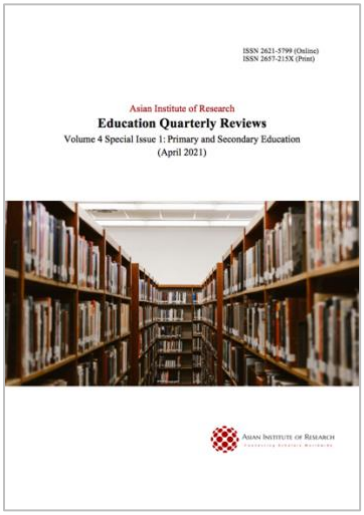

Education Quarterly Reviews

Aslan, Ayşe Esra, and Soysal, Sümeyra. (2021), Reliability and Validity of the Turkish Version of Aurora-a Intelligence Test Battery. In: *Education Quarterly Reviews*, Vol.4 Special Issue 1: Primary and Secondary Education, 214-225.

ISSN 2621-5799

DOI: 10.31014/aior.1993.04.02.241

The online version of this article can be found at:
https://www.asianinstituteofresearch.org/

Published by:
The Asian Institute of Research

The Asian Institute of Research *Education Quarterly Reviews* is a peer-reviewed International Journal. The journal covers scholarly articles in the fields of education, linguistics, literature, educational theory, research, and methodologies, curriculum, elementary and secondary education, higher education, foreign language education, teaching and learning, teacher education, education of special groups, and other fields of study related to education. As the journal is Open Access, it ensures high visibility and the increase of citations for all research articles published. The *Education Quarterly Reviews* aims to facilitate scholarly work on recent theoretical and practical aspects of education.

ASIAN INSTITUTE OF RESEARCH
Connecting Scholars Worldwide

The Asian Institute of Research
Education Quarterly Reviews
Vol.4 Special Issue 1: Primary and Secondary Education, 2021: 214-225
ISSN 2621-5799

DOI: 10.31014/aior.1993.04.02.241

Reliability and Validity of the Turkish Version of Aurora-a Intelligence Test Battery

Ayşe Esra Aslan[1] & Sümeyra Soysal[2]

[1] Istanbul University-Cerrahpaşa, Istanbul, Turkey. ORCID: 0000-0002-0113-1745
[2] Necmettin Erbakan University, Konya, Turkey. ORCID: 0000-0002-7304-1722

Correspondence: Ayşe Esra Aslan, Istanbul University-Cerrahpaşa Alkent 2000 Mah. Yiğittürk Cad. No:5/9/1 Büyükçekmece/ 34500 İstanbul-Turkey E-mail: aeaslan@hotmail.com

Abstract
The Aurora-a test battery was applied to 520 students who were between the ages of 9 and 12 attending public and private schools in Istanbul to create the Turkish version of the Aurora-a Intelligence Test Battery (Aurora-a_TR), which was developed for children aged 9-12 years based on the Triarchic Intelligence Theory. The three sub-test scores that measured verbal intelligence in the original form were excluded from the analyses since they were not suitable for statistical analysis. The validity evidence of Aurora-a_TR test was obtained by Confirmatory Factor Analysis and Reliability Analyses. The analyses supported the three-factor Triarchic Intelligence Model with strong evidence, which also included the fourteen subtests. It was recommended in this study to re-apply Aurora-a_TR on a different sampling with the tests that were re-adapted into Turkish for 3 verbal content tests, which could not be tested in the present study.

Keywords: Triarchic Intelligence Theory, Measurement of Intelligence, Adaptation of Intelligence Test, Factor Analysis

1. Introduction

Intelligence tests are one of the most important techniques for recognizing individuals. Galton, who was one of the pioneers of intelligence, which has been researched with scientific methods since 1884, argued that sensory processes are important for determining the intelligence levels, and that intelligence is passed on to future generations through heredity (Burt, 1968; narrated by Snyderman and Rothman, 1988:51). Spearman, on the other hand, defined general ability with “g” and the specific abilities of an individual with “s” factors (Baykul, 2010). Thorndike (1924, cited by Gottfredson & Saklofske, 2009) argued that there was no factor like general intelligence, and each mental activity required a number of different abilities. Guilford explained cognitive processes with the Cube Model, which included product, process, and content dimensions (Kaygın & Cetinkaya, 2015). Cattell-Horn-Carroll’s (CHC) mind abilities explained narrowing and specializing mental abilities with the three-layer intelligence theory; Gardner (1983) explained the effect of culture on intelligence in the context of problem-

solving, and interpersonal differentiation with the dominant intelligence types. Ceci's biological theory is based on the context to show cognitive abilities (1990, narrated by Pal, 2004).

Sternberg examined the evolution of intelligence theories in a three-level model. the fundamental question in level one is whether a theory of intelligence should be singular or plural; in the second level, the integration of the opinion in the first question in a hierarchical or non-hierarchical way in one sense; and in the third level, the question of what intelligence is and how it should be investigated is dealt with again, and possible new answers are evaluated (Sternberg, 1981). The measurement of the gifted is a separate subheading in the scope of the differentiation of intelligence tests. In the context of superior ability, it is emphasized that the concept of intelligence is a complex structure that consists of the interaction of different cognitive, social, emotional, and environmental factors (Heller, Pertelh & Lim, 2005). In the literature, the use and disadvantaged sides of tests including short (e.g., The Kaufman Brief Intelligence Test) or expanded ability criteria (e.g., Wechsler Tests) have long been discussed and study support was sought in the selection of students to gifted programs (Pierson, Kilmer, Rothlisberg, & McIntosh, 2012). Also, it is recommended that teachers' and parents' observations are included in the diagnoses (Renzulli,1978; Poweryeter, 2016).

