

Psychometric Properties of Turkish Version of Aggression Questionnaire Short Form: Measurement Invariance and Differential Item Functioning across Sex and Age

Yaşar KUZUCU * Özge SARIOT ERTÜRK **

Abstract

The aim of the present study was to test the psychometric properties of the Aggression Questionnaire Short Form for adolescents and adults in Turkish. The adaptation study was conducted with 778 adolescents aged between 15-18 and 1067 adults aged between 19 and 44. The construct validity of the questionnaire was tested via Parallel Analysis, Exploratory Factor Analysis and Confirmatory Factor Analysis. Furthermore, item-total correlations, test-retest score correlation, and internal consistency (Cronbach Alpha and McDonald's Omega) were calculated as reliability analyses. The Measurement Invariance test and Differential Item Functioning in male and female, adolescent and adult samples were also conducted. The results yielded that the Turkish version of the Aggression Questionnaire Short Form is a reliable questionnaire with four-factors, and without sex and age differences, it can be used to measure aggression among Turkish adolescents and adults.

Key Words: Aggression questionnaire short form, measurement invariance, differential item functioning

INTRODUCTION

Aggression is a multidimensional construct that develops within a complex interaction of biological, psychological, social, and cultural factors (Vitoratou, Ntzoufras, Smyrnis, & Stefanis, 2009) and has received great deal of attention in mental health area (Evren, Çınar, Güleç, Çelik, & Evren, 2011; Hinshaw; 1987; Johnson, Carve, & Joormann, 2013; Podubinski, Lee, Hollander, & Daffern, 2017). A large number of theoreticians and researchers tried to explain the origin and reason of aggression and association of aggression with other behaviors (Chang, Schwartz, Dodge, & McBride-Chang, 2003; Coie & Dodge, 1998; Maslow, 1943; Moyer, 1982; Sexton et al. 2019).

Several measurement tools were developed to measure this essential issue (Buss & Perry, 1992; Orpinas & Frankowski, 2001; Kang, Lim, Suh, Gang, & Pedersen, 2020; Palmstierna & Wistedt, 1987; Raine et al. 2006). The Buss-Perry Aggression Questionnaire (BPAQ; Buss & Perry, 1992) is one of the most frequently used measurement tool in the literature to measure aggression (Adıgüzel, Özdemir & Şahin, 2019; Kühn et al. 2019; Singh, 2017). Buss-Durkee Hostility Inventory (BDHI; Buss & Durkee, 1957) is the origin of the questionnaire. Researchers constructed BPAQ as a more current instrument in terms of psychometric properties. BPAQ is a 5-point Likert scale, consists of 29 items and has four factors. These factors are physical aggression, verbal aggression, anger, and hostility. Additionally, different from the other instruments developed to measure aggression, BPAQ has validity for both adolescent (Reyna, Sanchez, Ivacevich, & Brussino, 2011) and adult samples (Vitoratou et al. 2009). Moreover, it is used with both clinical (Evren et al. 2011) and nonclinical samples (Özdemir, Vazsonyi & Çok, 2017) rather than just with clinical or nonclinical ones (Palmstierna & Wistedt, 1987). BPAQ also provides valid and reliable data from offenders (Diamond, Wang & Buffington-Vollum, 2005). In terms of factor structure, the scale explains aggression with four structures that involve different forms of active and passive aggression, rather than just proactive or reactive aggression (Raine et al. 2006). The

* Associate professor, Aydın Adnan Menderes University, Education Faculty, Aydın-Turkey, e-mail: yasarku@yahoo.com, ORCID ID: 0000-0002-8487-9993

** Research assistant, Aydın Adnan Menderes University, Faculty of Science and Arts, Psychology Department, Aydın, Turkey, e-mail: ozge.sariot@adu.edu.tr, ORCID ID: 0000-0003-4565-8300

To cite this article:

Kuzucu, Y., & Sariot-Ertürk, Ö. (2020). Psychometric Properties of Turkish Version of Aggression Questionnaire Short Form in Adolescents and Adults. *Journal of Measurement and Evaluation in Education and Psychology*, 11(3), 243-265. doi:

Received: 01.02.2020

Accepted: 17.09.2020

psychometric properties of the BPAQ were tested with different methodologies and samples, and research results confirmed the original four-factor structure of the questionnaire (Bernstein & Gesn, 1997; García-León et al. 2002; Gerevich, Bácskai, & Czobor 2007; Harris, 1997; Reyna et al. 2011; Torregrosa et al. 2020). However, most of the studies reported better fit to original factor structure or better factor loadings when some items are omitted (Bernstein & Gesn, 1997; Gerevich et al. 2007; Harris, 1995). Additionally, researchers reported BPAQ as an inadequate measurement tool because of the explained common variance by these four factors (Bryant & Smith, 2001).

In order to develop an acceptable measurement model for the BPAQ, Bryant and Smith (2001) refined the questionnaire and proposed a 12 item version (short form) of the Aggression Questionnaire (AQ-SF). The new short form of the AQ-SF also has a four-factor structure model with the same names, physical aggression, verbal aggression, anger arousal and hostility. Each factor includes three items. Unlike the BPAQ, the AQ-SF is a 6 point Likert questionnaire (Bryant & Smith, 2001). However, most of the studies (e.g., Maxwell, 2007; Torregrosa et al. 2020) which includes AQ-SF preferred the 5 point Likert type version.

As BPAQ, the psychometric properties of the AQ-SF (12 item version of AQ) was tested with different methods and samples. The AQ-SF showed good construct validity in the offenders (Diamond & Magaletta, 2006) and mentally ill male prisoners (Diamond et al. 2005). Sex invariance of the questionnaire was also confirmed for the Argentinean adolescents (Reyna et al., 2011) and federal offenders (Diamond & Magaletta, 2006). Maxwell (2007) tested validity on the translated Chinese version AQ-SF with Chinese sample. Results indicated a good fit to the data and adequate internal reliability. The Dutch version of AQ-SF also has sufficient validity and reliability in the psychiatric patient and the student samples (Hornsveld, Muris, Kraaimaat, & Meesters, 2009).

In addition to the good psychometric properties of the AQ-SF, remarkable relations with aggression and other mental health issues were reported in the studies that used the 12-item version of the AQ-SF. The relation between aggression and collective narcissism (De Zavala, Cichocka, Eidelson, & Jayawickreme, 2009), hubristic pride (Carver, Sinclair, & Johnson; 2010) mindfulness and rumination (Borders, Earleywine, & Jajodia, 2010) were pointed out. Johnson et al. (2013) reported significant relation of anger and verbal aggression dimensions with borderline personality characteristics, anxiety symptoms and alcohol consumption.

As in varied languages the Turkish 29 item version of the BPAQ was also studied. In order to test the psychometric properties of BPAQ, studies were conducted with college students (Madran, 2012), adolescents (Önen, 2016) and male substance dependent inpatients (Evren et al. 2011). Despite their different sample profiles, all have a common result; the Turkish version of the BPAQ is a valid and reliable questionnaire to measure aggression. However, no studies have been conducted to test the psychometric properties of the AQ-SF in Turkish.

The AQ-SF was reported as acceptable to use in different cultures, sexes, clinical and nonclinical samples. The relation of aggression with both well-being and ill-being variables was pointed out when aggression was measured through the AQ-SF. Taking into account all of these, it seems essential to introduce the AQ-SF into Turkish. Therefore, this study aims to test the construct validity and reliability of the AQ-SF and to test sex and age invariance of the questionnaire in the Turkish sample.

METHOD

This study, which aims to adapt the AQ-SF into Turkish, is a descriptive study. Descriptive studies attempt to explain “what” events, objects, entities, institutions, groups, and areas are (Fraenkel, Wallen & Hyun, 2012). In this descriptive study, the validity and reliability analyses were conducted, and the psychometric properties of AQ-SF were determined. Detailed information about participants, the data collection instrument, and data analysis are presented below.

Study Group and Process

The AQ-SF was implemented to 778 students between the ages of 15 and 18 from five different high schools. The self-report measures were administered to the participants at their school. Participants were volunteers, and no personal information was assembled. The whole data was collected two times for Parallel Analysis (PA), Exploratory Factor Analysis (EFA), and Confirmatory Factor analysis (CFA). PA and EFA were conducted with 383 adolescents. In order to CFA, the data from 395 adolescents were used.

AQ-SF was also applied to the adult group. The adult group consists of overall 1067 people, undergraduate students from Aydın Adnan Menderes University, University of Ege and University of Ankara, graduated from university and participated in pedagogical formation training and trainees in the public training center. Participants were determined by convenience sampling, and they were voluntarily participating. Two different data sets were used for PA, EFA (n= 648) and CFA (n= 419). The distribution of the study groups is given in Table 1 and Table 2.

Table 1. The Distribution of the Study Group for the EFA

Adolescent (15-18 years of age)			Adult (19-35 years of age)		
Sex	F	%	Sex	F	%
Male	98	26.41	Male	220	33.95
Female	273	73.58	Female	428	66.04
Total	371	100.0	Total	648	100.0
Age	F	%	Age	F	%
15	97	25.3	19-23	510	78.70
16	77	20.1	24-30	122	18.82
17	128	33.4	31-35	16	2.46
18	81	21.1	Total	648	100.0
Total	383	100.0			

Table 2. The Distribution of the Study Group for the CFA

Adolescent (15-18 years of age)			Adult (19-44 years of age)		
Sex	F	%	Sex	F	%
Male	165	41.1	Male	130	31.63
Female	230	57.4	Female	281	68.36
Total	395	100.0	Total	411	100.0
Age	F	%	Age	F	%
15	201	50.1	19-23	291	74.44
16	124	30.9	24-30	71	18.15
17	65	16.2	31-35	14	3.58
18	6	1.5	36-44	15	3.83
Total	396	100.0	Total	391	100.0

The Adaptation Procedure

The original questionnaire was independently translated from English into Turkish by four experts in psychology. In addition to the individual translation, using the focus group technique, each item was evaluated by the same experts. The group members are composed of people who know both languages and cultures, have measurement tool development skills, and know the purpose of the translated measurement tool. Consensus was reached on a common draft by these experts. Then back-translated by bilingual psychiatry and psychology experts who are different from the experts in the translation process.

