The critical thinking skills test for 5-6 year-old children (CTTC): A study of validity and reliability

Kevser Tozduman Yaralı, *Kırklareli University,* Turkey, *kevsertozduman@hotmail.com* ORCID: 0000-0002-7765-0461

F. Abide Güngör Aytar, *Gazi University,* Turkey, *abidegungor@gmail.com* ORCID: 0000-0002-0327-3317

Abstract. In this study, it was aimed to develop a measurement tool to determine the critical thinking skills of children. The Critical Thinking Skills Test for 5-6-Year-Old Children (CTTC) is composed of a total of 41 items and six subscales including interpretation, explanation, evaluation, inference, analysis, and self-regulation. As a result of the Confirmatory Factor Analysis performed for the validity of the test, it was found that X2/df (1,19) value was less than 5, GFI, AGFI, NFI, and CFI values were higher than .90, and RMSEA value was below 0.08. It was found that the Spearman correlation coefficient value applied for the criterion validity was over the moderate level and positive (r=0.66; p<0.05). The reliability of the test was analyzed by using the test-retest method. As a result of the analyses, it may be asserted that "The critical thinking skills test for 5-6-Year-Old Children (CTTC)" is a valid and reliable measurement tool.

INTRODUCTION

Critical thinking is one of the most important skills to develop in order for individuals to achieve success today (Gray, 2016). Critical thinking is mental activities which enable us to reach the accurate information and make decisions accordingly by evaluating the claims/evidences or hypotheses (Huitt, 1998). Critical thinking includes skills such as questioning, estimating, investigating, developing a hypothesis, analyzing, reflecting, reviewing, comparing, evaluating and presenting an opinion (Greenberg, 2016). Different disciplines like philosophy, education and psychology have defined critical thinking on the basis of their own fields, causing the emergence of different definitions without any detail (Daniel and Auriac, 2011; Gündoğdu, 2009). In order to remove this complexity in the definition of critical thinking, a Delphi Project (1988-1989) was carried out under the leadership of Peter A. Facione, an important philosopher and writer on this issue. In the Delphi Project aiming to reveal the views of 46 professionals from different disciplines who specialized in critical thinking and to arrive at an agreement, the experts' views were determined and the skills comprised by critical thinking were clarified. As a consequence, the experts determined that critical thinking included cognitive skills such as explanation, analysis, interpretation, inference, evaluation, and self-regulation (Figure 1).

Critical thinking is a skill that is acquired starting from infancy and can be developed lifelong (Leon, 2015). Heyman (2008) argues that children start to develop reasoning skills at around the age of three and usually question the misinformation relayed by adults and underlines the importance of social experiences for critical thinking of children in this period. Kuhn (1999) states that critical thinking of children is grounded on the skill of theory of mind and even younger children can think about thinking. 3-6-year-old preschool children show success in thinking about their own behaviors and are interested in observing the incidents in their environment. Children in this period are competent to develop and comprehend the past and future mental images, concerning their desires. They can display the skill of explaining their own and other people's behaviors, using their experience and knowledge (Epstein, 1993). The observations and interviews conducted with 5-7-year-old children years also have revealed that children can have critical thinking at early ages (Davis-Seaver, et al., 2003; McCall, 2017).

Critical thinking is indicated to be a basic competence for every individual to choose the correct information in the present day, when the flow of information is intense. It helps the individual realize herself/himself and poses a necessity for a healthy individual and a healthy society (Gündoğdu, 2009). Because serious problems emerge concerning how to use the information, as it becomes easier to reach the information. Nowadays, children have necessary information to find their direction when they get lost in a forest; however, they are unable to use that information (Dewey, 1993). As it has become easier to reach the information, it is necessary for children to have critical thinking in order to choose, evaluate, and use the correct information, discern the wrong information and diverge from stereotypes (Halpern, 2003). Critical thinking is an important tool for producing scientific information and developing science, as well (Küçükali and Akbaş, 2015).
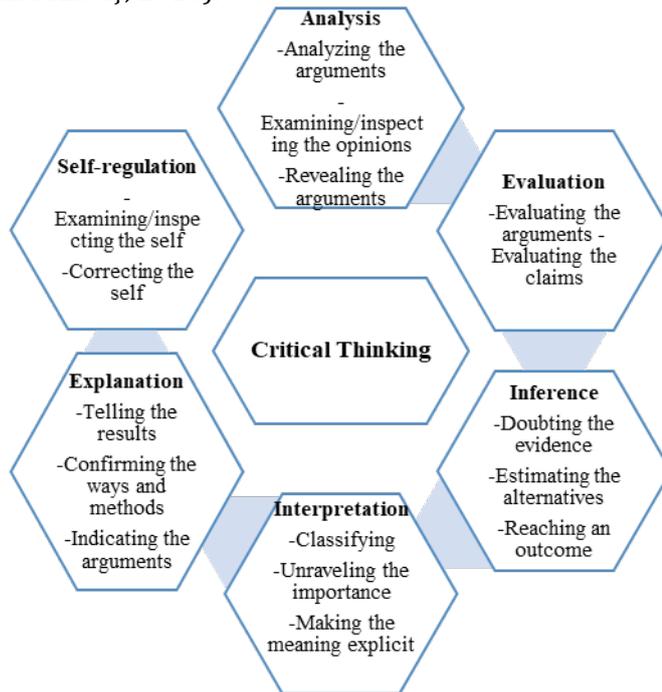
FIGURE 1. *Essence of critical thinking skills (Facione, 1990).*

