Ilkogretim Online - Elementary Education Online, 2020; 19 (4): pp. 2307-2321
http://ilkogretim-online.org.tr
doi:10.17051/ilkonline.2020.764229

Development of analytical thinking tendency scale: Validity and reliability study

Gökhan Aksu, *Adnan Menderes University*, Turkey, *gokhanaksu1983@hotmail.com* ORCID: 0000-0003-2563-6112
Mehmet Taha Eser, *Akdeniz University*, Turkey, *tahaeser@gmail.com* ORCID: 0000-0001-7031-1953

Abstract. This study aims to develop an analytical thinking tendency scale for university students. In this direction, a draft scale, consisting of 50 items, was applied to 574 university students. The obtained data were subjected to reliability tests, explanatory and confirmatory factor analysis (EFA-CFA). The factor analysis showed that the scale is composed of 2 factors. The percentage of variance explained by the factors was calculated as 43%. The fit indexes of the model obtained via CFA showed that the fit indexes of the two-factor structure are sufficient. Considering the reliability coefficients and AVE values calculated within the scope of the research, it can be said that the measurement results are reliable, and the divergent validity of the measurement results is provided. As a result of analyzes, a 5-score Likert type analytical thinking tendency scale consisting of 19-items was developed. According to the findings, it is determined that the analytical thinking tendency scale is a valid and a reliable assessment tool.

Keywords: Analytical thinking tendency, scale development, reliability, validity

Received: 06.01.2020	Accepted: 14.05.2020	Published: 15.09.2020

INTRODUCTION

Thinking is a human-specific concept that begins with the birth of the individual and can be developed and improved directly or indirectly in the process following the birth. Knowledge acquisition, ability to apply what is learned, analysis, synthesis and evaluation are the skills related to the communication process (Charoenwongsak, 2003; Thaneerananon, Triampo and Nokkaew, 2016). Although most of these skills are acquired spontaneously in childhood, the rate of putting these skills into practice and using them in real life depends largely on the education system, individual efforts and working on the subject (Dewey, 2007). The ability of an individual to overcome real-life problems and develop new approaches to the problems in the process of education, individual efforts and working on the subject is directly related to the concept of "Analytical Thinking" (Cottrell, 2005).

Analytical thinking is structurally based on reasoning and evaluation. All reasoning is of great importance in terms of making generalizations from certain observed phenomena called inductions and reaching certain conclusions from generalizations called deductions (Fathima & Rao, 2008). It can be considered as a design consisting of a mixture of two basic reasoning processes such as induction and deduction to analytical processes. In the analytical thinking process, the objects are first handled separately; then, in order to complete the process related to analytical thinking, the interaction with each other is examined in the sense of unity (Dewey, 2007). With analytical thinking, the individual can separate any subject or problem into subtitles using the deductive method and evaluate each information separately. The purpose of questioning the information by dividing the problem or topics into small subtitles is to analyze / develop the whole based on unity based on the parts (Sternberg, 2006; Art-in, 2012; Yulina et al., 2019).

Analytical thinking is one of the most basic high-level thinking skills serving to the development of 21st century skills, such as critical thinking, problem solving, creativity and decision making. The historical and philosophical foundations of the analytical thinking process are based on Aristotle's works on the disciplines of logic, in which he has set out the principles of systematic thinking (Shields, 2012). Cottrel (2005) stated that analytical thinking is related to the mental processing of the evidences that constitute an advantage or a disadvantage. According to

Sternberg (2006), analytical thinking is closely related to dividing a problem into pieces and giving a meaning to these pieces, evaluating and criticizing the characteristics of a concept. According to R. A. Dewey (2007), analytical thinking is the effective, consistent and careful evaluation of a belief or a knowledge in the light of the basis that supports it and, moreover, by predicting its direction. According to Üstündağ (2011), analytical thinking is a complement of creative thinking. According to Güneş (2011), analytical thinking includes the operations involved in the process of separating the parts that make up the whole from the body, re-defining and classifying these parts.

According to the constructivist approach, one of the main objectives of education is to realize analytical thinking. Analytical thinking is one of the predominant goals of pragmatic philosophy, of which John Dewey is one of the pioneers. According to Dewey, reflection gives meaning to the experience through the reorganization and restructuring of the existing experience and leads to the realization of subsequent experiences. Analytical thinking requires the learner to focus on the process rather than the result. For the realization of analytical thinking, first of all, learning should occur, and the individual should be able to convert what has learned into behavior. After that, the behavior should be evaluated, especially by the learner. It can be thought that an individual who is aware of the path he has taken in the process and can comment on the resulting product will benefit from this experience in his later performances (Dewey, 1933; Gürkaynak, Üstel & Gülgöz, 2008).

Especially in the 21st century, analytical thinking is blended with critical thinking as a part of the problem-solving process, to provide the skills required to prepare children for a more complex living and working environment. Therefore, it is very important to develop analytical thinking skills in addition to applied, conceptual, synthesis, creative and criticism-based thinking skills of individuals (Charoenwongsak, 2003; Chaijaroen, Kanjug and Samat, 2006). Bringing a scale identifying analytical thinking tendencies of university students to the literature was thought to be quite important considering the factors such as the studies aiming to define analytical thinking entirely in a structural sense, analytical thinking being one of the main elements of the constructivist approach and helping the assessment of the teachers for improving intellectual processes of the students, etc. The literature includes *"The Scale of Holistic and Analytical Thinking Styles in Problem Solving*" by Umay and Kızıltuğ (2007). Although there are scales for measuring critical thinking tendencies in the literature, there are almost no scales developed for analyzing analytical thinking tendencies. When the literature on analytical thinking is analyzed, the concept of analytical thinking is considered as a sub-dimension on the scale developed by Pacini and Epstein (1999). This scale, which tries to determine the level of analytical thinking in the context of the sub-dimension, is scored in a 5-point Likert type. The reliability value in terms of internal consistency was calculated as .87. When the scales developed in the literature are examined, there are several scales with analytical thinking styles, but these scales are called as "Analytical and Holistic Thinking Scale". The "Integrated and Analytical Thinking Scale in Problem Solving Scale" which was adapted by Umay and Kızıltuğ (2007) and consists of 5 items in total, is the only scale in the literature. Analytical thinking can be taught within the scope of educational systems and analytical thinking, especially problem solving, is thought to be beneficial for all subjects (Dewey, 2007). Students are directed to analytical thinking at universities (Arıol, 2009). In particular, individuals who are at the university level of education and training are expected to have the skills to solve the problems they encounter using sequential processes (Hammouri, 2003). However, each individual in the holistic way of thinking solves the problem with the help of similar examples of individuals with analytical thinking styles that intuitively and logically break down the problem and solve it by using sequential processes in the educational environment. When designing educational environments, it is necessary to reveal the characteristics of individuals in the education environment in terms of thinking style rather than preparing a program based on a single thinking style or based on a single thinking style (Sternebrg & Grigorenko, 2001; Sternberg & Grigorenko, 2004; Esmer & Altun, 2013). Measurement tools are needed to help determine the trend towards thinking styles such as Analytical Thinking Tendencies developed within the scope of research in terms of obtaining information in terms of thinking styles of individuals in an educational environment, preparing a training program according to the thinking styles of individuals, increasing the quality of education and measuring student achievements correctly. Therefore, bringing a scale determining analytical thinking tendencies of university students to the literature is a necessity. In this context, the research aims to develop the Analytical Thinking Tendency Scale (ATTS).