Data Collection Instruments

Aggression

Aggression was measured by using the AQ-SF. The AQ-SF containing 12 items comprised the refined four-factor measurement model. This questionnaire was developed from Buss and Perry's 29-item AQ, and it has a four-factor structure; physical aggression, verbal aggression, anger, and hostility. The physical aggression, involves nine items, factor loadings of these items changes between .44 and .84. The verbal aggression involves five items and factor loadings of these items changes between .35 and .56. The anger, consists of seven items and these items' factor loadings change between .35 and .75. Lastly, the hostility involves eight items and their factor loadings change between .37 and .70 (Buss & Perry, 1992). Although Buss and Perry (1992) did not report explained variance for the AQ, Garcia-Leon et al. (2002) supported four-factor structure of the questionnaire and reported variance explained by the whole questionnaire as 42.1 %. Cronbach Alpha values of the factors and the total score for the AQ-SF are .85, .72, .83, .77, and .89, respectively. Moreover, test-retest reliability estimates are .80, .76, .72, .72, and .80 for the four factors and total score, respectively (Buss & Perry, 1992).

Bryant and Smith (2001) explored the factor structure of the AQ. The researchers deleted items that displayed low or multiple loadings in a principal component analysis and excluded a number of reverse-scored items. This procedure yielded the AQ-SF (12 item), for which the hypothesized four-factor model produced an acceptable fit. The AQ-SF has the same factor structure with the AQ. Each dimension had three items. However, Bryan and Smith (2001) did not report factor loadings, explained variance and test retest reliability of AQ-SF. In addition to obtaining dimension scores, a total aggression score can also be calculated. Cronbach Alpha values for the dimensions of the original AQ-SF change between .70 and .83. In the original form (Buss & Perry, 1992) the questionnaire is a 5 point Likert questionnaire and Bryant and Smith (2001) adopted the questionnaire to a 6-point response tool ranging from 1 (extremely uncharacteristic of me) to 6 (extremely characteristic of me). Despite the adaptation of Bryant and Smith (2001) in the current study, the original 5-point questionnaire (1 = uncharacteristic of me, 5 = very characteristic of me) was sustained likewise previous adaptation studies (Abd-El-Fattah, 2013; Maxwell, 2007; Torregrosa et al. 2020) in order to compare the present results with earlier researches in a credible way.

Social problem solving

The Social Problem Solving Inventory-Revised Short-Form (SPSI-RSF; D'Zurilla, Nezu, & Maydeu-Olivares, 2002) was used. The scale has 25 self-administered questions that are developed to assess cognitive, emotional or behavioral reactions of individuals to real life problem-solving situations. It has five dimensions, each involves five items, comprising two problem orientations as positive and negative, and three problem-solving styles, as rational, impulsive/carelessness, and avoidance. In terms of the validity, Sorsdahl, Stein, and Myers (2017) reported the variance explained by SPSI-RSF as 57.9%. The inventory has good internal consistency ($\alpha=.84$), excellent test-retest reliability, ($r=.90$), and good discriminant validity tested on a sample of sexual offenders (Webster, Mann, Thornton, & Wakeling, 2007). The Turkish form of the tool (Eskin & Aycan, 2009) supported original factor structure. Factor loadings for positive orientation change between .52 and .67, for negative orientation .62 and .81, for rational orientation .60 and .72, for impulsive/carelessness orientation .38 and .76, lastly, for avoidance orientation .35 and .90. CFA results for Turkish form of the inventory is also acceptable; χ^2 / df 2.15, RMSEA = .04, CFI = .92. In the adaptation study, the coefficients of internal consistency and test-retest reliability differed from .62 to .92 and from .60 to .84, respectively (Eskin & Aycan, 2009). In the present study, the coefficient of internal consistency is ranged from .68 to .90 for adolescents and .69 to .80 for adults.

Trait anger

Trait Anger was assessed using the 10-item subscale of the Anger Expression Scale (Spielberger, 1985). Trait Anger and Anger Expression Style Scale (STAXI) is a self-report scale comprised of 44 items; 10 items of this 44 item scale define trait anger, 10 items define state anger, and 24 items define anger expression style (Anger control, Anger-out and Anger-in). The scale allows researchers to use each subscale independently. Trait Anger Scales (TAS) reports how angry they generally feel. The TAS correlates positively with a variety of anger and hostility measures such as the Buss-Durkee Hostility Inventory and with various state anger measures and discriminates high from low anger groups (Spielberger, 1988). The reliability study of the STAXI-2 with adult males from the general population reports alpha coefficients ranging from .73–.95 for the total scale scores and from .73–.94 for the subscales (Spielberger et al., 1985). In Turkish adaptation study (Özer, 1994), for anger control, the coefficients of internal consistency were calculated as .84. In the present study, the coefficient of internal consistency is .83 for adolescents and .87 for adults.

Data Analysis

SPSS 25.0 (SPSS Inc.), Factor Analysis 10.10 (Fernando & Lorenza-Seva, 2017), LISREL 8.80 (Jöreskog & Sorbom, 1993) and jMetrik Version 4.1.1 statistical package programs were used in the analysis. The data were analyzed using PA, EFA, and CFA techniques for the construct validity. Furthermore, item-total correlations, test-retest score correlation, internal consistency estimates of reliability (Cronbach Alpha and McDonald's Omega) were calculated. T-test was performed to test whether the items of the questionnaire distinguished between the lower and upper 27% groups. By examining the measurement invariance (MI) in female-male and adolescent-adult samples, it was tested whether the measurement tool was appropriate for the comparisons between groups. In order to test the validity of the questionnaire by item, Differential Item Functioning (DIF) tests were conducted for sex and age groups. Expert opinion was used to determine what the source of the DIF is for an item that gives DIF (Doğan & Öğretmen, 2008).

RESULTS

AQ-SF Adolescent Application

To test the psychometric properties of the measurement in adolescent, validity and reliability analyses was conducted. All analyses were explained in detail.

Pre-analyses

In order to determine whether the data showed normal distribution or not, measures of central tendency, Skewness and Kurtosis values were examined. The results about central tendencies, showed that Mean = 29.17, Median = 29, and Mode = 30. The similarity of these scores indicates the normal distribution of the data (Büyüköztürk, 2007). For aggression total score Skewness is .11, and Kurtosis is -.23 (n=778, data set for PA, EFA, and CFA). The fact that both values are between the range of -1, +1 implies that they show normal distribution.

Kaiser Meyer Olkin (KMO) coefficient was used to determine whether the data structure was appropriate for factor analysis in terms of the sample size of the application. As a result, KMO value was determined as 0.79. The fact that KMO value is high means that each variable in the questionnaire can be estimated well by the other variables (Field, 2013). Bartlett's test of Sphericity was significant ($\chi^2(66, n = 383) = 1261.459 p < .001$), and this value supported the factorability of the correlation matrix. Another indicator of the appropriateness of the data for factor analysis is the Anti-image Correlation Matrix. These values need to be above 0.5, and the values below this must be excluded from the analysis (Field, 2013). The diagonal values for each variable in the anti-image matrix vary between .70 and .89.

The fact that all the values of the intersection point are above 0.5 indicates that it is accurate to include all the items in the questionnaire.

The validity analysis

The factor structure for the construct validity of the questionnaire was determined by performing PA and EFA. The purpose of performing PA and EFA is to gather the variables that are related to each other and that measure the same quality together, and to reduce the number of items forming the questionnaire (Aksu, Eser, & Güzeller, 2017; Horn, 1965). CFA was performed to test whether the restricted structure defined by PA and EFA was verified as a model (Horn, 1965; Tabachnick & Fidell, 2013).

When the factor structure of the questionnaire is analyzed via PA and EFA the scree plots are also examined. As can be seen in Figure 1 the graph curve shows a sharp decrease till the fourth factor and that the curve proceeds horizontally after the fourth factor. It indicates that this finding supports the four-factor structure of the questionnaire.

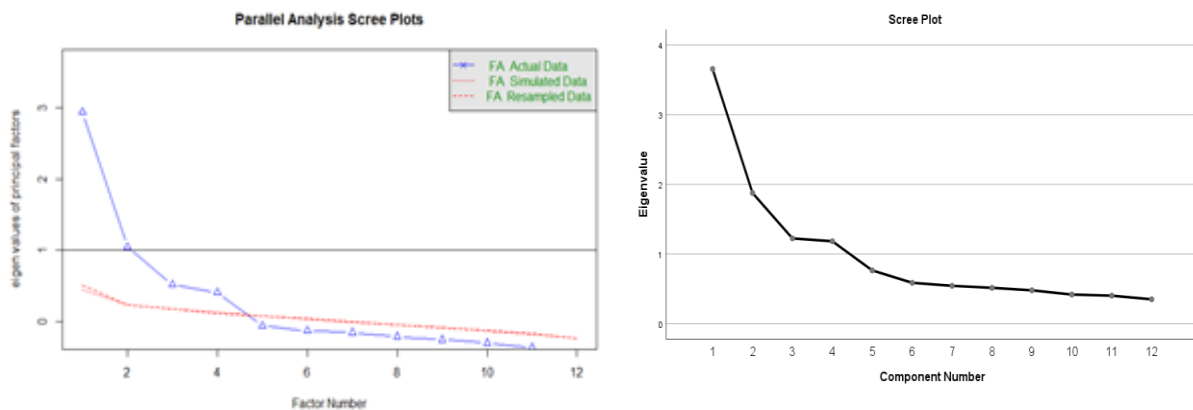


Figure 1. AQ-SF Parallel Analysis and EFA Scree Plots Graph of Adolescent Application

In PA, factor number is decided through comparing eigenvalues from real data and simulated random parallel data set that is produced based on the real data set. Factor number is accepted till the point in which the real data eigenvalue is larger than the parallel data eigenvalue (Akbaş, Karabay, Yıldırım-Seheryeli, Ayaz, & Demir, 2019). Depending on these explanations and the values mentioned in Table 3, the PA results indicated that, the adolescent application of AQ-SF has four factors.