Renzulli (1978) considered superior ability as a condition that stems from the interaction of above-the-average general or special ability, creative talent, and motivation. He also stated that the lack of one of these or being below a certain threshold prevents the formation of superior intelligence, and that there is no need to have very high potentials in all these three ability fields to have superior intelligence. In the light of Renzuli and Sternberg's opinions on intelligence and superior ability, a new perspective has been developed on creativity measurement. Mednick's Word Connotations Test (Benedek, & Neubauer, 2013), and Torrance Tests of Creative Thinking test batteries (Torrance,1974) were included in the creative thought intelligence test.

The successful intelligence theory of Sternberg defines intelligence as "a balanced abilities system allowing one to adapt, shape, and choose environments to achieve goals in the context of one's own culture or society". People who are intelligent and also successful need creative capabilities to produce new ideas, and to deal with a relatively new situation, analytical capabilities to detect the value of their new ideas, coping strategies, practical abilities to put these novel ideas into practice, and convince others regarding the values of their ideas (Sternberg, 1999). This opinion shows parallelism with the theory of Piaget on intelligence (Santrock, 2011) defining the ability to adapt to the child's environment by changing schematics as intelligence behaviors. For these reasons, Triarchic Intelligence Theory addresses intelligence under Analytic, Application, and Creative Intelligence headings (Sternberg, 2005). Intelligence requires that serial mental ability is combined to solve problems in a socio-cultural context (Sternberg & Grigorenko, 2004).

Aurora-a Test Battery was developed as part of a project that began at Yale University Child Study Center on the basis of Triarchic intelligence theory. The test, which was worked on for three years, is still a tool whose validity and reliability are tested in seven different languages and countries (Tan et al., 2009; Prieto, Ferrándiz, Ferrando & Bermejo, 2015; Gubbels, Segers, Keuning, & Verhoeven, 2016). As well as the general intelligence score, it can also measure the analytic, application, and creative intelligence types. The ability to measure creative intelligence in line with other intelligence types under one single scale is an important acquisition in recognizing an individual.

Tests, such as Wisc-R and Wisc-4 (Uluç et al., 2011), Basic Abilities Test, and ASIS (Sak, Bal Sezerel, Ayaş & Tokmak, 2016) are widely used in our country to evaluate the cognitive abilities of children who need special education. However, some of these tools that are used have limitations, such as being outdated, or simply measuring mere analytical intelligence. The individual needs to be measured in a multifaceted way to recognize him/her and support his/her development. Also, intelligence tests are needed in international form to make international comparisons and researches on Turkish children. Although the values of international intelligence tests are high, their adaptation requires different knowledge and care. The main objective of this study was to create the Turkish version of ARORA-a Test Battery (Aurora-a TR) for age groups (9-12 years) with analytic, application intelligence, and creative intelligence types and 17 subtests.

2. Method

The Method section describes in detail how the study was conducted, including conceptual and operational definitions of the variables used in the study, Different types of studies will rely on different methodologies; however, a complete description of the methods used enables the reader to evaluate the appropriateness of your methods and the reliability and the validity of your results, It also permits experienced investigators to replicate the study, If your manuscript is an update of an ongoing or earlier study and the method has been published in detail elsewhere, you may refer the reader to that source and simply give a brief synopsis of the method in this section.

2.1 Participant Characteristics

A total of *520* randomly selected volunteer students who studied in public and private schools in 2017-18 in the borders of the city of Istanbul, which received a lot of migration in Turkey, and which covers different regional cultures, constituted the study groups of the Turkish form.

2.2 *Measurement Tool*

Along with the description of subjects, give the mended size of the sample and number of individuals meant to be in each condition if separate conditions were used. State whether the achieved sample differed in known ways from the target population. Conclusions and interpretations should not go beyond what the sample would warrant.