METHODS

The universe of the study consisted of the students studying at Adnan Menderes University and purposeful sampling was used in the study.

Study Group

The research was based on the answers given by two different study groups of sample sizes 263 and 314, which were studying in Adnan Menderes University Faculty of Education, Faculty of Science and Literature and Faculty of Economics and Administrative Sciences in the Spring Semester of 2016-2017 Academic Year. The first study group consisted of 198 female, 65 male students, whereas the second group consisted of 210 female and 104 male students. In order to increase the external validity of the study, the sample was kept as large as possible since larger sample size increases the power of the model to be tested (Weston and Gore, 2006).

Data Analysis

Within the scope of the research, Exploratory Factor Analysis (EFA) was performed using the data obtained from the first study group, whereas Confirmatory Factor Analysis (CFA), reliability analyzes, and content validity analysis were performed using the data obtained from the second study group. In scale development process, analyzes were started with an item pool consisting of 50 items. The adaptation validity of the developed scale was checked using the scores related to the *Holistic Thinking* subscale of the *Scale of Holistic and Analytical Thinking Styles While Problem Solving* (Umay and Kızıltuğ, 2007). R program was used in the analyses performed within the scope of the research. *Sem, lavaan, onyxR, Performance Analytics, psych, psychometric, semtools, stats, user friendly science* packages were used in the analyzes performed within the scope of R program.

RESULTS

Statistical analyzes were performed to reveal the psychometric properties of the measurements. First of all, EFA was applied for the construct validity of the interpretations made from the measurements. Sample size was the first priority in this study. There is no consensus on the sample size of the study group that should be included in factor analysis studies. Cattell (1978) stated that 200 participants were acceptable in factor analysis studies; Comrey and Lee (1992) stated that the average sample size should be 200 participants; Kline (1994) stated that 200 participants would usually be sufficient. In the estimation of the appropriate sample size for performing a factor analysis, it is recommended to reach a sample size that meets at least two of the criteria present in the literature (Çokluk, Şekercioğlu and Büyüköztürk. 2014). In this study, data from the first study group consisting of 263 participants were used for EFA and data from the second study group obtained from 314 participants were used for CFA. Considering the criteria listed above, it can be said that the number of participants in the study groups is appropriate for factor analysis.

Findings of the Exploratory Factor Analysis

Kaiser-Meyer-Olkin (KMO) index was checked and Bartlett test was performed to determine the suitability of the data for factor analysis in terms of construct validity. The Bartlett's Test of Sphericity is based on the principle of testing the correlation matrix of the variables in the data set against the unit matrix for scale development and consequently gives an indication about the suitability of the data set for EFA. The Kaiser-Meyer-Olkin (KMO) value is a measure of whether the data can be illustrated by factor analytic model (Aksu, Eser and Güzeller, 2017; Büyüköztürk, 2018). As a result of the analyzes, Kaiser-Meyer-Olkin (KMO) value was found to be 0.92; the result of the Bartlett sphericity test was statistically significant ($\chi 2 = 255.41$; sd = 49; p = .000). Kaiser-Meyer-Olkin (KMO) values above 0.90 are interpreted as excellent (Çokluk, Şekercioğlu and Büyüköztürk, 2014; Aksu et al., 2017). In addition, the analysis of KMO value for each variable showed that these values varied between 0.80 and 0.96. It was concluded that factor analysis could be applied to the data set because of the high KMO value of the whole data set; whereas the significance of the Bartlett sphericity test confirmed the assumption that the study group had normal distribution (Aksu et al., 2017).

As a result of the EFA, overlapping problems was observed on the 2nd, 4th, 5th, 6th, 7th, 8th, 10th, 12th, 13th, 15th, 16th, 18th, 19th, 20th, 23rd, 26th, 27th, 30th, 33rd, 34th, 36th, 37th, 42nd, 43rd, 44th, 45th, 48th, 49th and 50th items (For the items to be included in the final form of the scale, the cut-off value of factor load for each item was set as 0.5). A second EFA was carried out by removing these items from the measuring tool initially consisting of 50 items, and at the end of the second EFA, an overlapping problem was observed on items 22 and 32.

A total of 31 items, observed to have overlapping problems in the 2 EFAs, were removed from the scale and a final EFA was performed to determine the number of factors and factor structures of the remaining 19 items. Scree Plot, Horn's Parallel Analysis, Very Simple Structure (VSS) and Velicer's Map Test were used to determine the number of factors. In determining the number of factors, the illustration of Horn's Parallel Analysis was done first.

In recent years, other methods that give accurate results in determining the number of factors such as Parallel Analysis, Smallest Mean Partial Correlation (MAP = Minimum Average Partial) test have been frequently preferred (Fabrigar et al., 1999; Ford, MacCallum, and Tait. 1986; Wang and Weng, 2002; Weng, 1995). In the graphic shown in Figure 1 in the study, both the Basic Components Analysis and Factor Analysis (Basic Axis Factoring) are performed together and the ideal number of factors is determined within the scope of the R program. However, in order to reveal the factor structure in the continuation of the study, the analysis results shown in Table 1 were obtained by the Basic Axis Factoring method. Although it was stated in the literature that factor analysis and principal components analysis were obtained as the reliability level of the measurement tool increased, principal components analysis was not used in order to reveal the factor structure (Thompson, 1992).

Figure 1 shows the Scree Plot for the ideal number of factors. Regarding Figure 1, it is seen that a significant part of the variance is explained by the first factor for both curves based on principal component analysis and factor analysis, and the eigenvalues of the factors following the second factor are close to each other. In other words, after the second factor the contribution of the factors to the variance is minimal and approximately the same. As a result, it can be said that the ideal number of factors is 2 based on Scree Plot.