Critical thinking is one of the higher-order thinking skills (Bruning et al., 1995). In Turkey, there are various measurement tools developed for skills such as problem solving (Oğuz and Köksal Akyol, 2015; Ömeroğlu, Büyüköztürk and Aydoğan, 2009) and creative thinking (Kaytez, 2015), which are among higher-order thinking skills, for preschool children. In addition, it is seen that measurement tools aimed at critical thinking skills are usually developed or adapted for primary and higher education children. The "Ennis-Weir Critical Thinking Composition Test" adapted by Koç (2007), "Critical Thinking Scale" developed by Demir (2006) for fourth and fifth grade primary education students, "California Critical Thinking Disposition Inventory" adapted by İskifoğlu and Ağazade (2013) for adults, "Watson-Glaser Critical Thinking Appraisal" adapted by Çıkrıkçı (1996) for high school and university students and the "Critical Thinking Scale" developed by Semerci (2000) for university students, are among the measurement tools developed or adapted for primary and higher education children and adults on critical thinking. In Turkey, a measurement tool has been developed by Karadağ, Demirtaş and Yıldız (2017) aimed at preschool children, enabling the evaluation of children's philosophical questioning on critical thinking, according to teachers' views. In addition, another measurement tool developed to determine the critical thinking skills of four-five year-old children has been encountered within the scope of a thesis written abroad (Chandra, 2008).

Paul and Nosich (1991) state that although it has some limitations, it is possible to evaluate critical thinking via multiple-choice tests, just as recognizing and defining basic skills. Additionally, they state that it is necessary to ask essential questions so that children can reveal the author's purpose in a text, determine unjust inferences that are probably accurate and have

not been proven adequately, discern evidences from hypotheses and results and evaluate evidences that are explained to be reliable as probably reliable, probably unreliable and reliable. From this point of view, this study aimed to develop the 'Critical Thinking Skills Test for 5-6-Year-Old Children (CTTC)' to determine the critical thinking skills of children as a result of applications directly conducted with them.

METHODS

The study was conducted as a survey design in the quantitative model. In this design, it is aimed to describe quantitatively attitudes, tendencies or views toward the population via researches that are performed on a sample group selected (Creswell, 2014).

1. Sample Group

The children who participated in the study were selected using the purposeful sampling method. The sample group of the validity and reliability studies of the CTTC consisted of 202 children in the 'age group of 5-6 years', attending nursery classes, kindergartens and first grade primary education within the body of a primary school affiliated with the Ministry of National Education (MNE) in the City Center of Kırklareli during the school year of 2017-2018 and showing a normal development. After obtaining permissions from the ethics committee and the Provincial Directorate for National Education; the schools where the application was planned to be performed were visited, the school administrators were met and appropriate dates and hours were established for the application. The sample group consisted of children from nine different schools (two kindergartens and seven primary schools). In the study, there were 16 schools attended by 5-6 year-old children who could participate from the City Center of Kırklareli; however, only nine schools were included in the study due to similar socioeconomic characteristics. Table 1 shows distribution of children who participated in the validity and reliability studies of the CTTC according to gender and age.

Table 1. *Distribution of children who participated in the validity and reliability studies of the CTTC according to gender and age*

	N	%
Gender		
Female	106	52.5
Male	96	47.5
Total	202	100
Age		
5 Years (48-60 months)	108	53.5
6 Years (60-72 months)	94	46.5
Total	202	100

As is seen in Table 1; the application was carried out with 202 children for the validity and reliability of the CTTC. Table 1 shows that 52.5% of the children (106) were girls, 47.5% (96) were boys, 53.5% (108) were 5 years old and 46.5% (94) were 6 years old.