Table 1. The Aurora-a Subtests Grouped by Target Ability and Domain

	Analytical	Practical	Creative
Images (visual/ spatial)	*Shapes (Abstract Tangrams):* complete shapes with missing pieces. (10 items) (MC) *Floating Boats:* identify matching patterns among connected boats. (5 items) (MC)	*Book Covers*: interpret an abstract picture and invent a story to accompany it. (5 items) (OE) *Multiple Uses:* devise three new uses for each of several household items. (5 items) (OE)	*Paper Cutting:* identify the proper unfolded version of a cut piece of paper. (10 items) (MC) *Toy Shadows:* identify the shadow that will be cast by a toy in a specific c orientation. (8 items) (MC)
Words (verbal)	*Words That Sound the Same (Homophone):* Blank sentence with two missing words using homonyms. (20 items) (RW) *Metaphors:* explain how two somewhat unrelated things are alike. (10 items) (OE)	*(Inanimate) Conversations*: create dialogues between objects that cannot typically talk. (10 items) (OE) *Interesting (Figurative) Language:* interpret what sentence logically comes next after one containing figurative language. (12 items) (MC)	*(Silly) Headlines:* identify and explain an alternative "silly" meaning of actual headlines. (11 items) (RW) *Decisions:* list elements given in a scenario on either "good" or "bad" side of a list in order to make a decision. (3 items) (RW)
Numbers (numerical)	*Number Cards (Letter Math):* find the single-digit number that letters represent in equations. (5 items) (RW) *Story Problems (Algebra):* (before any algebra training) devise ways to solve logical math problems with two or more missing variables. (5 items) (RW)	*Number Talk:* imagine reasons for various described social interactions between numbers. (7 items) (OE)	*Maps (Logistics Mapping):* trace the best carpooling routes to take between friends' houses and destinations. (10 items) (RW) *Money (Exchange):* divide complicated "bills" appropriately between friends. (5 items) (RW)

Note: MC = multiple choice; OE = open-ended items that need to be scored by an individual using a rating scale; RW = answers are either right or wrong; ( ) in subtest titles = subtest titles or portions of titles no longer in use.

Source: (Tan et al., 2009, p.448)

2.3. *Study Procedure*

The test was translated independently by two different experts who mastered both in English and Turkish, checked by two field specialists, and was then converted into one single Turkish form. The linguistic equivalence studies of the Aurora-a TR are prepared as separate articles. The items of the subtest of Words that sound the same were created again by the researchers based on words written the same in Turkish but which meant differently by considering cultural differences. In the preliminary application conducted at a state and a private school, it was observed that Headlines, Decision and Interesting language subtests could not be understood by children. For this reason, these subtests were not included in the study. The Turkish version of the test consisted of 14 subtests according to the results of the study. Scoring and interpreting the test requires psychological test usage competence and scoring expertise. The test can be applied as a group or as an individual test. The test consists of a total of 119 questions as Analytical Intelligence ($N_{question}$= 59), Practical Intelligence ($N_{question}$= 33), and Creative Intelligence ($N_{question}$=27) subtests, and the application time of the test is 159 minutes. Detailed information on the number of items, response styles, application times, and the characteristics measured by the subtests is given in Appendix 1

2.4 Data Analysis

The validity analyses of Aurora-a-TR Battery were examined as based on Factor Analysis. CFA was done to determine whether the data were in line with the structure defined by the American version. The Mplus 7 Program was used for the analyses. Weighted Least Square Mean and Variance Adjusted (WLSMV) estimation method was preferred because there were different item types with two and more categories. WLSMV has been recommended for estimating CFA model parameters with categorical variables (Muthén & Muthén, 2010). Firstly, single-dimensional DFA was carried out for each subtest, and then, second order CFA was conducted to check whether the subtests were predicted by the intelligence they were related to. Multiple evaluations were made for model-data compatibility according to the criteria of CFI >0.90 (Browne & Cudeck, 1993), TLI>0.90 (Bentler & Bonett, 1980), RMSEA <.08 (Jöreskog & Sörbom, 1993). In CFA analyses, one item was excluded from Paper Cutting sub-scale with standardized pathway coefficient below 0.30, and three items were excluded from Shapes subtest because of the same reason. The findings were reported over 115 items.

Convergent and divergent validity checks were done. Convergent validity was evaluated according to composite reliability (CR)>0.70 and AVE>0.50 (Fornel & Lacker, 1981). Divergent validity criteria: (1) If the correlation coefficient of two dimensions is less than the individual Cronbach α reliability coefficient (Gaski & Nevin, 1985), and (2) if the correlation coefficient of two dimensions is smaller than the square root of AVE, then these two dimensions have divergent validity (Fornel & Lacker, 1981; Gaski & Nevin, 1985).

Cronbach α coefficient was used in the sub-test reliability, and the Stratified Cronbach Alpha Coefficient was used for general scale reliability. Cronbach, Schonemann and Brennan (1965) proposed to use the Stratified Cronbach α coefficient for the reliability of the combined scores obtained from measurement tools with sub-dimensions. The Stratified Cronbach α coefficient was calculated by using the "sirt" package in program R (Robitzsch, 2017).

After the validity of the measurement and structural models for Analytical, Creative and Practical Intelligence was achieved, three model trials were conducted for all the data as Unidimensional Model, Three-Factor Model and General Factor Model. Although the $\chi 2$ difference test can be used to compare nested models in the WLSMV method, there is no statistic allowing the direct comparison of non-nested models. For this reason, improvement in the goodness of fit indices and in factor loadings were considered in the decision-making process in the comparison of the 3 non-nested models in question.

3. Results

The results of the confirmatory factor analysis of the measurement model of the subtests of the Analytical Intelligence Scale are summarized in Table 2.