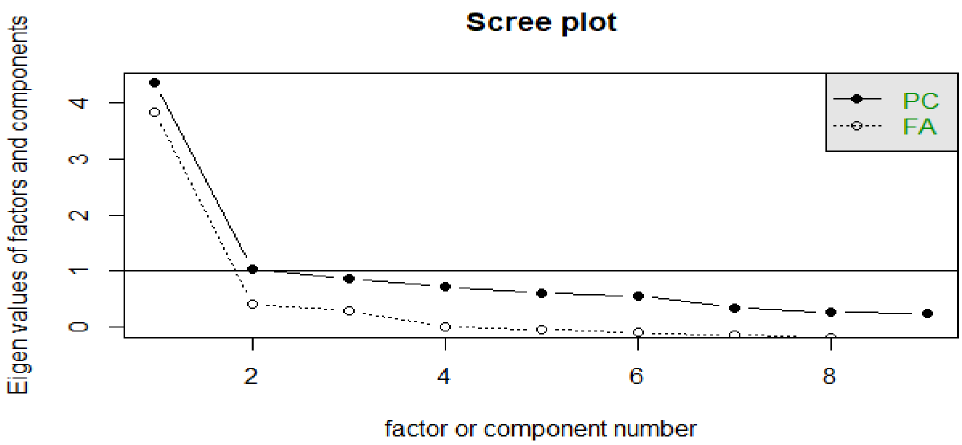

FIGURE 1. *Scree plot*

The working principle of Horn's Parallel Analysis requires a random series of correlation matrix with the same number of participants and variables as the actual data. Principal component analysis is performed on these random correlation matrices and the mean of the eigenvalues is found. The mean eigenvalues calculated using random matrices is compared with eigenvalues calculated using actual data and the ideal number of factors is the point where the eigenvalue calculated by actual data is greater than the eigenvalue calculated by random data (Piccone, 2009). Figure 2 shows the illustration of Horn's parallel analysis for setting the ideal number of factors. The blue line in Figure 2 shows the eigenvalues of the actual data set; the dashed line consisting of red dots shows the eigenvalues of simulated data set; and the dashed red line indicates the eigenvalues of the resampled data set. The point at which a big drop occurred in the actual data set and the point at which the gap between the actual data set and the simulated data set are minimal (i.e., the intersection point) tells us the number of ideal factors. From Figure 2, it can be concluded that the ideal number of factors is either 2 or 3.

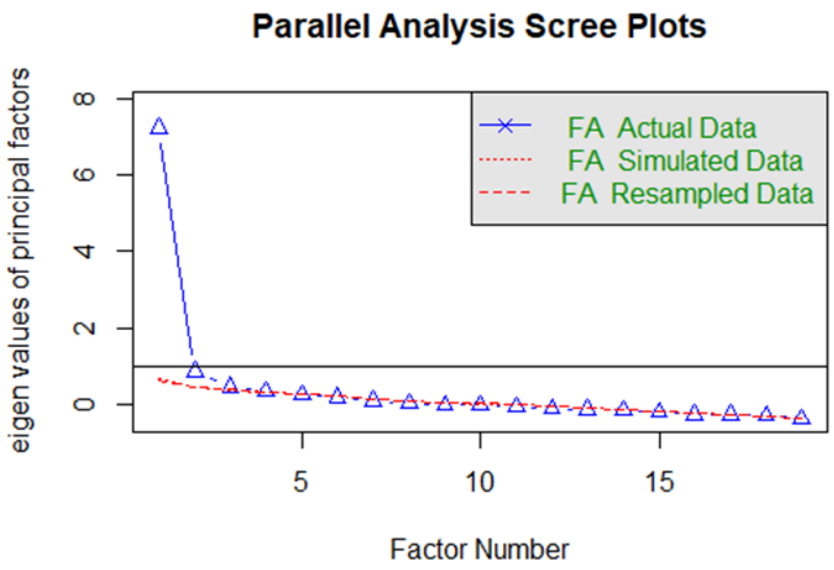

FIGURE 2. *Scree plot of Horn's Parallel Analysis*

In order to reach a clearer result regarding the determination of the ideal number of factors according to Horn's Parallel Analysis, Table 1, which provides information about the number of components with eigenvalues greater than one, should be examined. Table 1 shows the statistical results for Horn's Parallel Analysis. Considering that the eigenvalues of the two components in Table 1 that are greater than 1 (R program gives only the outputs the components whose eigenvalue is greater than one and therefore must be kept), the number of components is two. Therefore, it can be said that the ideal number of factors is 2 according to Horn's Parallel Analysis method.

Table 1. *Eigenvalues from Horn's Parallel Analysis*

Component	Adjusted eigenvalues	Unadjusted eigenvalues	Estimated bias
1	7.1888	7.7102	0.5213
2	1.1142	1.5344	0.4202

Very Simpe Structure (VSS) is another method used in collecting evidence for the validity of ideal number of factors determination. Table 2 shows to the output obtained by the VSS method. Regarding Figure 4, the maximum value of VSS Complexity 1 (0.75) was reached when the number of factors was one; whereas VSS Complexity 2 reached its maximum value (0.85) when the number of factors was two. The analysis output of VSS method also includes indications of the Bayesian Information Criteria. According to the analysis results, the smallest Bayesian Information Criteria (BIC) value was reached when the number of factors was 2. Considering the results of the Very Simple Structure method in determining the ideal number of factors, it can be said that the ideal number of factors is 2.

Table 2. *Results of very simple structure method*

Iteration nember	VSS1 value	VSS2 value
1	0.75	0.00
2	0.66	0.85
3	0.60	0.81
4	0.51	0.76
5	0.62	0.72
6	0.62	0.76
7	0.62	0.76
8	0.62	0.76

In order to obtain a clearer result for the ideal number of factors according to Very Simple Structure method, the output of the number of factors related to model fit in Figure 5 can be examined. Regarding Figure 3, it is seen that the maximum fit value (approximately 0.9) can be achieved by 2 factors.