2. Development Process of the Measurement Tool

In order to develop the CTTC, the references related with critical thinking (Facione, 1990; Facione, 1998) were used first and then the theoretical structure was examined and thinking skills of the children were evaluated. Stories and questions to be included in the test were prepared according to subscales of the critical thinking skill. These subscales consist of central/self-critical thinking skills including "interpretation, explanation, evaluation, inference, analysis and self-regulation" skills, which were revealed in the Delphi Project carried out by experts from the fields of philosophy, social sciences, education and physics under the leadership of Facione (1990), to provide a conceptual unity for critical thinking skills (Figure 1).

Questions in the CTTC were formed on the basis of goals and acquisitions included in the MNE Preschool Education Program for 36-72 Month-Old Children, and it was taken into consideration that the questions would be able to measure interpretation, explanation, evaluation, inference, analysis and self-regulation skills, which were determined as the subscales of critical thinking. Before forming the items/questions of the CTTC, opinions were taken from three academicians (two from the field of child development and one from the field of basic education), a preschool teacher and two experts in the area of assessment and evaluation. The experts whose opinions were taken were consulted on the questions to be written for the test, the answers to be obtained from children and the scoring of the answers. In accordance with the opinions taken from the experts, it was decided that all the questions in the CTTC would be open-ended as much as possible and the CTTC grading key would be formed after the preliminary study to be conducted with children.

Stories to be used in the CTTC were prepared by the researcher according to the subscales of critical thinking skills and then a CTTC question pool was created. Six stories were written by the researcher for the test as open-ended without presenting the answers directly in a way to encourage children to think and according to their attention span, upon literature review on the subject. While reading the stories to children, animal characters were used to make the process more fun and enable children to identify with characters. In addition, the application duration was considered for the attention span of children and unnecessary details were not included. The ends of all stories were written without an ultimate conclusion, in order not to direct and affect the answers of children.

As children could not remember the stories, sample drawings not indicating the answers directly but reminding of the story process were made (to be used while story reading and answering of the questions by children) for each story, so that children would not make mistakes. A graphic designer drew the drawings. As a result of the study, 3 drawings were prepared for the first story (regarding the interpretation, explanation and evaluation subscales); 3 drawings for the second story (regarding the inference and analysis subscales) and one drawing for each of the third, fourth, fifth and sixth stories (regarding the self-regulation subscale) (Figure 2). The drawings were prepared in the digital media. While performing the drawings, their colorful printouts were taken and they were reinforced with PVC. A preliminary study was conducted with 10 children, in order to examine whether the stories and questions were comprehensible or not. During the preliminary study, comprehensibility of the stories and questions for children was examined, unclear statements were corrected and expressions in the sentences were simplified. Throughout the application, all the answers given by children to the questions in the CTTC were recorded. The answers recorded were later used in the score coding of the test.

FIGURE 2. *Some drawings for the CTTC*

CTTC consists of 41 questions. The test consists of six subscales as interpretation, explanation, evaluation, inference, analysis, and self-regulation. These six subscales can be evaluated independently within themselves and also, the child's critical thinking skills score can be calculated with the total score obtained from sub-skills. The maximum score to be obtained from each question varies according to their content. In the grading key prepared for the practitioner, the maximum score to be obtained from each question is specified. For example; in the sixth question regarding the interpretation subscale of the test, the child is asked to find a

title for the story. If she/he gives a meaningless or irrelevant title (such as Bunny and Wolf), she/he will get 0 point; if she/he gives an indirectly relevant title that does not establish a direct connection (such as Apple), she/he will get 1 point; and if she/he gives a title directly related with the story (such as Filo's Escape from the Circus), she/he will get 2 points. In the CTTC; score range is between 0-9 points in the interpretation subscale, 0-7 points in the explanation subscale, 0-5 points in the evaluation subscale, 0-6 points in the inference subscale, 0-6 points in the analysis subscale, 0-20 points in the self-regulation subscale and 0-53 points in the overall CTTC. In the CTTC, it is accepted that as the child's score increases, her/his critical thinking skills increase.