Table 2: Fit Statistics and Factor Loadings for The Model Fit of the Subtests of the Analytical Intelligence Test

Analytical Intelligence Subtests	Fit Measures			Range of standardized path coefficient
	RMSEA	TLI	CFI	
Story Problems	0.02	0.99	0.99	0.47-0.80
Words that sounds the same	0.09	0.94	0.95	0.52-0.95
Metaphors	0.09	0.97	0.98	0.67-0.75
Number Cards	0.14	0.98	0.99	0.53-0.97
Shapes	0.04	0.97	0.98	0.40-0.80
Floating boats	0.08	0.94	0.96	0.58-0.87
Analytical Intelligence (second order model)	0.04	0.94	0.95	0.42-0.86

When Table 2 is examined, it is seen that the TLI and CFI indices of the subtests are above acceptable limits, and the RMSEA value is above the limit of 0.8 in some subtests. When the minimum factor loadings of 0.40 is considered, it can be argued that the subtests fitted an one-dimensional structure. The fit indices were obtained above the limit values in the second order CFA model, which was conducted after fitting the measurement models of the subtests. In this way, it was shown that Analytical Intelligence is the predictor of Story Problems, Words that Sound Same, Metaphors, Number Cards, Shapes and Floating Boats subtests.

Table 3. Cronbach Alpha, CR, AVE, square root of AVE (in bold) and correlations between subtests of analytical intelligence (off-diagonal).

Subtests	N of items	Alpha	Subtests							
			CR	AVE	F	SP	NC	S	WTS	Met
Floating Boats	10	0,82	0,92	0,54	**0,73**					
Story Problems	5	0,68	0,81	0,47	0,42	**0,69**				
Number Cards	5	0,80	0,95	0,79	0,35	0,65	**0,89**			
Shapes	7	0,64	0,79	0,47	0,42	0,68	0,65	**0,69**		
Words that Sounds	20	0,92	0,97	0,60	0,20	0,38	0,31	0,38	**0,77**	
Metaphors	9	0,90	0,91	0,53	0,33	0,62	0,51	0,61	0,42	**0,73**
Scale (Stratified Alpha)	56	0,94								

In Table 3, Cronbach α reliabilities of Analytical Intelligence subtests vary between 0.64-0.92; and composite reliabilities vary between 0.81-0.95. AVE values range from 0.47 to 0.79. According to Fornel and Larcker (1981), CR should be higher than 0.70, and AVE should exceed 0.50 so that is adequate for convergent validity. However, Fornel and Larcker argued that if AVE is less than 0.50 but CR is higher than 0.60, convergent validity of the construct is still adequate. It may be argued that the subtests of the analytical intelligence have convergent validity because CR values are ≥0.79, as well as reasonably goodness of fit values. Also, the reliabilities and the square root of the AVE are higher than the correlation coefficient of two dimensions, it is considered of proof of discriminant validity. When both Cronbach α and the square root of the AVE are compared with correlations, it may be accepted that the subtests of the analytical intelligence have divergence validity. The overall reliability of the Analytical Intelligence Scale was calculated to be 0.94 with Stratified Alpha.

The results of CFA regarding the measurement model of four subtests of the Creative Intelligence Scale are summarized in Table 4.

Table 4. Fit Statistics and Factor Loadings for the Fit of the Subtests of the Creative Intelligence Test

Creative Intelligence Subtests	Fit Measures			Range of standardized path coefficient
	RMSEA	TLI	CFI	
Multiple uses	0.17	0.98	0.99	0.85-0.86
Book cover	0.19	0.94	0.97	0.73-0.81
Number talk	0.23	0.92	0.94	0.76-0.85
Conversation	0.15	0.96	0.97	0.75-0.84
Creative Intelligence (second order model)	0.06	0.97	0.98	0.15-0.87

When Table 4 is examined, it is seen that the TLI and CFI goodness of fit indices of the subtests are above the acceptable limits, and the RMSEA value is above the limit of 0.8 in all subtests. When it is considered that factor loadings are ≥0.73, it may be accepted that the subtests fitted an unidimensional structure. Goodness of fit indices was obtained above the criterian values in the second order CFA model, which was conducted after fitting the measurement models of the subtests. In this way, it was shown that Creative Intelligence was the predictive of multiple uses, book cover, number talk and conversation subtests.

Table 5. Cronbach Alpha, CR, AVE, square root of AVE (in Bold) And Correlations Between Subtests of Creative Intelligence (Off-Diagonal).

Subtests	N of items	Alpha	Subtests					
			CR	AVE	MU	BC	NU	C
Multiple Uses	5	0,92	0,93	0,73	**0,85**			
Book cover	5	0,86	0,88	0,59	0,12	**0,77**		
Number Talk	7	0,91	0,93	0,65	0,51	0,09	**0,81**	
Conversations	10	0,94	0,95	0,65	0,71	0,13	0,54	**0,81**
Scale (Stratified Alpha)	27	0,96						

In Table 5, Cronbach αreliabilities and composite reliabilities of creative intelligence subtests are over 0.70; and AVE values are above 0.50. In this way, the convergence reliability of the subtests was proven. When Cronbach α reliabilities and the square root of the AVE (bold) are compared with the correlations, it may be accepted that the subtests of the creative intelligence have the divergence validity. Also, the overall reliability of the creative intelligence scale was calculated to be 0.96 with stratified alpha. The results of the CFA of the measurement model of the six subtests of the practical intelligence scale are summarized in Table 6.