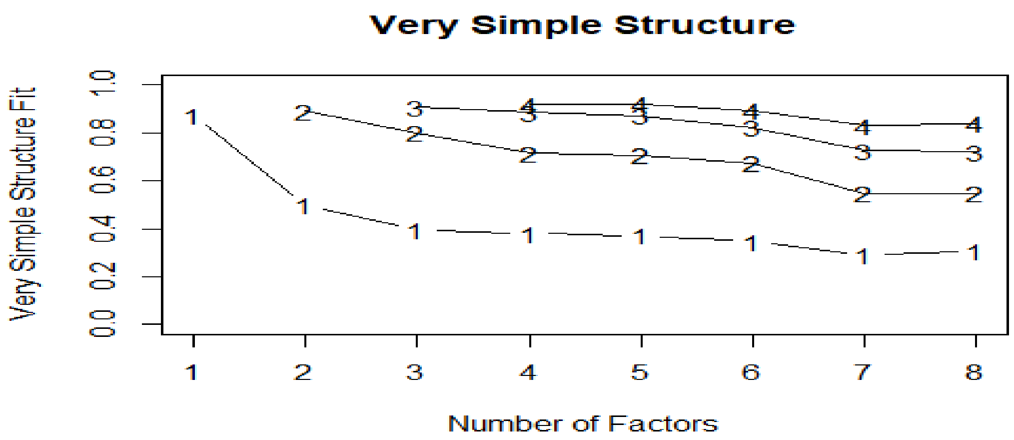

FIGURE 3. *Illustration of very simple structure method II*

As a final step in determining the ideal number of factors, the Eigenvalue-Component Graph in Figure 4, which allows comparison of various methods both visually and numerically, is examined. Regarding Figure 4, although the ideal factor number is found to be 1 according to the acceleration factor, the ideal factor number is 2 considering the eigenvalues, parallel analysis and optimal coordinates. The number of ideal factors according to the acceleration factor was found to be different than the ones found by other methods, which was thought to be related to the fact that the acceleration factor gives better results with large samples and high number of variables (Raiche, Riopel and Blais, 2006).

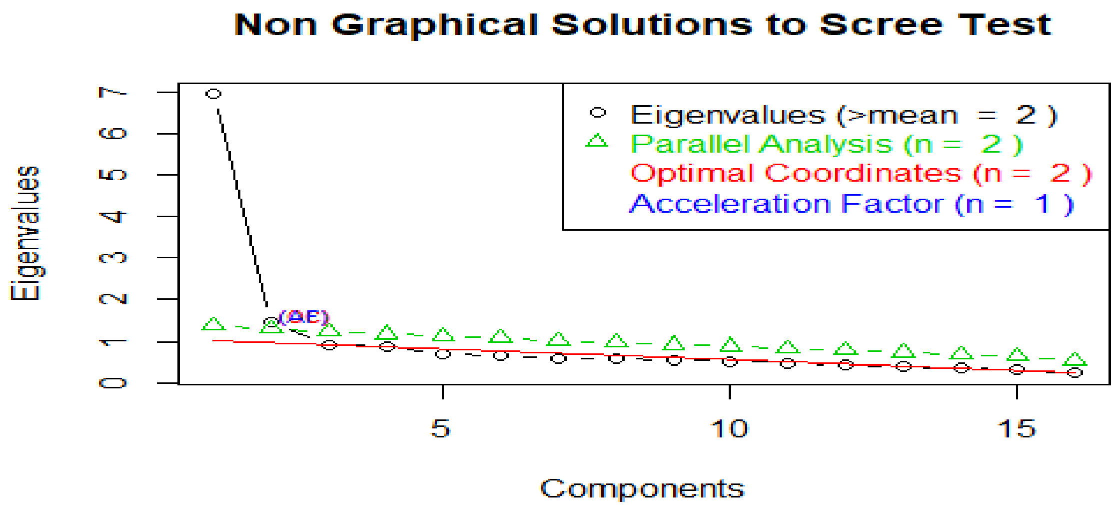

FIGURE 4. *Eigenvalue-component graph*

As a result of the evaluation of the outcomes of all methods used to determine the ideal number of factors in a holistic Figure, it was concluded that the ideal factor number of the scale developed within the study should be 2. The dimension consisting of 6 items was named as *Association Dimension* and the dimension consisting of 19 items was named as *Differentiation Dimension*. In the naming of dimensions, the meanings of the items and the definitions of Chaffee (1990) and Brookhart (2010) given about analytical thinking skill, dividing the whole into pieces and the relationships of the pieces with each other and with the whole, were taken into consideration.

Table 3 shows the factor loads of the items associated with the 2 dimensions, which was set as the ideal number of factors in EFA, the relative variance and the total variance. According to Table 3, factor loads of the items vary between 0.51 and 0.70. For EFA, factor loads should be at least 0.30 (Barnes et al., 2001) and the factors can be interpreted if their loads are 0.32 and above. Factor loads being above 0.71 is considered as excellent, above 0.63 as very good, above 0.55 as good, above 0.45 as sufficient, and above 0.32 as weak (Comrey and Lee, 1992). As mentioned before, the cut-off value of the items' factor load has been set as 0.5, accordingly, the items with a factor load below 0.5 have been removed from the scale. At the stage of deciding the rotation method to be used within the scope of the research, *Promax*, one of the orthogonal rotation methods, was preferred considering that there may be a relationship between the dimensions since the concept of analytical thinking requires association skills and the fact that in reality the unrelatedness of the factors is quite unlikely (Rennie, 1997).

Table 3 shows the factor loads of the items. Considering that the factor loads of the items in Table 1 vary between 0.51 and 0.70, it that the factor loads of the items can be said to be good. Table 3 also contains information about the explained variance. According to Table 3, the total variance explained by 2 factors constituting *Analytical Thinking Tendency Scale* is 0.43. In general, very high amount of explained variance cannot be reached in social sciences due to less information accuracy, thus the percentages varying between 40% and 60% are considered to be sufficient for explained variance (Gorsuch, 1983; Scherer et al., 1988; Dunteman, 1989; Kline, 1994). Accordingly, the total variance explained by the 2 factors constituting the scale was accepted as sufficient.

Table 3. *Results of the exploratory factor analysis (EFA)*

	Factor loads of association dimension	Factor loads of differentiation dimension
Item1		0.658
Item3		0.707
Item14		0.523
Item17		0.515
Item46		0.579
Item47		0.547
Item9	0.537	
Item11	0.534	
Item21	0.521	
Item24	0.539	
Item25	0.582	
Item28	0.634	
Item29	0.570	
Item31	0.653	
Item35	0.672	
Item38	0.596	
Item39	0.552	
Item40	0.680	
Item41	0.593	
SS loadings	5.032	3.224
Relative variance	0.265	0.170
Total variance	0.265	0.435

Findings of the Confirmatory Factor Analysis

CFA was conducted within the scope of the research in order to provide evidence about the construct validity of the data obtained from the scale and to verify the dimension structure of the model. Model fit indices of CFA are given in Table 4. Regarding the fit indices in Table 4, the relative chi-square value (χ^2/sd) is 2.45. χ^2 / df (371.552 / 151) < 2.5 indicates a perfect fit (Kline, 2011). According to Table 2, SRMR was 0.05; TLI was 0.86; CFI was 0.88; GFI was 0.86; AGFI was 0.83. According to the literature, SRMR < 0.08 indicates a good fit (Maydeu-Olivares, Shi and Rosseel, 2017); 0.85 < TLI < 0.9 indicates a good fit (Bentler and Bonett, 1980; Sharma, Mukherjee, Kumar and Dillon, 2005); 0.8 < CFI < 0.9 indicates an acceptable fit (Bentler, 1990); GFI < 0.9 indicates a good fit (Tanaka and Huba, 1985; Sharma et al., 2005); and AGFI ≤ 0.9 indicates an acceptable fit. As a result of the holistic evaluation of the figures in the Table and relevant criteria, it can be said that model-data fit was achieved.