Table 3. Eigenvalues from PA for Adolescents Application

Factors	1	2	3	4
Eigenvalues from sample correlation matrix	3.88	1.63	1.21	1.14
Average eigenvalues from parallel analysis	1.20	1.15	1.11	1.08
95th percentile eigenvalues from parallel analysis	1.25	1.19	1.14	1.10

Notes: n = 778

The result of the EFA with 12 items indicated that the items were collected in 4 sub-dimensions, with eigenvalues greater than 1. The items of each sub-dimension were examined, and it was determined that they were grouped under the factor to which they were related. To clarify the relationship among factors, the varimax rotation (the orthogonal rotation technique of Principal Component Analysis) is used. As a result of the EFA it was found that the eigenvalue of the factors from the first to the fourth were 2.12, 2.09, 1.88 and 1.83 respectively. Additionally, the variance explained by the factors from the first to the fourth were 17.74, 17.45, 15.70 and 15.25 respectively. The total variance explained by the

questionnaire was found at 66.16%. When the eigenvalues and cumulative variance percentages of the four factors were taken into consideration, it was determined that the questionnaire had four factors. The findings obtained as a result of the EFA performed for AQ-SF Adolescent Application revealed that the construct validity of the questionnaire was sufficient and factor structure was similar to the original form. The factors formed after EFA and the items collected under each factor are given in Table 4.

Table 4. Factor Loadings, Item-Total Correlations and Common Variances for Adolescent Application

Factors	PA Factor Loadings				EFA Factor Loadings				Item-Total Correlation	Common Variances
	1	2	3	4	1	2	3	4		
Physical	.80	-.25	.16	.00	.83	.19	.09	-.01	.45	.73
	.80	-.04	-.01	-.02	.82	.06	.18	.06	.46	.72
	.64	.03	.03	.02	.77	.21	.11	.11	.49	.67
Verbal	.04	.46	-.01	.04	.16	.79	.10	.03	.40	.67
	-.02	.79	-.04	-.03	.15	.76	.26	.01	.45	.67
	.01	.70	.04	-.00	.12	.65	-.03	.10	.30	.45
Anger	-.07	-.09	.70	-.04	.01	-.10	.82	.08	.33	.69
	-.00	-.01	.70	.05	.10	.23	.75	.14	.49	.65
	.08	.10	.63	-.01	.24	.31	.65	.12	.56	.60
Hostility	-.06	.04	.05	.66	.00	.06	.06	.86	.39	.76
	.04	-.04	-.10	.94	.00	.04	.04	.84	.36	.72
	-.00	.02	.11	.51	-.01	.06	.22	.76	.38	.62

Notes: PA = Parallel Analysis, EFA= Exploratory Factor Analysis

When Table 4 is examined, the results of PA and EFA reveal that each item is clustered under a factor that is related to a value that is more than twice as much as the factor loading value that they have in other factors. This finding, which shows that the items differentiate in terms of factors, supports the construct validity of the questionnaire. As can be seen in Table 4, each factor is composed of the three items. The factor loadings of the first factor vary between .80 and .64 for PA, .83 and .77 for EFA. The factor loading of the second factor values varies between .46 and .70 for PA, .79 and .65 for EFA. The factor loadings of the third factor vary between .70 and .63 for PA, .82 and .65 for EFA. The factor loadings of the fourth vary between .66 and .94 for PA, .86 and .76 for EFA. Following this phase, the items in each factor were examined as a whole, and a factor structure consistent with the original form of the questionnaire was observed. In order to determine whether there were significant correlations among the factors forming AQ-SF adolescent application, Pearson Correlation Analysis was performed. It was revealed that the relationship coefficients of “Physical aggression” factor with “Verbal Aggression”, “Anger”, and “Hostility” were found as .39, .38, and .25 respectively; and the relationship coefficient of “Verbal Aggression” with “Anger” and “Hostility” was found as .38 and .22 respectively; and lastly, the relationship coefficient between “Anger” and “Hostility” was determined as .34. The results obtained, consistent with the literature (Şahin, 2018), show a positive significant ($p \leq .001$) relationship among all the factors of the questionnaire.

First-order and second-order CFA was performed to evaluate the applicability of the four factors of AQ-SF Adolescent application. The models obtained from these analyses are given in Figure 2. Additional to the first and second-order CFA, 1- factor solution was also tested.

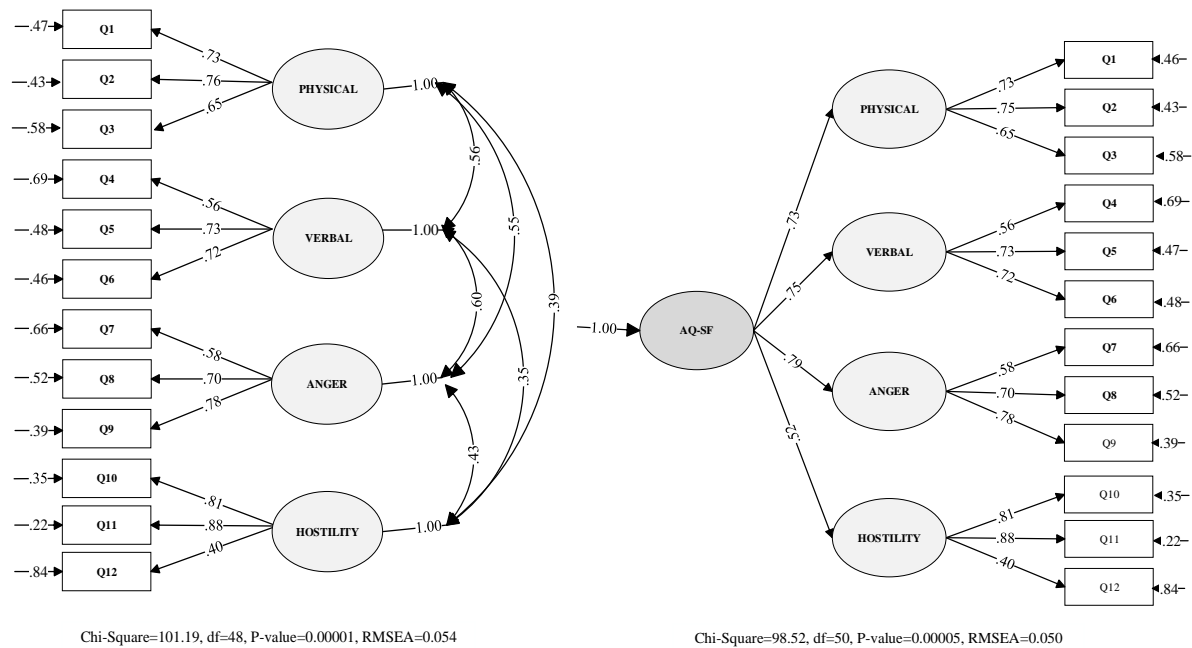


Figure 2. AQ-SF Adolescent 1st and 2nd Order CFA

First and second-order CFA were performed for four-factor structured AQ-SF adolescent application. When the CFA was evaluated, χ^2/sd ratios for the first and second-order were determined as 2 ($\chi^2/sd=96/48$) and 1.97 ($\chi^2/sd=98.52/50$), respectively. The fact that χ^2/sd ratios obtained as a result of first and second-order CFA are ≤ 2.0 , correspond to a good fit. RMSEA fit index values were determined as 0.051 and 0.050 as a result of first and second-order CFA, respectively. The fact that RMSEA fit index value is below and equal to 0.05 can be interpreted as a good fit (Kline, 2015). It was determined that, among the fit index values related to the model as a result of the first and second order CFA, AGFI was 0.93, GFI was 0.96, standardized RMR fit index value was 0.059, NFI fit index value was 0.96, and CFI fit index value was 0.98. There is no statistically significant difference between first and second-order CFA (less than 3.84 chi-square difference with one degree of freedom); however, the second-order was evaluated to be superior since it is more parsimonious. When all the values related to data fit of the model are taken into consideration, it can be seen that the model formed shows a sufficient order to fit with the data.

Another CFA was performed to support the multifactorial structure of AQ-SF adolescent application; the results of first and second-order factor analyses were compared with the one-factor analysis of the questionnaire. The questionnaire was assumed to have one dimension, and it produced the following statistics: χ^2/sd ratio of the fit values used in the model comparisons was calculated as 9.41 ($\chi^2/sd=508.48/54$, RMSEA= 0.15, GFI= 0.82, NFI= 0.79, CFI = 0.81). The results showed that the one-factor structure had poorer fit values than the multifactorial structure.

In order to determine the convergent validity of AQ-SF adolescent application, the relationship between AQ-SF scores with trait anger scores was examined with Pearson Product-Moment Correlation Analysis. The correlation of the AQ-SF with trait anger ($r=.54$) is moderate and statistically significant ($p\leq.001$). Additionally, to determine the divergent validity of AQ-SF, the relationship between AQ-SF scores and social problem-solving scores was examined in the same way. Results showed a negative ($r=-.30$) and statistically significant ($p\leq.001$) relationship between the two variables.

The reliability analysis

Item analysis was conducted with all adolescent data ($n=778$) to determine the contribution of the items in the questionnaire of the implicit structure they belong to and to measure the level of discrimination between the items with and without relevant characteristics of the structure they belong to (Erkuş, 2012).