3. Data Collection

The preliminary study was conducted with 10 children between 27.11.2017 and 01.12.2017in order to determine the answers to be accepted as correct in the grading key of the test and to test comprehensibility of the questions. Following the preliminary study, changes were made in the ordering of the elements of some sentences, for children to understand the questions in the CTTC better. For example, the question "What did the dog do when he realized that him and Filo would not reach out for the apples?" was corrected as "What did the dog which thought that him and Filo would not reach out for the apples, do?". Two researchers collected the data concerning the validity and reliability of the CTTC from 202 children between 11.12.2017 and 17.01.2018. The second practitioner was a child development specialist doing a doctorate in the field of child development. The researcher expressed implementation of the CTTC to the expert and the expert observed the researcher's three applications. Then three applications performed by the expert were also observed by the researcher and the expert worked with children individually during the following applications, saving time for the application. The children, who would participate in the application were met and informed before the application. Children who did not agree to participate in the study were not included in the sample group. The application was carried out in the music room, counseling room or the support education class, which are independent from the classroom. The application was performed on children individually. It took approximately 25-30 minutes for each child.

Ethical Principles in the Data Collection Process

The studies related with the application were started after obtaining an ethics committee approval and an MNE permission. Principals and classroom teachers of the schools were met regarding the place and time of the application. The researcher introduced herself to children before the application, informed them about the application and indicated that they were free to not participate in the application if they did not want or could return to their classes during the application. The data of each child who participated in the study were kept confidential.

4. Data Analysis

The data were analyzed using the statistical package program. Within the scope of the validity study of the CTTC; face validity, confirmatory factor analysis (CFA), content validity, and criterion validity were performed. In the face validity; content validity ratio and content validity index were determined. In the confirmatory factor analysis (CFA); model fit of the test was tested. In the content validity; correlations related with the subscales were calculated. In the criterion validity, correlations between total scores of the Scale for Problem Solving Skills (PSSS), developed by Oğuz and Köksal Akyol (2015) and the CTTC were examined. For the reliability, the test-retest reliability was applied.

The Problem Solving Skills Scale (PSSS)

The Problem Solving Skills Scale (PSSS) was developed by Oğuz and Köksal Akyol (2015) to evaluate the problem solving skills of 60-72 month-old children. The scale, consisting of 18 problem situations and 18 drawings is a five-point likert scale. Each problem situation is rated between 0-4 and the scores to be obtained from the test range between 0-72. The content

validity index was found to be 0.99 for convenience level of the items and 0.96 for convenience level of the items for drawings. As a result of the exploratory factor analysis, it was found that the scale had a single-factorial structure. It was determined that while the Cronbach's Alpha reliability coefficient of the scale was .86; the test-retest correlation coefficient was .60. According to these results, it is possible to state that the PSSS is a valid and reliable measurement tool.

RESULTS

In this section, the results concerning the validity and reliability are given separately.

1. Results Concerning the Validity

In the study; face validity, confirmatory factor analysis (CFA), content validity and criterion validity were performed within the scope of the validity of the test.

Face Validity

Opinions were taken from an expert in the field of children's literature for six stories prepared for the CTTC. After making necessary corrections on the stories based on the opinion of the expert, the test questions prepared for the CTTC were presented to eight academicians (six from the field of child development, one from the field of preschool education and one from the field of psychology) and to an expert preschool teacher. Following the experts' evaluations, the Lawshe content validity was applied to the CTTC. Face validity is used in transforming a qualitative study which is conducted on the basis of expert opinions into a quantitative study (Yurdagül, 2005). The CTTC form was sent to the experts by the researcher with the options of "appropriate, inappropriate, and explanation" across each question in the electronic media, in order to indicate the expert opinions. While evaluating the expert opinions, content validity ratio (CVR) was calculated for each question in the CTTC. While the questions whose CVR value was 0 or negative were eliminated; the CVR was averaged for the questions whose CVR value was positive and the content validity index (CVI) was found. This index is a method used by the experts in determining whether each item is necessary or not. As opinions were taken from nine experts, it can be concluded that the scale whose content validity index (CVI) was greater than 0.75 provided face validity. If there are more than one dimensions in the characteristic to be measured, CVI should be calculated for each dimension (Yurdugül 2005). When calculating the content validity index (CVI) values; the CVI concerning the convenience level of the questions was found to be 0.80 for the interpretation subscale, 0.77 for the explanation subscale, 0.77 for the evaluation subscale, 0.81 for the inference subscale, 0.77 for the analysis subscale, and 0.98 for the self-regulation subscale. It can be asserted that these results indicated that the test was statistically significant. Thus, it was determined that number of questions was 6 for **interpretation**, 6 for **explanation**, 5 for **evaluation**, 6 for **inference**, 6 for **analysis**, 12 for **self-regulation** and 41 in total. During the implementation of the CTTC, the researcher read the stories to children and then a graphic designer prepared 10 drawings not indicating the answers directly but reminding of the story process, in order to prevent the possibility of children to forget the stories in the test while answering the questions. Opinions were received from three experts for convenience of the drawings for the test. In accordance with expert opinions, corrections were made in some drawings (such as fixing the bear's pants, adding a tail to the bunnies, and changing the ground color of two drawings).