Table 4. *Model fit indices*

χ^2/sd	RMSEA	SRMR	TLI	CFI	GFI	AGFI
2.45	0.07	0.05	0.86	0.88	0.86	0.83

Table 5 shows CFA results. Table 5 includes numerical information about the estimates, standard error, and z values. According to Table 5, estimates were found slightly higher than 1 and the significance values (z values) were statistically significant.

Table 5. *Results of the confirmatory factor analysis*

	Estimate	Std.Err	z-value	P(>|z|)
Association				
Item1	1.000			
Item3	1.026	0.112	9.183	0.000
Item14	0.907	0.110	8.277	0.000
Item17	1.038	0.113	9.176	0.000
Item46	0.992	0.130	7.618	0.000
Item47	0.890	0.118	7.513	0.000
Differentation				
Item9	1.000			
Item11	1.002	0.119	8.394	0.000
Item21	0.805	0.106	7.569	0.000
Item24	0.971	0.111	8.716	0.000
Item25	0.980	0.112	8.727	0.000
Item28	0.926	0.116	8.000	0.000
Item29	0.989	0.124	7.962	0.000
Item31	1.190	0.127	9.372	0.000
Item35	1.034	0.115	8.958	0.000
Item38	0.974	0.107	9.077	0.000
Item39	1.018	0.116	8.795	0.000
Item40	1.165	0.124	9.376	0.000
Item41	0.855	0.102	8.415	0.000

Factor loads of the two-dimensional model obtained by CFA are given in Figure 5. As can be seen, factor loads of association dimension vary between .35 and .68 and factor loads of differentiation dimension vary between .30 and .55.

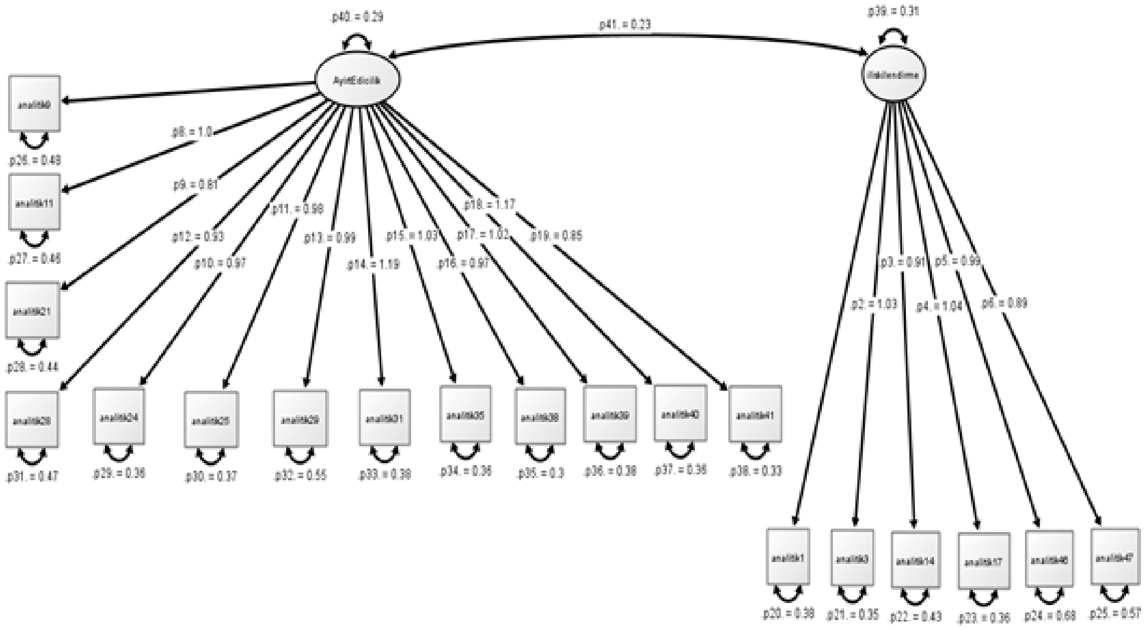

FIGURE 5. *ATTS measurement model*

Findings on Content Validity

In this study, Lawshe method was used in terms of content validity, for which the opinions of 8 different education assessment and evaluation experts were taken. The quality and number of experts (between 5 and 40) involved in the calculations performed to determine the validity of the scope is an important issue for the objectivity of the results (Lawshe, 1975; Veneziano and Hooper, 1997). Under the Lawshe (1975) method, experts were asked to rate each item as "essential", "useful but not essential" or "not necessary". As a result of the ratings, Content Validity Ratio (CVR), which was calculated using the formula and CVR values of the items, were found to be greater than 0. Content Validity Index (CVI) was calculated by taking the mean of the CVR values for all items and this value was found to be (0.87). Content Validity Index, which is a criterion used to determine the statistical significance of the items, was found to be 0.78 for 8 experts, at 0.05 significance level. In the light of these results, the CVI ≥ CVR criterion was validated and the content validity of the scale was found to be statistically significant.

Findings of Concurrent Validity

Within the scope of concurrent validity, the correlations between students' association and differentiation tendencies and holistic thinking were calculated. Figure 6 provides an illustration of the correlation of the scale's subdimensions with holistic thinking and the distribution of variables. Regarding Figure 8, there was a positive and statistically significant correlation between association and holistic thinking (r = 0.71; p <0.001) and a positive and statistically significant relationship between differentiation and holistic thinking (r = 0.60; p <0.001). Considering that analytical thinking and holistic thinking are two complementary concepts (Umay and Arıol, 2011; Brookhart, Nisbett et al., 2001), these correlation values can be taken as an evidence of the achievement of concurrent validity.