Table 6. Fit Statistics and Factor Loadings for The Fit of the Subtests of the Practical Intelligence test

Practical Intelligence Subtests	Fit Measures			Range of standardized path coefficient
	RMSEA	TLI	CFI	
Maps	0.10	0.85	0.88	0.40-0.96
Paper cutting	0.04	0.95	0.96	0.33-0.75
Toy shadows	0.04	0.97	0.98	0.50-0.88
Money	0.05	0.97	0.98	0.75-0.79
Practical Intelligence (second order model)	0.03	0.95	0.95	0.34-0.89

When Table 6 is examined, it is seen that the goodness of fit values of subtests other than Maps are above the acceptable limits. Goodness of fit values was slightly below acceptable limits in the Maps subtest; however, factor loadings were reasonable. It was decided that the Maps scale should remain in the main scale because its divergence and convergence validities were adequate. Goodness of fit values was obtained above the criteria values in the second order CFA model, which was conducted after the subtests fitted an unidimensional structure. In this way, it was shown that Practical Intelligence was the predictor of Maps, Paper Cutting, Toy Shadows and Money subtests.

Table 7. Cronbach Alpha, CR, AVE, square root of AVE (in bold) and Correlations Between Subtests of Practical Intelligence (Off-Diagonal).

Subtests	N of items	Alpha	Subtests					
			CR	AVE	TS	PC	MO	MA
Toy Shadows	8	0.77	0.88	0.49	**0.70**			
Paper Cutting	9	0.67	0.84	0.37	0.40	**0.61**		
Money	5	0.71	0.87	0.58	0.62	0.52	**0.76**	
Maps	10	0.72	0.86	0.39	0.24	0.19	0.30	**0.62**
Scale (Stratified Alpha)	32	0.84						

It is seen in Table 7 that the Cronbach α and composite reliabilities of the Practical Intelligence subtests were>0.70, and the AVE values were <0.50 in general. Fornel and Larcker (1981) as the reference, it was accepted that the convergence validity is still adequate if AVE is <0.50 but the CR is ≥0.60. When Cronbach αand the square root of AVE were compared with the correlations, it was accepted that the subtests of the Practical Intelligence have the divergence validity The overall reliability of the Practical Intelligence Scale was calculated to be 0.84 with Stratified Alpha.

Table 8. Goodness-of-Fit Statistics for the Confirmatory Factor Analysis Models.

Model	$\chi 2$	df	CFI	TLI	RMSEA
Single-factor model	19843.57	6440	0.70	0.70	0.06*
Three factor model	9479.51	6422	0.93	0.93	0.03
General factor model	9476.51	6422	0.93	0.93	0.03

*p<0.05

After the validity of the structural models of the analytics, application and creativity intelligence scales, different structural models were also tested by adding all the items and subtests. As summarized in Table 8, firstly, a one-dimensional CFA model (single factor model) was tested. In this model, it was seen the factor loadings of some items were below <0.30; and goodness of fit indices was insufficient. Then, a second-order CFA model, which was related to analytics, practical, and creativity intelligence scales, was tested. It was found that the goodness of fit indices of this three-factor model were above the criteria values, and the factor loadings varied between 0.40-0.96.

Finally, the third-order CFA model, which collected the three scales under a general factor, was also tested. In this general model, although factor loadings (0.39-0.96) and goodness indices were above the criteria values, the standardized path coefficient (i.e., the correlation) between analytical intelligence and general intelligence was calculated to be >1.00. It is considered that this negative condition, which is known as the Heywood Case, stemmed from a sampling error. Also, this model did not show improvements in goodness of fit compared to the three factor model. As a result, depending on the comparison of these three structural models, it was concluded that the Three-factor Model was more fitted to the data (Figure 1).

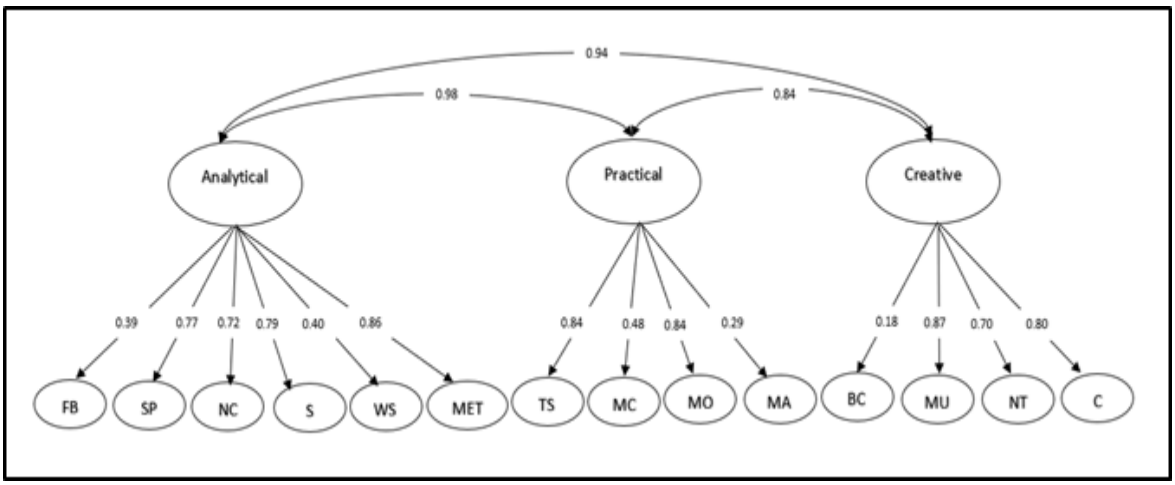

Figure 1: Standardized path coefficients for the 14 Aurora-a subtests in the three-factor model. (Observed variables omitted for simplicity)