The Cronbach Alpha and McDonald's Omega coefficients were calculated for all and each factor of the questionnaire. It is suggested that McDonald's Omega coefficient is more appropriate for multi-dimensional measures (Revelle, 2018). Cronbach Alpha and McDonald's Omega coefficients have following values for the first factor .76 and .73, for second factor .68 and .70, for the third factor .70 and .60, for the fourth factor, .74 and .64 respectively. Cronbach Alpha and McDonald's Omega were calculated as .80 and .76 for the total score.

Test-retest reliability was found as .99. Item total correlation coefficients varied between .57 and .62 for the first factor, .41 and .56 for the second factor, .49 and .54 for the third factor, .49 and .67 for the fourth factor, .30 and .56 for total.

It was also analyzed whether there was a significant difference between individuals with low scores and high scores. As a result of the *t*-test conducted to compare the responses of the individuals in the lower 27% group and the responses of the individuals in the upper 27% group to all the items in the questionnaire, the items' *t* values varied between 62.73 ($p < .001$) and 32.96 ($p < .001$) and a significant difference was found. In the analysis performed, it was found that the variances were heterogeneous. It can be seen that the reliability values of the overall and factors of the AQ-SF adolescent application are generally acceptable for social sciences.

AQ-SF Adult Application

To test the psychometric properties of the measurement in adults, validity and reliability analyses were conducted. All analyses were explained in detail.

Pre-analyses

In a similar manner with the adolescent application analysis, measures of central tendency, Skewness, and Kurtosis values were examined. The central tendency results showed that Mean = 29, Median = 29, and Mode = 29. Skewness and Kurtosis values were examined ($n = 1067$, data set for PA, EFA, and CFA). Skewness was found .30, and Kurtosis was found .02. As for the data of adolescence, the similarity of central tendency measures, Skewness, and Kurtosis values indicated normality for data of adults.

KMO value was determined as 0.78. It means that each variable can be estimated well by the other variable. Bartlett's test of Sphericity was significant ($\chi^2 (66, n = 648) = 1985.553$ $p < .001$) and this value supported the factorability of the correlation matrix. Besides, the Anti-Image Correlation Matrix intersection values were also analyzed and it was found that these values varied between .68 and .89. As the values at this intersection point were above 0.5, it was determined that it was accurate to include all the items in the questionnaire.

The validity analysis

PA and EFA were conducted for adult application data, too. When the "Scree Plots" graphs were examined (Figure 3), it can be seen that the curves show a sharp decrease till the fourth factor and that the curve proceeds horizontally after the fourth factor. The results are consistent with the previous results showing that the questionnaire has a four-factor structure.

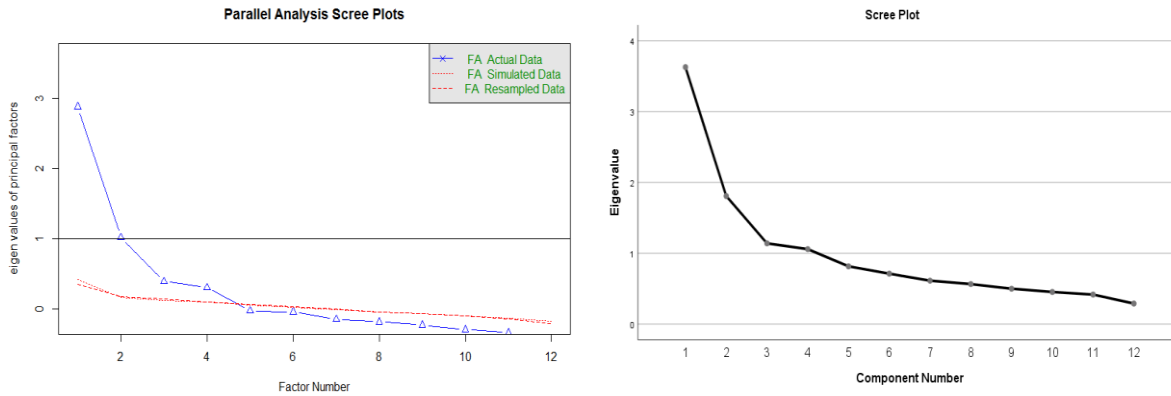


Figure 3. AQ-SF Parallel Analysis and EFA Scree Plots Graph of Adult Application

Accordingly, to the PA results, when eigenvalues from real data and stimulated parallel data were compared (see in Table 5), it indicates that consistent with the original structure, the adult application of AQ-SF has four factors.

Table 5. Eigenvalues from PA for Adult Application

Factors	1	2	3	4
Eigenvalues from sample correlation matrix	3.63	1.74	1.15	1.10
Average eigenvalues from parallel analysis	1.17	1.13	1.10	1.07
95 th percentile eigenvalues from parallel analysis	1.21	1.16	1.12	1.09

Notes: n=648

The items were grouped under the factor, with eigenvalues greater than 1, to which they were related. To clarify the relationship among factors, varimax rotation (the orthogonal rotation technique of Principal Component Analysis) is used.

As a result of the EFA it was found that the eigenvalue of the factors from the first to the fourth were 2.17, 1.92, 1.86 and 1.67 respectively. Additionally, the variance explained by the factors from the first to the fourth were 18.12, 16.01, 15.25 and 13.95 respectively. The total variance explained by the questionnaire was 63.61%. The findings revealed that the construct validity of the questionnaire was sufficient and factor structure was similar with the original form. The factors formed after EFA and the items collected under each factor are given in Table 6.

Table 6. Factor Loadings, Item-Total Correlations and Common Variances for Adult Application

Factors	PA Factor Loadings				EFA Factor Loadings				Item-Total Correlation	Common Variances
	1	2	3	4	1	2	3	4		
Physical	.68	-.19	.25	-.03	.83	.16	.06	.03	.39	.72
	.82	-.01	-.07	-.21	.77	.03	.27	.09	.44	.67
	.45	.02	.12	.04	.69	.18	.09	.05	.37	.52
Verbal	-.02	.32	.04	.10	.22	.77	.14	-.02	.41	.66
	-.04	.72	-.02	-.03	.20	.75	.19	.06	.46	.65
	.05	.70	-.00	-.01	-.01	.58	.09	.22	.33	.40
Anger	-.03	.07	.51	-.01	.17	.06	.79	.16	.41	.57
	.00	-.10	.77	.02	.22	.20	.73	.12	.48	.67
	.02	.06	.67	-.03	.03	.22	.72	.06	.56	.64
Hostility	-.01	.02	.14	.57	.03	.12	.06	.89	.44	.81
	-.01	.03	-.10	.97	.07	.03	.10	.84	.42	.73
	.02	-.04	.02	.75	.06	.11	.17	.76	.44	.62

Notes: PA= Parallel Analysis, EFA= Exploratory Factor Analysis

As can be seen in Table 4, each factor is composed of the three items. The factor loadings of the first factor vary between .82 and .45 for PA, .83, and .69 for EFA. The factor loading of the second factor values varies between .72 and .32 for PA, .77, and .58 for EFA. The factor loadings of the third factor vary between .77 and .51 for PA, .79 and .72 for EFA. The factor loadings of the fourth factor vary between .97 and .57 for PA, .89 and .76 for EFA. Following this phase, the items in each factor were examined as a whole and a factor structure consistent with the original form of the questionnaire was observed. In order to determine whether there were significant correlations among the factors forming AQ-SF adult application, Pearson Product-Moment Correlation Analysis was performed. It was revealed that the relationship of “Physical Aggression” factor with “Verbal Aggression”, “Anger”, and “Hostility” was found as .38, .38, .21, respectively; the relationship of “Verbal Aggression” with “Anger” and “Hostility” was found as .38 and .24 respectively, and lastly, the relationship between “Anger” and “Hostility” was determined as .29. The results obtained, consistent with the literature (Şahin, 2018), show a positive significant relationship among all the sub-dimensions of the questionnaire $p \leq .001$.

First and second-order CFA were performed to determine whether the 12-item, 4-factor structure of the questionnaire achieved after EFA performed for AQ-SF adult application would be verified. The models obtained from these analyses are given in Figure 4. One-factor solution was also tested.