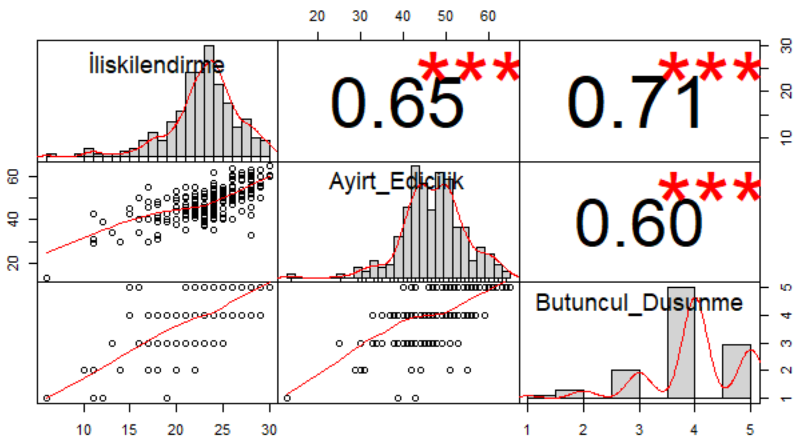

FIGURE 6. *Correlation values for concurrent validity*

Findings on reliability were discussed in the next section of the study.

Findings of Reliability

AVE and structural reliability values obtained from the analysis of the data collected by Analytical Thinking Tendency Scale are given in Table 6.

Cronbach Alpha (α), McDonald Omega Total (ωt), McDonald Omega Hierarchical (ωh), Revelle's Omega, Greatest Lower Bound (GLB), H coefficient were considered while analyzing the reliability in terms of internal consistency of the measurement results of ATTS. Regarding Cronbach Alpha values indicating reliability in terms of internal consistency, Cronbach alpha value was found to be approximately 0.81 for the association dimension; 0.91 for the differentiation dimension; and 0.92 for the whole measurement tool. Cortina (1993) and Schmitt (1996) reported that the Cronbach alpha coefficient above .70 may be sufficient.

Cronbach alpha coefficient is affected by the number of items that constitute the scale. Cronbach's alpha coefficient can give lower results in the scales with less items (Cortina, 1993;

Osburn, 2000). Therefore, it is recommended to avoid using Cronbach's alpha coefficient for the reliability analysis of the scales with factors comprised of low number of items (Erkuş, 1999). At the same time, Cronbach Alpha coefficient assumes that the measurements are parallel, and the reliability of the congeneric measurements is predicted lower than actual. In order to reveal the real reliability value of the related dimensions, the analysis was also performed for McDonald Omega coefficient which is not affected by the number of items. Considering the facts related to Cronbach's alpha and better actual reliability prediction of McDonald's Omega method (Lucke, 2005), McDonald’s Omega (Total) (1999) coefficient, which is defined for congeneric measurements, was also calculated in the study. McDonald Omega (total) coefficient was found to be approximately 0.81 for the association dimension; approximately 0.91 for the differentiation dimension; and 0.92 for the whole scale. McDonald's Omega (total) coefficient estimates the reliability by considering both general and specific factors. In addition to McDonald Omega (total) coefficient, McDonald’s Omega (hierarchical) values, which only consider the general factor of the scale, were calculated (Revelle and Zinbarg, 2009). McDonald Omega (hierarchical) coefficient was found to be approximately 0.75 for the association dimension; approximately 0.86 for the differentiation dimension; and 0.73 for the whole scale. Nunnally and Bernstein (1994) underlined that the reliability coefficient of a scale should be over 0.7 in order to be accepted as reliable. Considering that McDonald Omega coefficients for the whole scale and its sub-dimensions are above 0.7, it can be said that the reliability values of the scale are acceptable.

Apart from Omega coefficients, GLB coefficients also provide more reliable predictions that Cronbach alpha for the whole scale and its sub-dimensions (Dunn, Baguley and Brunsden, 2014; Peters, 2014). GLB coefficient was found to be 0.84 for the association dimension; 0.94 for the differentiation dimension; and 0.96 for the whole scale. As the other reliability coefficients calculated within the scope of the research, it can be said that the values related to GLB coefficients are also acceptable.

Within the scope of the study, Loevinger’s homogeneity coefficient, also known as H coefficient, was calculated for the reliability of the scale and its sub-dimensions. H coefficient was found to be 0.82 for the association dimension of Analytical Thinking Tendency Scale; 0.91 for the differentiation dimension; and 0.92 for the whole scale. H coefficient, which is a general scalability coefficient, of 0.5 and above means that the relevant structure is strong (Sijtsma & Molenaar, 2002). Therefore, it can be said that the structure of the scale is strong.

In the study, Average Variance Extracted (AVE) values were also calculated for the whole scale and its sub-dimensions. AVE value was found to be 0.39 for the association dimension; 0.42 for the differentiation dimension; and 0.41 for the whole scale. In terms of convergent validity, AVE values should be smaller than internal consistency reliability values (Structural Reliability) and each AVE value should be greater than 0.5 (Fornell and Locker, 1981). On the other hand, Fornell and Locker (1981) also stated that for AVE values less than 0.5, convergent validity of the structure is acceptable if the reliability in terms of internal consistency is greater than 0.6. In the light of this information, it can be said that AVE values obtained for association and differentiation dimensions and the whole scale are acceptable.

Table 6. *Reliability coefficients and AVE values of the scale*

	Association	Differentation	Total
Cronbach Alfa	0.81	0.91	0.92
McDonald’s Omega (Total)	0.81	0.91	0.92
McDonald’S Omega (Hierarchical)	0.75	0.86	0.73
Revelle’s Omega	0.92	0.85	0.93
Greatest Lower Bound (GLB)	0.84	0.94	0.96
H Coefficient	0.82	0.91	0.92
AVE	0.39	0.42	0.41

Fornell and Larcker (1981) stated that AVE values larger than shared variance estimates (square of inter-structural correlation) support divergent validity. In other words, the fact that the square roots of AVE values being higher than the correlation coefficients between latent variables is the evidence of discriminant validity. Table 7 includes the relationship coefficients between latent variables and the square roots of AVE values.

Table 7. *Relationship coefficients between latent variables and square roots of AVE values*

Dimension	Association	Differentiation
Association	**0.62***	
Differentiation	0.55	**0.64***

When Table 7 is examined, the diagonal elements of the matrix corresponding to the square roots of AVE values are larger than the non-diagonal elements of the matrix. According to this result, it can be said that divergent validity is supported within the scope of the research.

DISCUSSION AND CONCLUSIONS

In this study, it is aimed to develop a measurement tool that will allow to make valid and reliable measurements for the analytical thinking tendencies of university students. In order to ensure the content and face validity of the scale, expert opinion was taken and a draft measuring tool of 50 items was obtained. The items in the scale was evaluated by the students in the study group by a Likert-type scale of 5 (Always) – 1 Never (1).