Confirmatory Factor Analysis (CFA)

In the confirmatory factor analysis which was performed to examine the model fit of the CTTC, the status of the data to support the six-dimension factor structure revealed in the original scale was examined. Table 2 shows the goodness of fit indices found.

When examining Table 2, as the chi-square test was affected by sample size, the X^2/df (7.48/4=1.87) ratio was used to decide on the model data fit. One of the indicators of goodness of fit is that this value is less than 5. While evaluating goodness of fit; values such as Root Mean Square Error of Approximation (RMSEA), Goodness of Fit Index (GFI), Adjusted Goodness of Fit

Index (AGFI), Normed Fit Index (NFI) and Comparative Fit Index (CFI) can be used. Among these values, if the RMSEA index is 0,08 or below and other indices are higher than .90 and close to 1, it signifies that the fit is good as well. In other words, the difference between population covariance matrix and covariance matrix produced is close to each other (Hu & Bentler, 1999; Hooper, Coughlan & Mullen, 2008; Osterlind, Tabachnich & Fidell, 2001;). In this study, as indicated by Hu and Bentler (1998); precise (X^2, RMSEA, GFI, AGFI) and comparative (NFI, CFI) fit indices were used together. When examining the CFA results in Table 2, it was seen that the X^2/df (1,19) value in the model was less than 5, the GFI, AGFI, NFI and CFI values were higher than .90 and the RMSEA value was below 0.08. Thus, it is possible to state that the CTTC was confirmed in terms of its subscales (interpretation, explanation, evaluation, inference, analysis and self-regulation). Figure 3 shows standard solutions for the six-dimension model of the CTTC.

Table 2. *Goodness of fit indices of the model*

χ² (sd)	χ² / sd	RMSEA	GFI	AGFI	NFI	CFI
10.74 (9)*	1.19	0.031	0.98	0.95	0.92	0.99

*p=0.64,
p>0.05
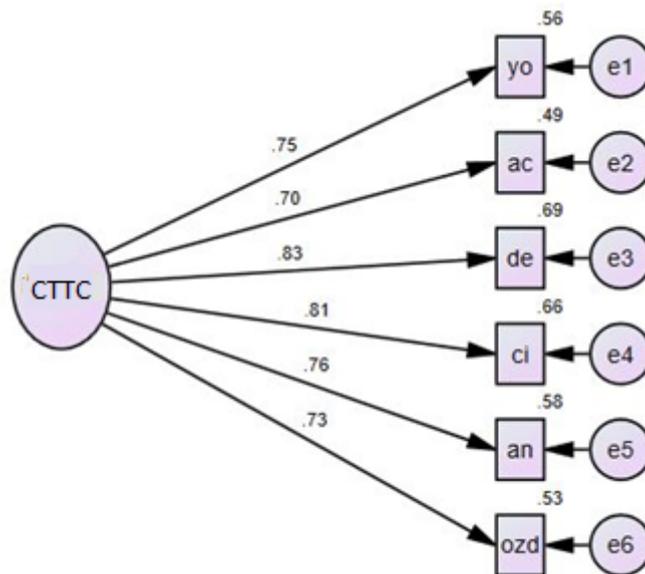
FIGURE 3. *Standard solutions for the six-dimension model of the CTTC*

When examining Figure 2, it is possible to state that t values of the CTTC regarding the interpretation (int), explanation (ex), evaluation (ev), inference (in), analysis (an), and self-regulation (sr) subscales were significant and greater than 0,30, whereas the error values were smaller than 0.90, making it appropriate for these subscales to be included in the model. Considering the results acquired; the CTTC, the latent variable, was significantly predicted by the interpretation (int), explanation (ex), evaluation (ev), inference (in), analysis (an) and self-regulation (sr) subscales of the test.

Content Validity

Table 3 shows correlations regarding the subscales which were calculated to determine content validity for the purpose of evaluating whether the CTTC was fit for purpose or not. As the scores obtained from the explanation, inference, analysis and self-regulation subscales were normally distributed, the Pearson's correlation coefficient was used for the pairwise correlation calculations between these subscales. On the other hand; as the scores obtained from the

interpretation and evaluation subscales were not normally distributed, the Spearman's correlation coefficient was used between these subscales and other subscales.