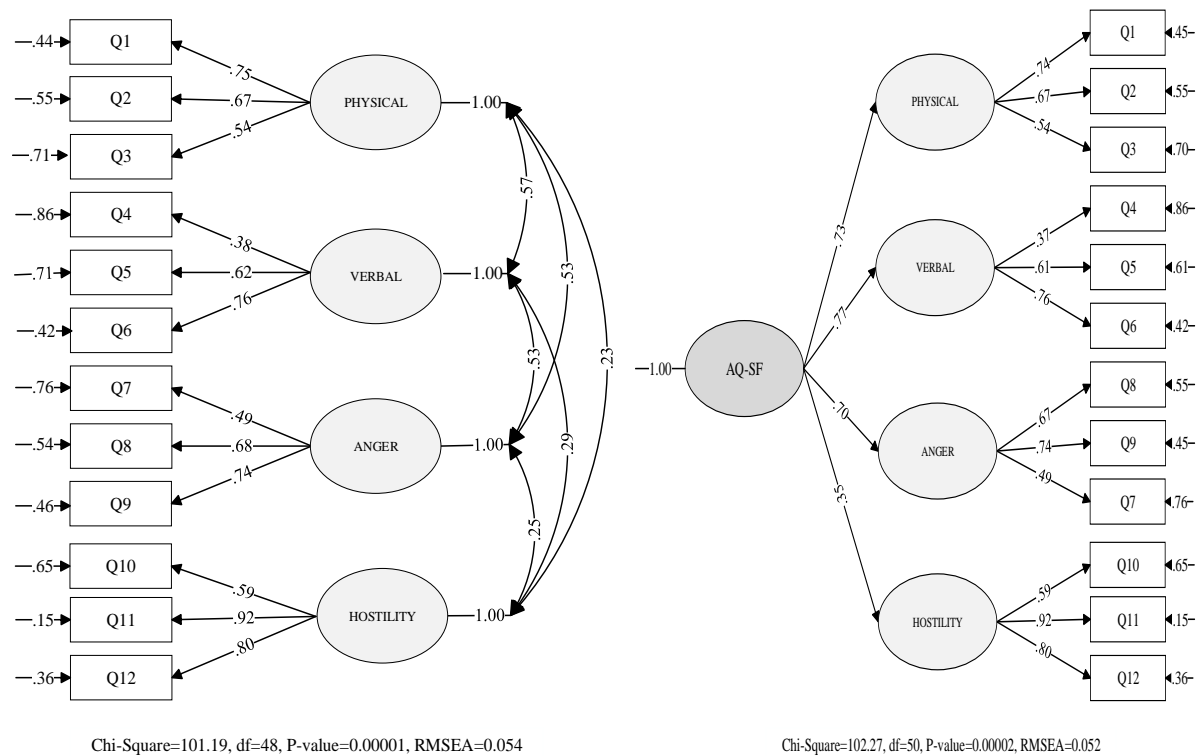


Figure 4. AQ-SF Adult 1st and 2nd Order CFA

First and second-order CFA were performed for AQ-SF adult application. When the result of CFA was evaluated, χ^2/sd ratios for the first and second-order were determined as 2.11 ($\chi^2/sd=101.19/48$) and 2.05 ($\chi^2/sd=102.27/50$), respectively. RMSEA fit index values were as determined as 0.054 and 0.052 as a result of first and second second-order CFA, respectively. It was determined that, among the fit index values related to the model as a result of the first and second-order CFA, AGFI was 0.93, GFI was 0.96, standardized RMR fit index value was 0.063, NFI fit index value was 0.94, and CFI fit index value was 0.97. When all the values related to data fit of the model are taken into consideration, it can be seen that the model has good fit indices.

An additional CFA was performed to support the multifactorial structure of AQ-SF adult application; the results of the first and second-order factor analyses were compared with the one-factor analysis of the questionnaire. The questionnaire was assumed unidimensional and it produced following statistics: χ^2/sd ratio of the fit values used in the model comparisons was calculated as 11.41 ($\chi^2/sd=616.26/54$, RMSEA= 0.17, GFI= 0.79, NFI= 0.66, CFI = 0.67). Consistent with the model comparison in the adolescent group, the second-order CFA was considered to be superior since it has higher degrees of freedom, i.e., having more parsimony. The results also showed that the one-factor structure had poorer fit values than the multifactorial structure.

In order to determine the convergent validity of AQ-SF adult application, the relationship between trait anger scores and AQ-SF scores from the adult application was examined with Pearson Correlation Analysis, and it was found that there is a positive ($r=.56$) and statistically significant ($p\leq.001$) relationship between the two variables. Additionally, to determine the divergent validity of AQ-SF adult application, the relationship between social problem solving and AQ-SF scores from the adult application was examined, and aggression has a statistically significant relationship with social problem solving ($r =-.31$, $p\leq.001$).

The reliability analysis

The reliability analysis of each factor and overall of the AQ-SF adult application was also conducted. Cronbach Alpha and McDonald's Omega coefficients have the following values for first factor .70 and .68, for second factor .60 and .60, for the third factor .68 and .62, for the fourth factor, .80 and .65 respectively. Cronbach Alpha and McDonald's Omega were calculated as .78 and .72 for all questionnaire.

Test-retest reliability was found as .98. Item total correlation coefficients varied between .45 and .56 for the first factor, .29 and .48 for the second factor, .44 and .53 for the third factor, .56 and .74 for the fourth factor, .33 and .56 for the total.

Item analysis was performed to compare the responses of the individuals with low scores and high scores. As a result of the *t*-test performed for this purpose, *t* values of the items varied between 8.16 ($p < .001$) and 2.83 ($p < .001$), and there was a significant difference. It can be seen that the reliability values of the overall and sub-dimensions of the AQ-SF adult application are generally acceptable values for social sciences.

Measurement Invariance for Sex and Age

For the questionnaire to show this it measures in the same manner for two subgroups MI is tested (Vandenberg & Lance, 2000). In the MI process, the aim is to test the factor structure of the questionnaire for different groups and to reach to a similar factor structure for compared groups. MI is frequently checked via multi-group confirmatory factor analysis (MG-CFA) (Jöreskog & Sörbom, 1993; Meredith, 1993). Additionally, these models are based on the increasingly restrictive assumptions regarding to the relations between the observed variables and the latent factor(s). These hierarchical models are named structural invariance, metric invariance, strong invariance, and strict invariance respectively through the least strict one to the most. For structural invariance an equal factor structure (i.e., constraining the number of factor(s) and the pattern of fixed and free loadings) across groups is required. When this requirement is met, it means respondents from various groups employ the same conceptual framework when responding (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000). Metric invariance requires invariant factor loadings across groups. This would have accepted that the content of the factors is the same across groups and that relationships between variables can justifiably be compared across groups (Iurino & Saucier, 2020; Milfont & Fischer, 2015). The third step, strong invariance necessitates equivalent intercepts (for continuous variables) or equivalent thresholds (for ordinal variables), invariant intercepts across groups and it suggests that means across groups can be compared (Gustavsson, Eriksson, Hilding, Gunnarsson, & Östensson, 2008; Iurino & Saucier, 2020). In the most rigid model, a strict invariance implies equivalent residual variances and indicates that the systematic measurement error is invariant across groups (Iurino & Saucier, 2020; Meredith, 1993). Among these models, in a hierarchical way, to meet the requirements of a model means to meet the requirements of the previous model(s).

Additionally, in the decision of how well MI models fit the data, several model indexes are used. Chi-square (χ^2), the root mean square of error of approximation (RMSEA), comparative fit index (CFI), non-normed fit index (NNFI), are some of them (Emerson, Guhn, & Gadermann; 2017; Guo et al., 2017). In the acceptable MI conditions, it is expected that differences between indexes (RMSEA, CFI, NNFI) of ensuing models should be equal or smaller than -0.01, Δ RMSEA, Δ CFI, Δ NNFI \leq -0.01 (Guo et al., 2017; Wu, Li, & Zumbo, 2007) and χ^2 show insignificant change from previous model (Guo et al., 2017).

In this manner, the present study tested whether participants from different groups having the same aggression level will have the same scores from AQ-SF or not through MI. In other words, to determine whether the properties of the questionnaire are invariant among males and females, MI was examined in terms of sex. In addition to this sex comparison, the questionnaire was tested in different age groups. To test the MI of the factor structure of the questionnaire was being measured for the sex groups (male, female) and age (adolescents and adults), MG-CFA was used. For this purpose, four hierarchical models

were tested respectively: structural invariance, metric invariance, strong invariance, and strict invariance.

Moreover, in this study, it was examined whether the invariance conditions of $\Delta RMSEA$, ΔCFI , ΔNFI ≤ -0.01 for MG-CFA study files which are compatible with the data were obtained. The fact that $\Delta RMSEA$, ΔCFI , and ΔNFI values obtained as a result of the comparison of the two models are equal to -0.01 or below can be used as the evidence that the MI is achieved (Wu, Li, & Zumbo, 2007).

The findings regarding the invariance steps tested are present in Table 7. “The Structural Invariance Model” in the table represents the factor loads, regression constant, and the error variances free model; “The Weak Invariance Model” in the table represents the factor loads constant, regression constants, and error variances free model; “The Strong Invariance Model” in the table represents the factor loads, regression constants, and error variance free model; and “The Strict Invariance Model” in the table represents the factor loads, regression constants, and error variances constant model.

Table 7. Fit Statistics Regarding MI

Steps	χ^2	Df	RMSEA (CI)	$\Delta RMSEA$	CFI	ΔCFI	NNFI	ΔNFI
Sex								
SI	162.70	108	0.036 (0.02; 0.04)		0.98		0.98	
MI	232.91	120	0.049 (0.04; 0.05)	-0.013	0.97	0.01	0.97	0.01
SgI	246.35	126	0.049 (0.04; 0.05)	0.000	0.97	0.00	0.97	0.00
StI	246.61	126	0.050 (0.04; 0.05)	0.001	0.97	0.00	0.97	0.00
Age								
SI	282.19	108	0.064 (0.05; 0.07)		0.96		0.95	
MI	333.32	120	0.067 (0.05; 0.07)	0.003	0.95	0.01	0.95	0.00
SgI	338.54	126	0.065 (0.05; 0.07)	0.002	0.95	0.00	0.95	0.00
StI	380.06	126	0.071 (0.06; 0.08)	-0.006	0.94	-0.01	0.94	0.01

Notes: n= 782 (for sex), 792 (for age) CI= Confidence Interval, SI= Structural Invariance, MI= Metric Invariance, SgI= Strong Invariance, StI= Strict Invariance

As can be seen in Table 7, the fit indexes obtained as a result of multi-group RMSEA, CFI, NNFI and $\Delta RMSEA$, ΔCFI , ΔNFI values obtained as a result of the CFI difference test can be interpreted for each step as follows. According to the results, it is seen that the structural invariance is provided, and this finding shows that the measured structures use the same conceptual perspectives in responding to the questionnaire items of the adolescents and adults; males and females. The result regarding the metric invariance indicates that the factor structures of the variables taken in the model are the same in the adolescent and adult; male and female groups. It is confirmed that strong invariance is provided, and the constant number in the regression equations formed for the items is invariant between the groups. In the last stage, considering the $\Delta RMSEA$, ΔCFI , ΔNFI values calculated with the fit indexes, it is accepted that the error terms regarding the items forming the measurement tool are invariant between the comparison groups. Hierarchical analysis results, factor structure, and pattern of the questionnaire, factor loads, regression constants, and error variances are seen to be invariant for the adolescent and adult; male and female groups.