EFA and CFA were used to test the construct validity of the interpretations obtained from ATTS measurements. As a result of EFA, a two-factor structure consisting of 19 items explaining 43% of the total variance was obtained. Regarding the meaning of the items grouped under the factors and the theoretical structure, the first factor was named as association and the second factor as differentiation. The validation of the theoretically constructed measurement model by the data was made by CFA. The fit indexes of the model obtained via CFA showed that the fit indexes of the two-factor structure are sufficient.

Considering that being above 40% was taken as the criterion for the explained variance in EFA (Dunteman, 1989; Kline, 1994), factor loadings of the scale items were over .45, which has been set as the lower limit for sufficiency (Comrey and Lee, 1992), and the fit indexes calculated in the CFA were within acceptable limits; therefore it can be said that the construct validity of the measurements obtained from ATTS is provided.

Cronbach Alpha, McDonald Omega Total, McDonald Omega Hierarchical, Revelle's Omega, Greatest Lower Bound, and H coefficient were analyzed for the reliability of the measurements obtained from ATTS; moreover, AVE values were analyzed for its validity. Usually, measurements with a reliability coefficient of .70 and above are considered to be reliable (Fornell and Larcker, 1981; Nunnaly and Bernstein, 1994). Considering the reliability coefficients and AVE values calculated within the scope of the research, it can be said that the measurement results are reliable, and the divergent validity of the measurement results is provided. Lawshe method was used for content validity of the measurements obtained from ATTS; CVI ≥ CVR criterion vas validated and the content validity of the scale was found to be statistically significant. The findings of the study suggest that ATTS can be used as a tool that produces valid and reliable measurements in determining analytical thinking tendencies of university students.

The literature review revealed the lack of a measurement tool that can be used to measure analytical thinking tendencies in international studies. However, a scale that could be used for this purpose was found in the Turkish literature. This study, which aims to develop ATTS, is thought to be important in terms of filling this gap in the literature. In other words, helping to identify trends in analytical thinking is one of the strengths of this study.

Providing more than one evidence for the construct validity and reliability of the measurements obtained from ATTS is another strength of the study. For example, Cronbach's

Alpha, McDonald Omega Total, McDonald Omega Hierarchical, Revelle's Omega, Greatest Lower Bound and H coefficients were used for the reliability of the measurement results. In terms of construct validity, both EFA and CFA were applied. Divergent validity and content validity were also tested within the scope of the study.

Using the analytical thinking scale, which has a structure that examines metacognitive levels, information about the analytical thinking styles of individuals in the educational environment can be obtained. Considering that the separation of the concepts into pieces and the evaluation of the separation of the separated pieces is the basis of analytical thinking, a scale developed for such a structure can also be seen as a major importance in terms of decomposition and association of two sub-dimensions that can be considered as association and its importance for the structure. Using the scale obtained as a result of the research, obtaining information about the analytical thinking styles of the individuals in the educational environment, creating the education program by considering the profile of each individual's thinking style, developing the analytical thinking tendencies of the individuals and analytic thinking tendency of each individual in the educational environment exams can be created by keeping in front of them.

Beside above-mentioned strengths of the study, there are some limitations, which bring various suggestions for future research. First, this study was conducted on a study group comprised of university students only. Considering that reliability is a concept related to the measurements and validity is a concept related to the appropriateness of interpretations made through measurements, validity and reliability studies to be carried out through the data collected from different samples are of great importance.

Another suggestion for future research may be about the analysis of the concurrent validity of ATTS. Within the scope of the study, the relationship between the scores obtained from the association and differentiation subscales and holistic thinking was analyzed regarding the concurrent validity. Departing from this point, it may be suggested to carry out studies trying to determine the relationship between the sub-dimensions of analytical thinking and holistic thinking or similar variables. The realization of the studies in which ATTS will be used is of great importance in terms of contributing to the measurement power of the developed scale.

REFERENCES

Aksu, G., Eser, M. T., Güzeller, C. O. (2017). *Açımlayıcı ve doğrulayıcı faktör analizi ile yapısal eşitlik modeli uygulamaları*. [Structural equation modeling applications with exploratory and confirmatory factor analysis]. Ankara: Detay Yayıncılık.

Art-in, S. (2012). Development of Teachers' Learning Management Emphasizing on Analytical Thinking in Thailand.. *Procedia - Social and Behavioral Sciences*. 46 (3339-3344).

Barnes, J., Cote, J., Cudeck, R. & Malthouse, E. (2001). Factor analysis—checking assumptions of normality before conducting factor analysis. *Journal of Consumer Psychology* 10(1,2), 79–81.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences* 42(5), 815-824.

Baumgartner, H., Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International Journal of Research in Marketing* 13(2), 139-161.

Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological Bulletin* 107(2), 238-246.

Bentler, P., & Bonett, D. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin* 88(3), 588-606.

Charoenwongsak, K. (2003). *Critical thinking*. Bangkok: Successmedia Ltd.

Chaijaroen, S., Kanjug, I., Samat, C. (2006). Development of Learning Innovations That Enhancing Thinking Potential. Faculty of Education, Khonkaen University.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Alexandria, VA: ASCD.

Browne, M., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research* 21(2), 230-258.

Byrne, B. (1989). *A primer of LISREL*. New York: Springer.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum.