Table 3. *Pearson's and Spearman's correlation coefficients regarding the subscales of the CTTC*

Subscales	Explanation	Inference	Self-Regulation	Analysis	Interpretation	Evaluation
Explanation	1					
Inference	0.55**	1				
Self-Regulation	0.53**	0.64	1			
Analysis	0.50**	0.68**	0.59**	1		
Interpretation	0.50**	0.61**	0.55**	0.58**	1	
Evaluation	0.32**	0.25**	0.29**	0.26**	0.38**	1

As seen in Table 3, it was determined that correlation coefficients between the subscales of the CTTC were significant. Correlations between the explanation, inference, analysis, self-regulation and interpretation subscales were moderate and positive; whereas, the correlations between the self-regulation subscale and other subscales were weak and positive.

Criterion Validity

One of the ways of providing the validity for the measurement tool is to examine the correlation between the test developed and similar measurement tools (Başol and Gencel, 2013). As critical thinking skills and problem solving skills are indicated to be correlated skills in the literature, the Problem Solving Skills Scale (PSSS), developed by Oğuz and Köksal Akyol (2015) to evaluate the problem solving skills of children, was used here to provide criterion validity. The correlations between total scores of the PSSS and the CTTC were examined. The data concerning the PSSS were collected from 40 children who participated in the study and were selected randomly. While performing correlation studies, it is recommended that the number of sample group should not go down below 30 in providing the external validity when considering that significance of the coefficient is affected by sample size (Büyüköztürk, 2018). As skewness and kurtosis values of the scores concerning the data were not between -1 and +1, they were not normally distributed. Thus, the correlations between total scores of the CTTC and the PSSS were calculated with the Spearman's correlation coefficient. Table 4 shows the results.

Table 4. *Correlations between total scores of the CTTC and the PSSS: The spearman's correlation coefficient*

	PSSS	p
CTTC	0.66**	0.00

When examining the correlation coefficient between total scores of the CTTC and the PSSS in Table 4, it was determined that the Spearman's correlation coefficient was over the moderate level and positive ($r=0.66$; $p<0,05$) and was also statistically significant. Thus, it was confirmed that these two measurement tools measuring similar structures, indeed, measured the two very close structures. The results pointed out the validity of the CTTC based on similar scales.

2. Results Concerning the Reliability

The test-retest reliability was used to test the reliability of the CTTC. In the test-retest method, if the same test is reapplied in the same sample group after a certain time, it is stated that the test is reliable at the similarity rate of the scores obtained by participants from the test (Şeker and Gençdoğan, 2014).

Test-Retest Reliability

The test-retest method was used to reveal whether the test performed a determined measurement in the course of time or not and the CTTC was reapplied to 40 children who were

selected randomly from 202 children in the sample group, one month after the first application. The Pearson's correlation coefficient was calculated for the scores obtained from the two applications. Table 5 shows results concerning the test-retest reliability coefficients regarding the scores obtained.

Table 5. *The Pearson's correlation coefficient for the scores obtained from the CTTC*

	Interpretation	Explanation	Evaluation	Inference	Analysis	Self-Regulation	Crit. Thin. Sk. Total Score
Test-Retest Reliability Coefficient	0.71	0.78	0.74	0.89	0.79	0.68	0.82

As seen in Table 5, it was determined that correlation coefficients regarding the subscales of the test were significant. If correlation coefficients are close to 1, this is accepted for an indicator of the fact that it is a reliable measurement tool.

DISCUSSION and CONCLUSIONS

In this study, it was aimed to develop the 'Critical Thinking Skills Test for 5-6-Year-Old Children (CTTC)', which enables to determine the critical thinking skills of the children. Before developing the CTTC, a literature review was conducted on critical thinking and principles, acquisitions and indicators in the MNE Preschool Education Program were examined. The application was carried out with a total of 202 children (106 girls, 96 boys, 108 aged five, 94 aged six), for the validity and reliability of the CTTC. The CTTC was developed on the basis of the stories and questions in the test and subscales of the critical thinking skills. These subscales consisted of "interpretation, explanation, evaluation, inference, analysis and self-regulation" skills revealed in the Delphi Project under the leadership of Facione (1990). First, stories were created by the researcher for the CTTC and then the questions / item pool were created for the stories. Five stories (1st story for the interpretation, explanation and evaluation skills; 2nd story for the inference and analysis skills; 3rd, 4th, 5th, and 6th stories for the self-regulation skill) were written in accordance with the development of 5-6-year-old children. Number of items formed for the subscales of the test was found to be 6 for interpretation, 6 for explanation, 5 for evaluation, 6 for inference, 6 for analysis, 12 for self-regulation and 41 in total.