4. Discussion

The structural characteristics of the Turkish form of the Aurora-a Intelligence Test, which was tested in different countries as an international scale, and which was the main problem of the study, adapted for Turkish children (between 9-11 years) studying at primary school 4th-6th grades were tested. Analytical, Practical, and Creative Intelligence, and a general intelligence above these three intelligences, were defined in the psychometric structure of the original Arora-a Battery. Firstly, the measurement models of Analytical, Practical, and Creative Intelligence subtests were tested in the Turkish form. As a result of the CFA analyses done for story problems, words that sound the same, metaphors, number cards, shapes and floating boats subtests in Analytical Intelligence, all items except for the "story problem" showed reasonable factor loadings. Three items were excluded from the analyses in the Story Problem subtest. In addition to Cronbach Alpha reliability, the convergence and divergence validity of analytical intelligence subtests was also adequate. The second order CFA model that was identified for these six subtests proved that the Aurora-a TR battery could measure analytical intelligence.

Similar structural evidence was obtained for Practical Intelligence, which consisted of Maps, Paper Cutting, Toy Shadows, and Money subtests. Structural evidence was achieved for the subtests (Multiple Uses, Book Covers, Talking Figures, Conversations) except for three subtests (Decision, Headings, Interesting Language) under the Practical Intelligence. Only one item that produced low factor loading in Paper Cutting subtest was excluded from the analyses. In Decision Making, Titles, Interesting Language subtests, which include items based on verbal skills, children were not able to answer the items of these subtests. It is estimated that they cannot understand the items of these subtests due to the fact that the idioms and expressions in the original form were preserved and adapted to Turkish.

Tan et al. (2009) emphasized that although it is accepted that Sternberg's intelligence understanding components and intelligence representations are universal, how the outward indicators of these components are reflected in a certain culture as a language and behaviors should also be considered. He also argued that although each country started with translations in the international journey of the Arora-a test Battery, it was noticed that a new tool requiring adaptation had to be created. The reflection of socio-cultural effects on the scores in the adaptation of intelligence tests to another language and country is discussed in the literature (Malda et al., 2010; Vijver, 1997). It is reported that using original or almost-originally translated tools will save cost and time; however, an adaptation with cultural knowledge, values and practices is still required to idealize a tool for a particular cultural context (Malda et al., 2010).

Three structural model were tested for Aurora-a_TR battery. The data did not fit to a one-dimensional model. Standardized path coefficients and goodness of fit indices of the three-factor model showed that the model adequately fit to the data. A general factor model was applied to determine whether the general dimension defined

in the original Aurora-a Battery complied with the Turkish version. However, the goodness of fit indices obtained for the general factor did not show much differentiation from the three-factor model. It was also observed that the correlation between the general factor and analytical intelligence was >1.00. Since the conditions that might have caused this, such as multicollinearity and outliers were checked, it was considered that this problem might have stemmed from sampling error. Although the data of the Aurora-a_TR battery did not support the general factor model, the high correlations obtained from the three-factor model between Analytical, Creative and Practical Intelligence (Figure 1) can be shown as evidence that a total score can be obtained.

Similar to the present study, Aljughaiman and Ayoub (2012) conducted a test of a gifted students program in Saudi Arabia in line with the Triarchic Intelligence Theory obtained from 5th-6th graders. To test the effectiveness of the program, some subtests of the Arora-a Test in the scope of Analytical (Metaphors, Shapes, Number Cards), Creative (Speeches, Book Covers, Numbers), and Practical Intelligence (Decision, Toy Shadows, Money) Intelligence types were used. In this context, although the number of the subtests in the original test was less, it was observed that the Goodness of Fit of the Triarchic Intelligence Structure was excellent. Specifically, this model produced a non-significant $\chi 2/df = 34.99$, $p = .069$. In addition, the values of RMSEA = 0.048, GFI = 0.96, AGFI = 0.93, and NFI = 0.97 indicated the suggested model for Aurora fits with the data.

High correlations were found between the total scores of Arora and the Analytic (0.91), Creative (0.94) and Practical (0.84) intelligence in the confirmatory factor analysis study that was conducted by Aghababaei, Malekpour, Kajbaf, & Abedi (2016) with the data of 400 Iranian students. The Cronbach Alpha coefficient was 0.92. The three-factor structure was thus confirmed in Iranian culture.

Similarly, Gubbels, Segers, Keuning and Verhoeven (2016) tested the three-factor structure in Dutch adaptation; however, they reported that the analytics/practical and creative subtests of the Aurora-a Battery were better for the two-factor structure (RMSEA= .09, CFI= .88, $\chi 2 /df = 4.91$).