Chaffee, J. (1990). Thinking critically. Houghton Mifflin Company, Boston, MA.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology* 78, 98-104.
Cottrell, S. (2005). *Critical thinking skills*. Basingstoke: Palgrave.
Çokluk, Ö., Şekercioğlu, G., ve Büyüköztürk, Ş. (2014). *Sosyal bilimler için çok değişkenli istatistik SPSS ve LISREL uygulamaları (3.baskı)*. [Multivariate statistical applications with SPSS and LISREL in Social Sciences]. Ankara: Pegem Akademi Yayınları.
Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology* 105(3), 399-412.
Dunteman, G. H. (1989). *Principal components analysis.* Thousand Oaks, CA: Sage Publications.
Dewey, J. (1933). *How we think. A restatement of the relation of reflective thinking to the educative process*. Boston: Heath and Company.
Dewey R. A. (2007). Psychology: An introduction. California: Wadsworth.
Doll, W. J., Xia, W., Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly* 18(4), 357–369.
Esmer, E., & Altun, S. (2013). Öğretmen adaylarının zihinsel stil tercihlerine yönelik bir araştırma: Zihinsel stiller değişiyor mu? *Trakya Üniversitesi Eğitim Fakültesi Dergisi*, 3 (2), 21-30.
Fathima, S. and Rao, D. B. (2008). *Reasoning Ability of Adolescent Students*. India: Discovery Publishing House.
Fletcher, T. D. (2010). *Psychometric: Applied Psychometric Theory.* R package version 2.2. https://CRAN.R-project.org/package=psychometric
Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with un observable variables and measurement error. *Journal of Marketing Research* 18(1), 39-50.
Fox, J., Nie, Z. And Byrnes, J. (2017). *SEM: Structural Equation Models*. R package version 3.1-9. https://CRAN.R-project.org/package=sem
Gorsuch, R. L. (1983). *Factor analysis.* Hillsdale, NJ: Lawrence Erlbaum.
Güneş, F. (2012). Öğrencilerin Düşünme Becerilerini Geliştirme. [Developing students' thinking skills]. Türklük Bilimi Araştırmaları-XXXII.
Gürkaynak, İ., Üstel, F., Gülgöz, S. (2008). *Eleştirel Düşünme* [Critical thinking]. Erg Raporları.
Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2019). *SemTools: Useful tools for structural equation modeling*. R package version 0.5-2. Retrieved from https://CRAN.R-project.org/package=semTools
Kenny, D., Kaniskan, B., & McCoach, D. (2014). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research* 44(3), 486-507.
Kline, P. (1994). *An easy guide to factor analysis*. UK: Routledge.
Kline, R. B. (2011). *Principles and practice of structural equation modeling (3th edition)*. New York: The Guilford.
Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology* 28, 563–575.
Marsh, H. W., Balla J. R. & McDonald, R. P. (1988). Goodness of fit indexes in confirmatory factor analysis: The effect of sample size. *Pshychological Bulletin* 103, 391-410.
Maydeu-Olivares, A., Shi, D., & Rosseel, Y. (2017). Assessing fit in structural equation models: A Monte-Carlo Evaluation of RMSEA versus SRMR Confidence Intervals and tests of close fit. *Structural Equation Modeling: A Multidisciplinary Journal* 25(3), 389- 402.
McDonald, R. P. (1999). *Test theory: a unified treatment*. Mahwah, NJ: LEA Publisher.
Mujtaba, A., Azam, S. & Rana, H. K. (2017). Performance evaluation of various data mining classification techniques that correctly classify banking transaction as fraudulent. *Gub Journal of Science and Engineering* 4 (1), 59-63.
Nisbett, R. E., Peng, K., Choi, I. & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review* 108(2), 291-310.
Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods* 5, 343–355.
Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, 76, 972-987.
Peters, G. J. Y. (2014). The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist* 16(2), 56-69.
Peterson B. G. and Peter C. (2019). *PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis.* R package version 1.5.3. https://CRAN.R-project.org/package=PerformanceAnalytics

Piccone, A. V. (2009). *A comparison of three computational procedures for solving the number of factors problem in exploratory factor analysis*. Published Doctor of Philosophy dissertation, University of Northern Colorado.
R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Raiche, G., Riopel, M. and Blais, J.-G. (2006). Non graphical solutions for the Cattell's scree test. Paper presented at the Intrenational Annual meeting of the Psychometric Society, Montreal.
Revelle, W. (2018) Psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.12.
Revelle, W. & Rocklin, T. (1979). Very simple structure-alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research* 14(4):403–414.
Revelle, W. & Zinbarg, R.E. (2009). Coefficients alpha, beta, omega and the glb: Comments on Sijtsma. *Psychometrika* 74(1), 145-154.
Scherer, R. F., Wiebe, F. A., Luther, D. C., & Adams, J. S. (1988). Dimensionality of coping: Factor stability using the ways of coping questionnaire. *Psychological Reports* 62(3), 763-770.
Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment* 8, 350-353.
Sharma, S., Mukherjee, S., Kumar, A., & Dillon, W. (2005). A simulation study to investigate the use of cutoff values for assessing model fit in covariance structure models. *Journal Of Business Research* 58(7).
Shields, C. (2012). Aristotle. Edward N. Zalta (ed.). The Stanford Encyclopedia of Philosophy. 15 Decemeber 2019 retrieved from http://plato.stanford.edu/archives/sum2012/entries/aristotle/
Sijtsma, K., & Molenaar, I. W. (2002). *Introduction to nonparametric ıtem response theory*. Sage Publications, Thousand Oaks.
Sternberg , R. J. (2006). The rainbow project: Enhancing the SAT through assessments of analytical,practical, and creative skills. *Intelligence* , 34, 321 – 350.
Sternberg, R. J., & Grigorenko, E. L. (1997). Are cognitive styles still in style? *American Psychologist*, 52 (7), 700-712.
Sternberg, R. J., & Grigorenko, E. L. (2001). *Perspectives on thinking, learning and cognitive styles*. New York: Taylor&Francis Group.
Tanaka, J., & Huba, G. (1985). A fit index for covariance structure models under arbitrary GLS estimation. *British Journal Of Mathematical And Statistical Psychology* 38(2), 197- 201.
Thaneerananon, T., Triampo, W., & Nokkaew, A. (2016). Development of a test to evaluate students'analytical thinking based on fact versus opinion differentiation. *International Journal of Instruction* 9(2), 123-138.
Thompson, B. (1992). A partial test distrubition for cosines among factors across samples. B. Thompson, (Ed.), Advances in social science methodology (in Bolume 2), 81-97. Greenwich, CT: JAL.
Umay, A. ve Arıol, Ş. (2011). Baskın olarak bütüncül stilde düşünenlerle baskın olarak analitik stilde düşünenlerin problem çözme davranışlarının karşılaştırılması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi* 30(30), 27-37.
Umay, A. ve Kızıltuğ, Ş. (2007). Analitik ve Bütünsel Düşünenlerin Problem Çözme Davranışı Üzerine Bir Çalışma. Hacettepe Üniversitesi Eğitim Fakültesi ve Azerbaycan Devlet Pedagoji Üniversitesi Uluslararası Öğretmen Yetiştirme Politikaları ve Sorunları Sempozyumu, 12 – 14 Mayıs 2007, Bakü, Azerbeycan.
Üstündağ, T. (2005). *Yaratıcılığa yolculuk (5. Baskı).* [Journey to Creativity]. Ankara: PegemA Yayıncılık.
Veneziano L. & Hooper J. (1997). A method for quantifying content validity of health-related questionnaires. *American Journal of Health Behavior* 21 (1), s. 67-70.
Yulina, I. K., Permanasari, A., Herhnani, H., & Setiavan, W. (2019). Analytical thinking skill profile and perception of pre service chemistry teachers in analytical chemistry learning. *International Conference on Mathematics and Science Education*.
Yves Rosseel (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48 (2),1-36. URL http://www.jstatsoft.org/v48/i02/.
Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's, β and McDonalds ω: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika* 70(1).