Differential Item Functioning for Sex and Age

In order to provide evidence for the validity of the items included in the measurement tools used in the study, it was examined whether each item showed bias according to the sex and age variables. In this context, it has been examined how the responses given to the items according to sex and age variables with the help of logistic functions by using the Mantel-Haenszel technique, which is based on the Item Response Theory. The change in the likelihood that individuals with the same level of ability will

respond correctly to an item is based on two reasons item bias or differences of actual knowledge, skill, etc. Determining whether items give DIF is a more commonly used technique, as it is seen as a more objective approach to bias (Doğan & Öğretmen, 2008).

DIF results for sex

As a result of the determination of males as focus groups and females as reference groups; the comparison variable is accepted as the score obtained from the questionnaire's each item. The chi-square values, significance values, and statistics showing the level of DIF obtained as a result of the analysis are presented in Table 8.

Table 8. DIF Results for Sex

Item	χ^2	Error	CI Lower	CI Upper	Class
Q1	4.03*	0.16	0.01	0.31	AA
Q2	74.63***	0.66	0.50	0.81	CC+
Q3	6.22**	0.20	0.04	0.36	BB+
Q4	0.18	-0.03	-0.17	0.11	AA
Q5	5.16**	0.20	0.03	0.36	AA
Q6	12.73***	0.27	0.11	0.42	BB+
Q7	6.56**	-0.28	-0.48	-0.09	BB-
Q8	12.03***	-0.28	-0.45	-0.11	BB-
Q9	2.54	-0.12	-0.28	0.04	AA
Q10	4.61*	-0.20	-0.39	-0.02	BB-
Q11	2.14	-0.14	-0.31	0.04	AA
Q12	2.49	-0.15	-0.33	0.02	AA

Notes: n= 1825, *= $p < .05$, **= $p < .01$, ***= $p < .001$, CI= Confidence Interval

When Table 8 is analyzed, it is seen that the χ^2 values obtained for all the items except Q2 coded item among the items in the measurement tool are not statistically significant in the determined degree of freedom. In other words, in the AQ-SF it was found that six items showed negligible (AA) DIF, six items showed medium (BB) DIF and one item showed high (CC) DIF (Güzeller, Eser & Aksu, 2018). This result explains that the 12 items in the measurement tool do not work in favor of female or male participants and the results obtained from the measurement tool didn't differ for both groups. However, it was determined that the Q2 coded item in the measurement tool showed DIF in favor of the focus group at the CC + (high) level. In order to say that an item produces biased results for or against one of the subgroups in the study universe, it should show at least C (high) DIF (Koyuncu, Aksu, & Kelecioğlu, 2018). Therefore, it is necessary to examine whether the item is biased according to the sex variable. The characteristic curve obtained for the second item determined to show a high level of DIF is shown in Figure 5.

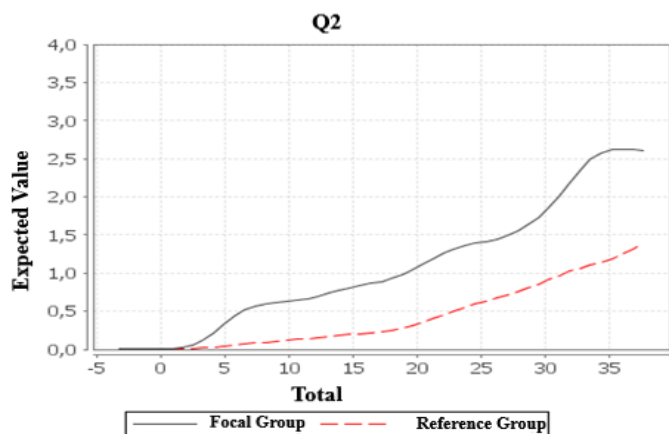


Figure 5. Item Characteristic Curve for the Related Item

When Figure 5 is examined, it is determined that the scores obtained from this item show DIF in favor of male participants who are determined as the focus group at all ability levels. In other words, Q2 measures aggression differently for males from aggression for females. Item impact means that respondents in different groups answer one item correctly express the real differences in their probabilities. This difference is explained by the knowledge or experience that one of the groups has (Gök, Kelecioğlu, Doğan & 2010). Item impact is also evident when examinees from different groups have differing probabilities of responding correctly to (or endorsing) an item because there are true differences between the groups in the underlying ability being measured by the item (Zumbo, 1999).

DIF results for age

It was analyzed whether each item differs depending on the age variable. As a result of the determination of the fewer adolescents as focus groups and adults as reference groups, the comparison variable is accepted as the score obtained for each questionnaire item. The chi-square values, significance values, and statistics showing the level of DIF obtained as a result of the analysis are presented in Table 9.

Table 9. DIF Results for Age

Item	χ^2	Error	CI Lower	CI Upper	Class
Q1	7.13**	0.11	0.02	0.19	AA
Q2	1.76	0.05	-0.03	0.13	AA
Q3	30.85***	0.23	0.15	0.32	BB+
Q4	0.99	-0.04	-0.11	0.04	AA
Q5	0.22	0.02	-0.07	0.10	AA
Q6	14.53***	0.16	0.07	0.24	AA
Q7	3.47	-0.10	-0.20	0.00	AA
Q8	3.45	0.08	-0.01	0.17	AA
Q9	3.88*	-0.09	-0.16	-0.01	AA
Q10	1.63*	-0.06	-0.16	0.04	AA
Q11	29.00***	-0.24	-0.33	-0.15	BB-
Q12	6.35**	-0.12	-0.21	-0.02	AA

Notes: n= 1825, *= $p < .05$, **= $p < .01$, ***= $p < .001$, CI= Confidence Interval

Table 9 indicated that the χ^2 values obtained for all the items in the measurement tool are not statistically significant in the determined degree of freedom. In the AQ-SF, it was found that 10 items showed negligible (AA) DIF and two items showed medium (BB) DIF (Güzeller, Eser & Aksu, 2018). This result explains that the 12 items in the measurement tool do not work in favor of female or male students and the results obtained from the measurement tool didn't differ for both groups.

Accordingly, when the results obtained regarding the reliability and validity of the measurement tool were analyzed as a whole, it was determined that the aggressive characteristics of the adolescents and adults were measured with a valid and reliable measurement tool.

DISCUSSION and CONCLUSION

This study aims to make the adaptation of the Aggression Questionnaire Short Form-in Turkish with adolescent and adult samples. In order to test the construct validity of the questionnaire, PA was conducted. The four factor structure of the questionnaire was confirmed via PA, which was defined as the best way to determine factor numbers to retain (Ledesma & Valero-Mora, 2007). This analysis has been indicated consistently accurate in determining the threshold for significant components, variable loadings, and analytical statistics when decomposing a correlation matrix (Franklin, Gibson, Robertson, Pohlmann, & Fralish, 1995). Moreover, the factor structure of the questionnaire was tested through EFA. EFA findings indicated that the questionnaire has a four-factor structure of adolescent and adult samples similar to the original form of the questionnaire (Bryant & Smith, 2001). Additionally, the results of the CFA, which were conducted for both adolescents and adults confirmed the four-factor structure of the questionnaire. These results also parallel the findings of Braynt & Smith (2001) that about the CFA for

the original form of the questionnaire. The four factors structure of the questionnaire was also approved via CFA in the study, which includes Spanish (Morales-Vives, Codorniu-Raga, & Andreu Vigil-Colet, 2005), Egyptian, Omani (Abd-El-Fattah, 2013), Dutch (Hornsveld et al. 2009) adolescents. In studies conducted with adults by Maxwell (2007) and Vitoratou et al. (2009), CFA results indicated four factor structure. McKay, Perry, and Harway (2016) tested both unidimensional and four-factor models of AQ-SF and reported limited evidence for unidimensional models beside four-factor model supported results. Different from the studies which support four-factor structure of AQ-SF via CFA, Kožený, Tišanská, & Csémy (2017) reported one component, Reyna et al. (2011) indicated two-component structure for AQ-SF.

For validity analysis, convergent and divergent validity of AQ-SF was examined. The moderate and significant correlation of AQ-SF scores with trait anger and social problem-solving scores in adolescent and adult applications confirmed the construct validity of AQ-SF. A significant and moderate correlation between AQ scores and trait anger level was reported by Wang et al. (2018). Similarly, Kuzucu (2016) reported a significant correlation between AQ-SF scores and social problem-solving scores. These results are not only evidence for convergent and divergent validity of the AQ-SF, but also show the correlation of questionnaire both with ill-being and well-being variables.

In terms of reliability, internal consistency and test-retest reliability scores were calculated. While the Cronbach Alpha scores in the present study are acceptable similar to the original form (Bryant & Smith, 2001), the test-retest reliability scores are higher than the original form of the questionnaire (Buss and Perry, 1992) and most of the previous studies (Harris, 1997; Surís, Borman, Lind, & Kashner, 2007; Webster et al. 2014). The differences were found between the responses of the individuals with low and high scores in adolescent and adult groups.

To test invariant measurement models of the AQ-SF between different sex and age groups, MI of the questionnaire was also tested in terms of sex and age. In the present study, there is sex invariance for measurement through AQ-SF between males and females. It is consistent with the other findings in the literature. Sex differences about the type and magnitude of aggressive behaviors seem as common results of the studies (Björkqvist, Österman, & Lagerspetz, 1994; Eron, Huesmann, Dubow, Romanoff, & Yarmel, 1987). The invariance of sex was also mentioned by Bryant & Smith (2001). Moreover, among Greek adults (Vitoratou et al. 2009) and federal offenders (Diamond & Magaletta, 2006), sex invariance was reported. Different from the sex invariance results of the present study, partial sex MI of AQ-SF for Argentinean (Reyna et al. 2011), Egyptian (Abd-El-Fattah, 2013) adolescents, and adolescents from Singapour (Ang, 2007) and Liverpool (McKay et al. 2016) was reported. The previous studies tested and showed MI of the questionnaire also with several samples from similar demographic backgrounds (Ang, 2007; Bryant & Smith, 2001; Vitoratou et al., 2009).

