012, from *www.ustunveozel.com/matyar.ppt*

Singh, B. (1987). The development of tests to measure methematical creativity. *International Journal of Mathematical Education in Science and Technology*, *18*(2) 181–186.

Singh, B. (1988). *Teaching-learning strategies and mathematical creativity*. Delhi, India: Mittal Publications.

Sternberg, R. J., Kaufman, J.C. & Grigorenko, E. L. (2008). *Applied intelligence*. New York: Cambridge University Pres.

Treffinger, D.J. (2004). *Creativity and Giftedness.Essential Readings in Gifted Education.* Thousand Oaks, California: Sage.

Türkan, Y. (2010). Matematiksel Üretkenlik Testi (MÜT)'nin İlköğretim 6., 7. ve 8. Sınıflar düzeyinde psikometrik özelliklerinin İncelenmesi[A research study on psychometric properties of the creative mathematical ability test (c-mat)]. Yüksek Lisans Tezi. Eskişehir, Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü.

Matematik Yaratıcılık Ölçeği Geliştirmeye Yönelik Bir Çalışma

Atıf:

Akgul. S., & Kahveci, N. G. (2016). A study on the development of a mathematics creativity scale. *Eurasian Journal of Educational Research*, 62, 57- 76 http://dx.doi.org/10.14689/ejer.2016.62.5

Özet

Problem Durumu: Her ne kadar fen ve matematik gibi belirli alanlarda yaratıcılık konusuna ilişkin sürekli artan bir ilgi olsa da, bu alanlarda öğrencinin yaratıcılığını ölçebilecek yeterli sayıda geçerli ve güvenilir ölçek bulunmamaktadır. Matematiksel yaratıcılık potansiyelinin ölçülebilmesi, öğretim programının da yaratıcı öğrencilerin ihtiyaçlarına uygun olarak düzenlenmesine imkân sağlayabilir. Bu potansiyelin ölçülmesinin yanı sıra yaratıcı öğrencileri mümkün olduğu kadar erken yaşta tespit edebilmek ve onların mevcut potansiyellerini destekleyecek bir matematik öğretim programı geliştirebilmek de eşit derecede önem taşımaktadır (Balka, 1974; Mann, 2005). Bu nedenle ortaokul kademesindeki (5., 6., 7., ve 8. sınıflar: 10-15 yaş grubu) öğrencilerin matematiksel yaratıcılıklarını tespit edebilmek amacına hizmet eden geçerli ve güvenilir ölçeklerin geliştirilmesi bir gerekliliktir.

Araştırmanın Amacı: Bu araştırmanın amacı ortaokul kademesindeki (5., 6., 7., ve 8. sınıflar: 10-15 yaş grubu) öğrencilerin matematiksel yaratıcılıklarını tespit edebilecek geçerli ve güvenilir bir Matematiksel Yaratıcılık Ölçeği geliştirmektir.

Yöntem: Araştırmanın öncelikli amacı ortaokul kademesindeki öğrenciler için geçerli ve güvenilir bir Matematiksel Yaratıcılık Ölçeği geliştirmektir. Pilot çalışmanın

örneklem grubunu 5.-6.-7.-8. sınıflarda eğitim gören 50 ortaokul öğrencisi oluşturmaktadır. Alan çalışmasındaki örneklem kümesi ise İstanbul'daki dört okulda yine 5.-6.-7.-8. sınıflarda eğitim gören 297 öğrenciden oluşmaktadır. Ölçeğin faktör yapısını belirlemek Faktör Analizi (AFA) uygulanmıştır. Madde analizleri kapsamında, her bir maddenin ölçeğin bütünü ile tutarlığını belirlemek için düzeltilmiş madde toplam korelasyonları incelenmiştir. Maddelerin ayırt edicilik gücü ise ilişkisiz örneklemler için *t* testi ile belirlenmiştir. Ölçeğin güvenirliğini belirlemek için Cronbach Alpha katsayısı ve test-tekrar test güvenirliği katsayısı hesaplanmıştır. Puanlama güvenirliğinin hesaplanması için puanlar arasındaki Pearson korelasyon katsayısı hesaplanmıştır.

Bulgular: Madde analizi; madde ayırt edicilik, madde-toplam ve madde-kalan değerlerinin hesaplanmasını kapsamaktadır ve analiz sonuçları her maddenin ölçeğin tamamı ile tutarlı olduğunu ve maddelerin ayırt edicilik gücünün kabul edilebilir düzeyde olduğunu göstermektedir. Ölçeğin Cronbach Alpha iç tutarlılık değeri 0.80 olarak hesaplanmıştır ve bu değer ölçeğin güvenilir olduğuna işaret etmektedir. Doğal olarak bu durum subjektif bir puanlama sürecini kapsadığı için, ölçek iki farklı kişi tarafından puanlanmıştır. Ayrıca aynı kişiler farklı zamanlarda ölçeği tekrar puanlamışlardır. Bu süreç sonunda elde edilen korelasyon değeri hem değerlendiriciler arası güvenilirliğin hem de tek hakem güvenilirliğinin yüksek olduğunu göstermektedir. Test- Tekrar Test güvenirlik değerleri ölçek ölçümlerinin tutarlı olduğuna işaret etmektedir. Kapsam, yapı ve görünüş geçerlik sonuçları ölçeğin geçerliği kapsamında yapılan çalışmalar içerisinde sunulmuştur. Öğrenci test sonuçlarının kabul edilebilirliğine de ayrıca yer verilmiştir.