As a conclusion, the Turkish form (Aurora-a TR) was obtained compatible with the original structure of the Aurora-a Test Battery, which measures the analytical, application and creative intelligence types. It is recommended for future studies to re-apply the Aurora-a TR with a different sampling and with adapted tests by adapting the 3 verbal content tests to Turkish again, which could not be tested in this study. Also, determining whether the battery has the ability to differentiate gifted students can be considered to be a new study topic.

References

Aljughaiman, A. M., & Ayoub, A. E. A. (2012). The effect of an enrichment program on developing analytical, creative, and practical abilities of elementary gifted students. *Journal for the Education of the Gifted, 35*(2), 153-174.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588–606. doi:10.1037/0033-2909.88.3.588

Benedek, M., & Neubauer, A. C. (2013). Revisiting Mednick's model on creativity-related differences in associative hierarchies. Evidence for a common path to uncommon thought. *The Journal of creative behavior, 47*(4), 273-289.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Chart, H., Grigorenko, E. L., & Sternberg, R. J. (2008). Identification: The aurora battery. In J. E. Plucker & C. M. Callahan (Eds.), *Critical issues and practices in gifted education: What the research says* (pp. 281-301). Waco, TX: Prufrock.

Cronbach, L. J., Schonemann, P., & McKie, D. (1965). Alpha coefficients for stratified-paralel tests. Educational and Psychological Measurement, 25, 291-312.

Fornell, C., & Larcker, D. F. (1981).Evaluating structural equation models with unobservable variables and measurement error. Journal of marketing research, 39-50. doi:10.2307/3151312

Gardner, H. (1983*). Frames of mind. the theory of multiple intelligence*. New York: Basic Books.

Gaski J. F. & Nevin, J. R. (1985). The differential effects of exercised and unexercised power sources in a marketing channel. *Journal of marketing research,* 130-142. doi:10.2307/3151359

Gottfredson, L., & Saklofske, D. H. (2009). Intelligence: Foundations and issues in assessment. *Canadian Psychology, 50*(3), 183-195. doi: 10.1037/a0016641

Grigorenko, E. L. (2009). Considering language, culture, and cognitive abilities: The international translation and adaptation of the Aurora Assessment Battery. In E. L. Grigorenko (Ed.), *Assessment of abilities and competencies in the era of globalization* (p. 443–468). New York, NY: Springer.

Gubbels, J., Segers, E., Keuning, J. & Verhoeven, L. (2016). The Aurora-a Battery as an Assessment of Triarchic Intellectual Abilities in Upper Primary Grades. *Gifted Child Quarterly, 60*(3), 26–238. doi: 10.1177/0016986216645406

Güçyeter, S. (2106). Gifted identification researches in Turkey and the measurement equipments that used in identification. *Turkish Journal of Education, 5*(4), 235-254.

Hein, S., Tan, M., Aljughaiman, A., & Grigorenko, E. L. (2015). Gender differences and school influences with respect to three indicators of general intelligence: Evidence from Saudi Arabia. *Journal of Educational Psychology*, 107(2), 486.doi: /10.1037/a0037519.supp

Heller, K. A., Pertelh, C., & Lim, T. K. (2005). The Munich model of giftedness designed to identify and promote gifted students. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 147–170). Cambridge: Cambridge University Press.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago, IL: Scientific Software International.

Kaygın, B., & Çetinkaya, Ç. (2015). Yaratıcılığın değerlendirmesinde yeni yaklaşımlar. *Journal of Gifted Education and Creativity, 2*(1), 1-11. doi: 10.18200/JGEDC.2015210883

Malda, M., van de Vijver, F. J., Srinivasan, K., Transler, C., & Sukumar, P. (2010). Traveling with cognitive tests: testing the validity of a KABC-II adaptation in India. *Assessment, 17*(1), 107-115.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus: Statistical analysis with latent variables user's guide 6.0.* Los Angeles, CA: Muthén & Muthén.

Pal, H. R., Pal, A., & Tourani, P. (2004). Theories of intelligence. *Everyman's Science, 39*(3), 181-192.

Pierson, E. E., Kilmer, L. M., Rothlisberg, B. A., & McIntosh, D. E. (2012). Use of brief intelligence tests in the identification of giftedness. *Journal of Psychoeducational Assessment, 30*(1), 10-24.

Prieto, D., Ferrándiz, C., Ferrando, M., & Bermejo, M. R. (2015). Aurora Battery: A new assessment of successful intelligence. [La Batería Aurora: una nueva evaluación de la inteligencia exitosa] *Revista de Educación, 368*, 183-210. doi: 10.4438/1988-592X-RE-2015-368-294.

Robitzsch, A. (2020). *sirt: Supplementary Item Response Theory Models*. R package version 3.9-4. http://r.meteo.uni.wroc.pl/web/packages/sirt/sirt.pdf

Sak, U. BalSezerel, B. Ayaş, M. B. Tokmak, F. (2016). Anadolu Sak Zekâ Ölçeği (ASIS) uygulayıcı kitabı. Eskişehir: Centre for Gifted Education of Anadolu University.

Santrock, J. W. (2011). Life-Span Development [Yaşam Boyu Gelişim]. (G.Yüksel, Trans.) Ankara: Nobel.