There is an age invariance for measurement through AQ-SF between adolescents and adults. In literature, adolescents are reported no more aggressive than adults. Adults are not less hostile than adolescents, but they use different and more latent means of aggression (Björkqvist et al., 1994). Torregrosa et al. (2020) showed age invariance between 8-9 and 10-11 aged children. Moreover, longitudinal studies emphasized the continuity of aggressive behaviors through adolescence to adulthood (Eron et al. 1987; Huesmann, Eron, Lefkowitz, & Walder, 1984, Huesmann, Eron, & Dubow, 2002). The present findings confirmed the invariant measurement of aggression between adolescents and adults via AQ-SF. However, to our knowledge, there is no study in which age invariance was tested for AQ-SF among adolescents and adults.

The DIF analysis for sex showed that the item of AQ-SF coded as Q2 'There are people who pushed me so far that we came to blows' measure aggression in a biased way between the sexes. With the aim of explaining whether this difference is item bias or true difference, expert opinion was obtained. The expert group interview conducted with the consideration of it is a physical aggression related item and they focused that it measures physical aggression in favor of males. In conclusion, this difference should be accepted as the real difference due to biological reasons; as a result, males are more likely to respond to this item. Similar to the DIF results and experts' opinions about Q2, it was reported that males are more physically aggressive than females related to the testosterone level (Björkqvist, 2018). Despite the

focused age group, Lansford et al., (2012) reported more physical aggression among boys than girls, consistently across nine different countries.

The DIF analysis for age supported that there is no bias in the AQ-SF items for adolescents and adults. In addition to the power of the questionnaire in terms of factorial structural that MI for age results showed, DIF results reinforced this power by items for different age groups. All items of the questionnaire measure aggression in an unbiased way for age. This result has support in the literature. With the evidence from longitudinal studies (Eron et al. 1987; Huesmann et al., 1984; Huesmann, Eron, & Dubow, 2002) it is known that aggression has persisted from adolescence to adulthood. Moreover, aggression is a topic that is investigated in the life span approach. Several studies were conducted with different aged group participants, from toddlerhood to old-adulthood (Liu, Lewis, & Evans, 2013). This wide range of studies of aggression, both in terms of time and age could explain the power of AS-QF about giving reliable measurements for different ages.

Despite the contributions to literature, this study has limitations. The results for the AQ-SF were not compared with BPAQ (29 item version). In the current study, participants came from a nonclinical sample. In further studies, MI for clinical and nonclinical samples can be tested. In addition to the cross-sectional data set, testing sex and age invariance in aggression with longitudinal data is another suggestion for the researchers. All results for validity and reliability tests confirmed four factors and 12 items structure of the questionnaire. The findings also presented that the AQ-SF is a valid and reliable questionnaire, and it can be used for male, female, adolescent, and adult populations.

REFERENCES

- Abd-El-Fattah, S. M. (2013). A cross-cultural examination of the Aggression Questionnaire–Short Form among Egyptian and Omani Adolescents. *Journal of Personality Assessment, 95*(5), 539-548.
- Adıgüzel, V., Özdemir, N., & Şahin, Ş. K. (2019). Childhood traumas in euthymic bipolar disorder patients in Eastern Turkey and its relations with suicide risk and aggression. *Nordic Journal of Psychiatry, 73*(8), 490-496.
- Akbaş, U., Karabay, E., Yıldırım-Seheryeli, M., Ayaz, A., ve Demir, Ö. O. (2019). Türkiye ölçe araçları dizininde yer alan açımlayıcı faktör analizi çalışmalarının paralel analiz sonuçları ile karşılaştırılması. *Kuramsal Eğitimbilim Dergisi, 12*(3), 1095-1123.
- Aksu, G., Eser, M. T., & Güzeller, C. O. (2017). *Açımlayıcı ve doğrulayıcı faktör analizi ile yapısal eşitlik modeli uygulamaları*. Detay Yayıncılık.
- Ang, R. P. (2007). Factor structure of the 12-item aggression questionnaire: Further evidence from Asian adolescent samples. *Journal of Adolescence, 30*(4), 671-685.
- Bernstein, I. H., & Gesn, P. R. (1997). On the dimensionality of the Buss/Perry Aggression Questionnaire. *Behaviour Research and Therapy, 35*(6), 563-568.
- Björkqvist, K., Österman, K., & Lagerspetz, K. M. (1994). Sex differences in covert aggression among adults. *Aggressive Behavior, 20*(1), 27-33.
- Björkqvist, K. (2018). Gender differences in aggression. *Current Opinion in Psychology, 19*, 39-42.
- Borders, A., Earleywine, M., & Jajodia, A. (2010). Could mindfulness decrease anger, hostility, and aggression by decreasing rumination? *Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 36*(1), 28-44.
- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: A measurement model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality, 35*(2), 138-167.
- Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 21*(4), 343.
- Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology, 63*(3), 452.
- Büyüköztürk, Ş. (2007). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem A Yayıncılık.
- Carver, C. S., Sinclair, S., & Johnson, S. L. (2010). Authentic and hubristic pride: Differential relations to aspects of goal regulation, affect, and self-control. *Journal of Research in Personality, 44*(6), 698-703.
- Chang, L., Schwartz, D., Dodge, K. A., & McBride-Chang, C. (2003). Harsh parenting in relation to child emotion regulation and aggression. *Journal of Family Psychology, 17*(4), 598.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255.

- Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon & N. Eisenberg (Ed.), *Handbook of child psychology: Social, Emotional, and Personality Development* (pp. 779-862). Hoboken, NJ, US: John Wiley & Sons Inc.
- D’Zurilla, T. J., Nezu, A. M. & Maydeu-Olivares, A. (2002). *Manual for the Social Problem Solving Inventory-revised (SPSI-R)*. North Tonawanda: Multi-Health Systems.
- De Zavala, A. G., Cichocka, A., Eidelson, R., & Jayawickreme, N. (2009). Collective narcissism and its social consequences. *Journal of Personality and Social Psychology*, 97(6), 1074.
- Diamond, P. M., & Magaletta, P. R. (2006). The short-form Buss-Perry Aggression Questionnaire (BPAQ-SF) a validation study with federal offenders. *Assessment*, 13(3), 227-240.
- Diamond, P. M., Wang, E. W., & Buffington-Vollum, J. (2005). Factor structure of the Buss-Perry Aggression Questionnaire (BPAQ) with mentally ill male prisoners. *Criminal Justice and Behavior*, 32(5), 546-564.
- Doğan, N., & Öğretmen, T. (2008). Degisen madde fonksiyonunu belirlemede mantel - Haenszel, ki-kare ve lojistik regresyon tekniklerinin karsilastirilmasi. *Egitim ve Bilim*, 33(148), 100-112.
- Emerson, S. D., Guhn, M., & Gadermann, A. M. (2017). Measurement invariance of the Satisfaction with Life Scale: reviewing three decades of research. *Quality of Life Research*, 26(9), 2251-2264.
- Erkuş, A. (2012). *Psikolojide ölçme ve ölçek geliştirme*. Ankara: Pegem Akademi Yayınları.
- Eron, L. D., Huesmann, L. R., Dubow, E., Romanoff, R., & Yarmel, P. W. (1987). Aggression and its correlates over 22 years. *Childhood Aggression and Violence*, 249-262.
- Eskin, M., & Aycan, Z. (2009). The adaptation of the revised social problem solving inventory into Turkish: A reliability and validity analysis. *Turkish Journal of Psychology*, 12(23), 11-13.
- Evren, C., Çınar, Ö., Güleç, H., Çelik, S., & Evren, B. (2011). The validity and reliability of the Turkish version of the Buss-Perry’s Aggression Questionnaire in male substance dependent inpatients. *The Journal of Psychiatry and Neurological Sciences*, 24(4), 283-295.
- Ferrando, P. J., & Lorenzo-Seva, U. (2017). 10 años del programa FACTOR: Una revisión crítica de sus orígenes, desarrollo y líneas futuras. *Psicothema*, 29(2), 236-240.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.
- Fraenkel, J.R., Wallen, N.E., & Hyun, H.H. (2012). *How to design and evaluate research in education* (Eight Edition). New York: McGraw-Hill
- Franklin, S. B., Gibson, D. J., Robertson, P. A., Pohlmann, J. T., & Fralish, J. S. (1995). Parallel Analysis: a method for determining significant principal components. *Journal of Vegetation Science*, 6(1), 99-106.
- Gök, B., Kelecioğlu, H., & Dogan, N. (2010). The Comparison of Mantel-Haenszel and Logistic Regression Techniques in Determining the Differential Item Functioning. *Egitim ve Bilim-Education and Science*, 35(156), 3-16.
- García-León, A., Reyes, G. A., Vila, J., Pérez, N., Robles, H., & Ramos, M. M. (2002). The Aggression Questionnaire: A validation study in student samples. *The Spanish Journal of Psychology*, 5(1), 45-53.
- Gerevich, J., Bácskai, E., & Czobor, P. (2007). The generalizability of the Buss-Perry Aggression Questionnaire. *International Journal of Methods in Psychiatric Research*, 16(3), 124-136.
- Guo, B., Kaylor-Hughes, C., Garland, A., Nixon, N., Sweeney, T., Simpson, S., ... & Morriss, R. (2017). Factor structure and longitudinal measurement invariance of PHQ-9 for specialist mental health care patients with persistent major depressive disorder: Exploratory Structural Equation Modelling. *Journal of Affective Disorders*, 219, 1-8.
- Gustavsson, J. P., Eriksson, A. K., Hilding, A., Gunnarsson, M., & Östenson, C. G. (2008). Measurement invariance of personality traits from a five-factor model perspective: multi-group confirmatory factor analyses of the HP5 inventory. *Scandinavian Journal of Psychology*, 49(5), 459-467.
- Harris, J. A. (1995). Confirmatory factor analysis of the Aggression Questionnaire. *Behaviour Research and Therapy*, 33(8), 991-993.
- Harris, J. A. (1997). A further evaluation of the aggression questionnaire: Issues of validity and reliability. *Behaviour Research and Therapy*, 35(11), 1047-1053.
- Hinshaw, S. P. (1987). On the distinction between attentional deficits/hyperactivity and conduct problems/aggression in child psychopathology. *Psychological Bulletin*, 101(3), 443.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185.
- Hornsveld, R. H., Muris, P., Kraaimaat, F. W., & Meesters, C. (2009). Psychometric properties of the aggression questionnaire in Dutch violent forensic psychiatric patients and secondary vocational students. *Assessment*, 16(2), 181-192.
- Huesmann, L. R., Eron, L. D., & Dubow, E. F. (2002). Childhood predictors of adult criminality: are all risk factors reflected in childhood aggressiveness?. *Criminal Behaviour and Mental Health*, 12(3), 185-208.
- Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). Stability of aggression over time and generations. *Developmental Psychology*, 20(6), 1120-1134.