Upon calculating the content validity index (CVI) values for content validity of the CTTC; the CVI concerning the convenience level of the items was found to be 0.80 for the interpretation subscale, 0.77 for the explanation subscale, 0.77 for the evaluation subscale, 0.81 for the inference subscale, 0.77 for the analysis subscale, and 0.98 for the self-regulation subscale. Construct validity of the test was performed with the confirmatory factor analysis (CFA). According to the CFA results; it was seen that the X^2/df (1,19) value in the model was less than 5, the GFI, AGFI, NFI and CFI values were higher than .90 and the RMSEA value was below 0.08. It was seen that the Pearson's and Spearman's correlation coefficients regarding the subscales of the test, which was performed for content validity of the CTTC, were significant. It was seen that correlations between the explanation, inference, analysis, self-regulation and interpretation subscales of the test were moderate and positive; whereas correlations between the self-regulation subscale and the other subscales were weak and positive. Correlation coefficient between total scores of the CTTC and the Problem Solving Skills Scale (PSSS) was examined for criterion validity and it was determined that the Spearman's correlation coefficient was over the moderate level and positive (r=0.66; p<0.05) and was also statistically significant. As a result of the test-retest reliability which was performed for the reliability of the CTTC, it was found that correlation coefficients regarding the subscales of the test were significant.

Critical thinking is a type of thinking that is becoming more and more important to examine, due to our era. Preschool period is a critical period in the development of critical thinking skills (Matthews, 2000; McCall, 2017). Critical thinking skills do not appear automatically or suddenly in adulthood and they develop under the influence of maturation and

environment as from early childhood. As the CTTC enables the evaluation of the critical thinking skills of children, it is predicted that reviews on this subject and studies on supporting thinking skills will increase.

As a consequence, it is possible to state that the "Critical Thinking Skills Test for 5-6-Year-Old Children (CTTC)", which consists of the 'interpretation, explanation, evaluation, inference, analysis and self-regulation' subscales, is a valid and reliable measurement tool. The researchers can use the CTTC to determine the critical thinking skills of 5-6 year-old children in descriptive and experimental studies. It is recommended to conduct validity and reliability studies of the CTTC for older age groups.

REFERENCES

Başol, G., & Gencel, İ. E. (2013). Yansıtıcı düşünme düzeyini belirleme ölçeği: Geçerlik ve güvenirlik çalışması. *Kuram ve Uygulamada Eğitim Bilimleri,* 13(2), 929-946.

Bruning, R., Schraw, G., & Ronning, G. G. (1995). *Cognitive psychology and instruction*. New Jersey: Prentice-Hall Inc.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. & Demirel, F. (2018). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Akademi Yayıncılık.

Chandra, J. S. (2008). *A Vygotskykian perspective on promoting critical thinking in young children through mother-child interaction*. (Doktora tezi). www. proquest. com sayfasından erişilmiştir.

Creswell, J. W. (2014). *Araştırma deseni*. (Selçuk Beşir Demir, Çev. Ed.), Ankara: Eğiten Kitap.

Çıkrıkçı-Demirtaşlı, N. (1996). *Eleştirel düşünme: Bir ölçme aracı bir araştırma*. III. Ulusal Psikolojik Danışma ve Rehberlik Kongresi Kongre Kitapçığı.

Daniel, M., & Auriac, M. (2011). Philosopy, critical thinking and philolophy for children. *Educational Philosophy and Theory*, 43(5), 415-435.

Davis-Seaver, J., Smith, T. & Leflore, D. (2003). Constructivism: A path to critical thinking in early childhood. *International Journal of Scholarly Academic Intellectual Diversity*, 7(1), 1-7.

Demir, M. K. (2006). *İlköğretim dördüncü ve beşinci sınıf öğrencilerinin sosyal bilgiler derslerinde eleştirel düşünme düzeylerinin çeşitli değişkenler açısından incelenmesi*. (Doktora tezi). http://tez.yok.gov.tr sayfasından erişilmiştir.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process.* Boston: MA: Heath.

Epstein, A. S. (1993). Training for quality: Improving early childhood programs through.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assesment and instruction-executive summary-The Delphi Report*. www. researchgate.net sayfasından erişilmiştir.

Facione, P. A. (1998). *Critical thinking: What it is and why it counts*. California: California Academic Press.