Sonuç ve Tartışma: Öncelikle öğrencilerin matematiksel yaratıcılık puanları onların akıcılık, esneklik ve orijinallik boyutunda ölçek maddelerine verdiği cevaplara göre hesaplandığı için araştırmacılar bu üç boyutun hem birbirleri ile hem de yaratıcılık puanı olarak ifade edilen toplam puanı ile yüksek korelasyona sahip olması gerektiğini öngörmüşlerdir. Sonuçlar da tahmin edildiği gibi sırasıyla .78 ve .96 korelasyon değeri olduğunu göstermektedir. Diğer taraftan, orijinallik ve esnekliğin yaratıcılık ile neredeyse 1.00'a yakın bir korelasyon değerine sahip olduğu görülmektedir ki bu durum sadece bu iki boyutun yaratıcılığı tanımlamak için kullanılabileceğine işaret etmektedir. Bu noktada herhangi bir yaratıcılık ölçeğinde orijinallik ya da esnekliğin tanımlanması zor olsa da, yaratıcılığın sadece esneklik ya da sadece orijinallik ise ortaya çıkarılması mümkündür.

Kapsam geçerliliğinin belirlenmesinde yedi uzmandan en az beşinin kabul ettiği maddeler ölçekte tutulmuştur. Yapı geçerliği, test sonuçlarının faktör analizi ile belirlenmiştir. Bu testten elde edilen bilgiler ve ana bileşenler ile yapılan faktör analizinin korelasyonu ölçeğin tek faktörlü bir yapıda olduğuna işaret etmektedir. Tüm maddelerin büyük ölçüde tek bir faktörde birleştiği (.60 - .71) ve toplam varyansın yüzde 42'sini açıkladığı görülmüştür. Görünüş geçerliliği belirlemek amacıyla 40 matematik eğitimi araştırmacısı ve matematik öğretmeninden görüş alınmıştır. Araştırmacı ve öğretmenlerin cevapları testin görünüş geçerliliğinin yüksek olduğuna işaret etmektedir. Öğrencilerin kabul edilebilirliğini belirlemek içinse 80 öğrenciye "testte hangi soruları ilginç buldukları" sorusu yöneltilmiştir.

Öğrencilerin cevapları öğrencilerin kabulünün oldukça tatmin edici olduğunu göstermektedir.

Madde analizinde, düzeltilmiş madde-toplam korelasyonları .46 ila .62 arasında hesaplanmıştır ve madde–kalan korelasyon değerlerinin ise .69 ila .73 arasında olduğu görülmüştür.

Madde ayırt ediciliğinin hesaplanmasında üst ve alt yüzde 27'lik grubu içeren "*t*" değeri hesaplanmış ve sonuçlar matematiksek yaratıcılık ölçeğinin her sorusunun ve toplam puanının ayırt ediciliğinin yüksek olduğunu göstermiştir. İç tutarlılık iki farklı yöntem ile incelenmiştir: Pearson Çarpım Moment Korelasyon Katsayısının yanı sıra madde-toplam katsayıları da hesaplanmıştır. Korelasyon Katsayısı .39 - .72 arasında bulunmuştur. Cronbach-Alpha iç tutarlılık katsayısı hesaplanmış ve .80 olarak bulunmuştur. Bu değer sadece beş soru içeren bir test için yeterli düzeydedir.

Ölçeğin değerlendirilmesinde sübjektifliğin sağlanması için puanlama güvenilirliğini belirlemek için üç yöntem birlikte kullanılmıştır. İlk olarak her bir madde için iki değerlendiricinin bağımsız olarak puanladığı Pearson Çarpım Moment Korelasyon Katsayısı hesaplanarak tek hakem güvenirliği hesaplanmış ve korelasyon değerinin. 81 ve üzerinde olduğu sonucuna ulaşılmıştır. Bu sonuç, ölçeğin tek hakem güvenilirliğinin olduğunu göstermektedir. İkinci aşamada her bir madde için aynı değerlendiricinin farklı zamanlardaki puanlaması için Pearson Çarpım Moment Korelasyon Katsayısı hesaplanarak değerlendiriciler arası güvenirlik belirlenmiş ve korelasyon değerlerinin. 88 ila .96 arasında olduğu görülmüştür. Puanlamalar arasındaki korelasyonun orta değeri ise .91'dir. Bu sonuçlar aynı zamanda ölçeğin değerlendiriciler arası güvenirliğinin de yüksek düzeyde olduğunu göstermektedir. Son olarak test-tekrar test güvenirliği ise iki aylık bir zaman aralığında testin uygulandığı 40 kişilik öğrenci grubundan alan veriler ile hesaplanmıştır. Bu iki uygulamadan elde edilen sonuçların Pearson Çarpım Moment Korelasyon Katsayısı .58 ila .74 olarak hesaplanmıştır. Sonuç olarak test-tekrar test güvenirliğinin de yeterli düzeyde olduğu görülmüştür.

Testin geçerliği konusunda ise sadece geçerli ya da geçerli değil şeklinde kesin bir ayrım yapılamayacağı için testin geçerliğine ilişkin daha fazla çalışma yapmak gerekmektedir. Geçerlik için ayrıca Doğrulayıcı Faktör Analizi yapılmalıdır. Bununla birlikte araştırmacılar olarak bu ölçeğin ortaokul düzeyi (5., 6., 7., 8. sınıflar) öğrencilerinin matematiksel yaratıcılıklarını ölçmek için uygun olduğu kanısındayız. Ölçek hem gerekli müdahalelerde bulunmak için hem de matematikte üstün yetenekli öğrencilerin tespiti için öğrencilerin matematik alanındaki yaratıcılıklarının tespitinde kullanılmaya elverişlidir.

Anahtar Kelimeler: Matematiksel Yaratıcılık, Üstün Zekalı ve Yetenekli Öğrenciler, Matematik Yaratıcılık Ölçeği.