Snyderman, M., & Rothman, S. (1988). *The IQ controversy, the media and public policy*. New Brunswick, NJ: Transaction Books.

Sternberg R. J. (1995). Styles of thinking and learning. *Language Testing, 12*(3), 265-291. doi:10.1177/026553229501200302

Sternberg R. J. (1999). The Theory of successful ıntelligence. *Review of General Psychology. 13*(4), 292-316. doi:10.1037/1089-2680.3.4.292

Sternberg, R. J. (1981). The evolution of theories of intelligence. *Intelligence, 5*(3), 209-230.

Sternberg, R.J.& Grigorenko, E.L. (2004) Successful intelligence in the classroom. *Theory into Practice, 43*(4), 274-279. Retrieved from http://www.jstor.org/stable/3701535.

Tan, M. T., Aljughaiman, A. M., Elliott, J. G., Kornilov, S. A., Ferrando-Prieto, M., Bolden, D. S., Adams-Shearer, K., Chart, H. E., Newman, T., Jarvin, L., Sternberg, R. J., & Grigorenko, E. L. (2009). Considering language, culture, and cognitive abilities: The international translation and adaptation of the Aurora Assessment Battery. In E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (p. 443–468). New York, NY: Springer.

Torrance, E. P. (1974). *Torrance Test of Creative Thinking, Verbal Tests Forms A and B (Figural A and B),* IL: Scholastic Service Inc. Bensenville.

Uluç, S., Öktem, F., Erden, G., Gençöz, T., & Sezgin, N. (2011). Wechsler intelligence scale for children-iv: a new era for turkey in evaluation of intelligence in the clinical context. *Turkish Psychological Articles, 14*(28), 49.

Vijver, F. V. D. (1997). Meta-analysis of cross-cultural comparisons of cognitive test performance. *Journal of Cross-Cultural Psychology*, 28(6), 678-709.

Appendix A

Aurora-a_TR subtests and the characteristics it measures

Table A1: Aurora-a_TR subtests and the characteristics it measures

Types of intelligence measured	Subtests	N of items	Scoring type	Time of app.	Feature measured by test
Analytical	Story Problems	5	0-1	12'	With this subtest, analyzing the problem, making a logical comparison and mathematical thinking skills are measured.
	Words that Sound the Same	20	0-1	15'	With this subtest, linguistic intelligence is measured. To accomplish the assignment, the individual needs to use the skills of reading and understanding, questioning what s/he understands, and placing words appropriately to provide semantic integrity.
	Metaphors	9	0-2 accuracy * 0-4 ability**	12'	With this subtest, it is desired to reveal the similarities and differences between words, and also to consider the secondary meanings that words take in context. This subtest also reveals the ingenuity or weakness of the individual in using the language. In addition, the test measures the individual's analytical thinking ability.
	Number cards	5	0-1	8'	With this subtest, the power of visualizing relations and performing mathematical operations between the figures are measured. In addition, abstract thinking skills are required to be successful in this test scoring is done as Correct (1) Wrong (0) in two categories.
	Shapes	10	0-1	18'	Visual-spatial intelligence and abstract thinking skills are measured with this subtest. In order to achieve this assignment, it is necessary to think from part to whole, and whole to part.
	Floating Boats	10	0-1	7'	Understanding the relationship between elements by using visual-spatial intelligence skills and problem-solving skills are measured.
	7 tests	59	0-104	72'	
Practical	Maps	10	0-2	7'	With this subtest, the spatial perception, the ability to think and calculate the visual material are measured.
	Paper Cutting	10	0-1	7'	This subtest measures the ability of the individual to establish relationships between parts through visual and mental imagery. At the same time, the test also provides information about the individual's visual attention and organizational skills.
	Toy Shadows	8	0-1	6'	The use of visual-spatial intelligence and the ability to imagine geometric shapes in mind from the previous visual information capacity are measured with this subtest.

	Money	5	0-1	15'	The ability to use mathematical thinking skills in everyday life is measured by the money subtest.
	4 subtest	33	0-43	35'	
Creative	Multiple Uses	5	0-2 accuracy * 0-4 ability**	15'	The ability to use creative thinking skills in daily life problems is measured. The skills of thinking outside the box, originality, unconventional thinking are required.
	Book Covers	5	0-2 accuracy * 0-4 ability**	12'	With this subtest, the connotations of the given visual stimuli and verbal creative thinking skills, imaginativeness are measured.
	Number Talk	7	0-2 accuracy * 0-4 ability**	10'	Mathematical creativity is measured by this subtest. This test requires to identify relations between numbers differently, with the help of imagination and with a multifaceted perspective.
	Conversations	10	0-2 accuracy * 0-4 ability**	15'	With this subtest, being able to distinguish the characteristics of the surrounding objects, and the ability to see their relations from different angles are measured.
	4 subtest	27	0-204	52'	
4 tests	14 subtest	119	0-351	159'	
*"Accuracy" is the accuracy score of the answer given in accordance with the instruction. The answers are scored from the wrong answer (0) to the exact correct answer (2) by three-point grading. **"Ability" is the ability to be evaluated in terms of the expected ability or creativity. The answers are scored from the lowest (0) to the highest (4) by four-point grading					