Gray, A. (2016). *The 10 skills you need to thrive in the Fourth Industrial Revolution*. https://www.weforum.org/agenda/2016/01/the-10-skills-you-need-to-thrive-in-the-fourth-industrial-revolution sayfasından erişilmiştir.

Greenberg, J. (2016). *Teaching children to think: Meeting the demands of the 21st century.* Retrieved from http://www.hanen.org

Gündoğdu, H. (2009). Eleştirel düşünme ve eleştirel düşünme öğretimine dair bazı yanılgılar. *Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü Dergisi,* 7(1), 57-74.

Halpern, D. F. (2003). *Thoughts and knowledge: an ıntroduction to critical thinking*. New Jersey-London: Lawrence Erlbaum Associates.

Heyman, G. D. (2008). Children's critical thinking when learning from others. *Current Directions in Psychological Science*, 17(5), 344-347.

Hooper, D., Coughlan, J., & Mullen, M. (2008). *Structural equation modelling: Guidelines for determining model* fit. https://arrow.dit.ie/cgi/viewcontent.cgi sayfasından erişilmiştir.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424.

Huitt, W. (1998). *Critical thinking: An overwiew. Educational Psychology Interactive,* 3, http://www. edpsycinteractive.org/topics/cognition/critthnk.html sayfasından erişilmiştir.

İskifoğlu, G., & Ağazade, A. S. (2013). Translation and validation of a Turkish version of the California Critical Thinking Disposition Inventory. *Social Behavior and Personality: an International Journal,* 41(2), 187-196.

Karadağ, F., Demirtaş, Y. V., & Yıldız, T. (2017). 5-6 yaş çocuklar için felsefi sorgulama yoluyla eleştirel düşünmenin değerlendirilmesi ölçeğinin geliştirilmesi. *International Online Journal of Educational Sciences,* 9(4), 1025-1037.

Kaytez, N. (2015). *Beş yaş çocuklarının yaratıcılıklarına 'Scamper Eğitim Programı'nın' etkisinin incelenmesi.* (Doktora tezi). http://tez.yok.gov.tr sayfasından erişilmiştir.

Koç, C. (2007). *Aktif öğrenmenin okuduğunu anlama, eleştirel düşünme ve sınıf içi etkileşim üzerindeki etkileri.* (Doktora tezi). http://tez.yok.gov.tr sayfasından erişilmiştir.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, 28(2), 16-26.

Küçükali, R., & Akbaş, H. (2015). Eleştirel düşüncenin bilimlerin gelişmesine katkıları. *Journal of Graduate School of Social Sciences,* 19(3), 1-10.

León, J. M. (2015). A baseline study of strategies to promote critical thinking in the preschool classroom (Un Estudio de Base sobre Estrategias para la Promoción de Pensamiento Critico en las Aulas de Preescolar). *GIST Education And Learning Research Journal,* (10), 113-127.

Matthews, G. B. (2000). *Çocukluk felsefesi.* İstanbul: Gendaş Kültür.

McCall, C. C. (2017). *Düşünmeyi dönüştürmek.* (Çev. Kurtul Gülenç ve Nihal P. Boyacı). Ankara: Nobel Akademik Yayıncılık

Oğuz, V., & Köksal Akyol, A. (2015). Problem Çözme Becerisi Ölçeği (PÇBÖ) geçerlik ve güvenirlik çalışması. *Çukurova University Faculty of Education Journal*, 44(1), 105-122.

Osterlind, S. J., Tabachnick, B. G., & Fidell, L. S. (2001). *SPSS for window workbook to acompany: using multivariate statistics.*--4th.--Tabachnick and Fidell. Allyn and Bacon.

Ömeroğlu, E., Büyüköztürk, S., Aydogan, Y., & Özyürek, A. (2009). Development of a Problem Solving Scale for Children Attending Class 1—5 of Primary Education and Norm Study for Turkey. The *International Journal of Learning*, 1(1), 739-742.

Paul, R. W., & Nosich, G. M. (1991). *A Proposal for the National Assessment of Higher-Order Thinking at the Community College, College, and University Levels*. National Center for Education Statistics (ED), Washington, DC.

Semerci, N. (2000). Kritik Düşünme Ölçeği. *Eğitim ve Bilim*, 25(116), 23-26.

Şeker, H., & Gençdoğan, B. (2014*). Psikolojide ve eğitimde ölçme aracı geliştirme*. Ankara: Nobel.

Yurdagül, H. (2005). *Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması.* XIV. Ulusal Eğitim Bilimleri Kongresinde sunulmuş bildiri, Pamukkale Üniversitesi, Denizli.