- Iurino, K., & Saucier, G. (2020). Testing measurement invariance of the Moral Foundations Questionnaire across 27 countries. *Assessment, 27*(2), 365-372.
- Johnson, S. L., Carver, C. S., & Joormann, J. (2013). Impulsive responses to emotion as a trans diagnostic vulnerability to internalizing and externalizing symptoms. *Journal of Affective Disorders, 150*(3), 872-878.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International Inc, Chicago.
- Kang, J. H., Lim, C. H., Suh, Y. I., Gang, A. C., & Pedersen, P. M. (2020). Establishing a web-based measurement of aggression (WTCRTT): Examining the validity of a modified Taylor's competitive reaction time test. *International Journal of Applied Sports Sciences, 32*(1), 37-48.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford Publications.
- Koyuncu, İ., Aksu, G., & Kelecioğlu, H. (2018). Mantel-Haenszel, Lojistik Regresyon ve Olabilirlik Oranı Değişen Madde Fonksiyonu İnceleme Yöntemlerinin Farklı Yazılımlar Kullanılarak Karşılaştırılması. *Elementary Education Online, 17*(2), 909-925.
- Kožeňý, J., Tišanská, L., & Csémy, L. (2017). A Rasch Analysis of The Buss-Perry Aggression Questionnaire-Short Form: An evidence from Czech adolescents' sample. *Ceskoslovenska Psychologie, 61*(3), 257-266.
- Kuzucu, Y. (2016). Do anger control and social problem-solving mediate relationships between difficulties in emotion regulation and aggression in adolescents? *Educational Sciences: Theory & Practice, 16*(3).
- Kühn, S., Kugler, D. T., Schmalen, K., Weichenberger, M., Witt, C., & Gallinat, J. (2019). Does playing violent video games cause aggression? A longitudinal intervention study. *Molecular Psychiatry, 24*(8), 1220-1234.
- Lansford, J. E., Skinner, A. T., Sorbring, E., Giunta, L. D., Deater-Deckard, K., Dodge, K. A., ... & Uribe Tirado, L. M. (2012). Boys' and girls' relational and physical aggression in nine countries. *Aggressive Behavior, 38*(4), 298-308.
- Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation, 12*(2), 1-11.
- Liu, J., Lewis, G., & Evans, L. (2013). Understanding aggressive behaviour across the lifespan. *Journal of psychiatric and mental health nursing, 20*(2), 156-168.
- Madran, H. A. D. (2012). Buss-Perry Saldırganlık Ölçeği'nin Türkçe formunun geçerlik ve güvenilirlik çalışması. *Türk Psikoloji Dergisi, 24*(2), 1-6.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*(4), 370.
- Maxwell, J. P. (2007). Development and preliminary validation of a Chinese version of the Buss-Perry Aggression Questionnaire in a population of Hong Kong Chinese. *Journal of Personality Assessment, 88*(3), 284-294.
- McKay, M. T., Perry, J. L., & Harvey, S. A. (2016). The factorial validity and reliability of three versions of the Aggression Questionnaire using Confirmatory Factor Analysis and Exploratory Structural Equation Modelling. *Personality and Individual Differences, 90*, 12-15.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-543.
- Meyer, J. P. (2018). *JMETRİK ile Ölçme Uygulamaları*. (Çev. C. O. Güzeller, M. T. Eser, G. Aksu). Maya Akademi (Orijinal yayın tarihi 2014).
- Milfont, T. L., & Fischer, R. (2015). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*(1), 111-130.
- Morales-Vives, F., Codorniu-Raga, M. J., & Vigil-Colet, A. (2005). Características psicométricas de las versiones reducidas del cuestionario de agresividad de Buss y Perry. *Psicothema, 17*(1), 96-100.
- Moyer, K. E. (1982). *The origins of aggression in child nurturance* (pp. 243-260). Springer, Boston, MA.
- Orpinas, P., & Frankowski, R. (2001). The Aggression Scale: A self-report measure of aggressive behavior for young adolescents. *The Journal of Early Adolescence, 21*(1), 50-67.
- Önen, E. (2016). Saldırganlık Ölçeği'nin Psikometrik Niteliklerinin Türk Ergenleri İçin İncelenmesi. *Türk Psikolojik Danışma ve Rehberlik Dergisi, 4*(32).
- Özer, A. K. (1994). Sürekli öfke ve öfke ifade tarzı ölçekleri ön çalışması. *Türk Psikoloji Dergisi, 9*(31), 26-35.
- Özdemir, Y., Vazsonyi, A. T., & Cok, F. (2017). Parenting processes, self-esteem, and aggression: A mediation model. *European Journal of Developmental Psychology, 14*(5), 509-532.
- Palmstierna, T., & Wistedt, B. (1987). Staff observation aggression scale, SOAS: presentation and evaluation. *Acta Psychiatrica Scandinavica, 76*(6), 657-663.
- Podubinski, T., Lee, S., Hollander, Y., & Daffern, M. (2017). Patient characteristics associated with aggression in mental health units. *Psychiatry Research, 250*, 141-145.
- Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., ... & Liu, J. (2006). The reactive-proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in

- adolescent boys. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, 32(2), 159-171.
- Reyna, C., Sanchez, A., Ivacevich, M. G. L., & Brussino, S. (2011). The Buss-Perry Aggression Questionnaire: construct validity and gender invariance among Argentinean adolescents. *International Journal of Psychological Research*, 4(2), 30-37.
- Revelle, W. (2018). *Psych: procedures for personality and psychological research*. Northwestern University, Evanston.
- Sexton, M. B., Davis, A. K., Buchholz, K. R., Winters, J. J., Rauch, S. A. M., Yzquibell, M., . . . Chermack, S. T. (2019). Veterans with recent substance use and aggression: PTSD, substance use, and social network behaviors. *Psychological Trauma: Theory, Research, Practice, and Policy*, 11(4), 424-433.
- Singh, S. (2017). Understanding aggression among youth in the context of mindfulness. *Indian Journal of Health and Wellbeing*, 8(11), 1377-1379.
- Sorsdahl, K., Stein, D. J., & Myers, B. (2017). Psychometric properties of the Social Problem Solving Inventory-Revised Short-Form in a South African population. *International Journal of Psychology*, 52(2), 154-162.
- Spielberger, C. D. (1985). The experience and expression of anger: Construction and validation of an anger expression scale. *Anger and hostility in cardiovascular and behavioral disorders*, 5-30.
- Spielberger, C. D. (1988). *Manual for the State-Trait Anger Expression Scale (STAX)*. 1988. Odessa, FL: Psychological Assessment Resources.
- Spielberger, C. D., Johnson E.H, Russell SF, Crane RJ, Jacobs GA, & Worden TJ. (1985). The experience and expression of anger: construction and validation of an anger expression scale. In: Chesney MA, Rosenman R.H, eds. *Anger and Hostility in Cardiovascular and Behavioral Disorders*. New York, NY: Hemisphere, 5-30.
- Şahin, C. (2018). *Bireyi tanıma teknikleri: Psikolojik testler, test dışı teknikler*. Ankara: Pegem Akademi.
- Surís, A., Borman, P. D., Lind, L., & Kashner, T. M. (2007). Aggression, impulsivity, and health functioning in a veteran population: equivalency and test-retest reliability of computerized and paper-and-pencil administrations. *Computers in Human Behavior*, 23(1), 97-110.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistic* (6th ed.). USA: Pearson.
- Torregrosa, M. S., Gómez-Núñez, M. I., Inglés, C. J., Ruiz-Esteban, C., Sanmartín, R., & García-Fernández, J. M. (2020). Buss and Perry Aggression Questionnaire-Short Form in Spanish Children. *Journal of Psychopathology and Behavioral Assessment*, 1-16.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-70.
- Vitoratou, S., Ntzoufras, I., Smyrnis, N., & Stefanis, N. C. (2009). Factorial composition of the Aggression Questionnaire: A multi-sample study in Greek adults. *Psychiatry Research*, 168(1), 32-39.
- Wang, X., Yang, L., Yang, J., Gao, L., Zhao, F., Xie, X., & Lei, L. (2018). Trait anger and aggression: A moderated mediation model of anger rumination and moral disengagement. *Personality and Individual Differences*, 125, 44-49.
- Webster, G. D., DeWall, C. N., Pond Jr, R. S., Deckman, T., Jonason, P. K., Le, B. M., ... & Smith, C. V. (2014). The brief aggression questionnaire: Psychometric and behavioral evidence for an efficient measure of trait aggression. *Aggressive Behavior*, 40(2), 120-139.
- Webster, S. D., Mann, R. E., Thornton, D., & Wakeling, H. C. (2007). Further validation of the Short Self-esteem Scale with sexual offenders. *Legal and Criminological Psychology*, 12(2), 207-216.
- Wu, A. D., Li, Z. & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation*, 12(1), 1-26.
- Zumbo, B. (1999). A handbook on the theory and methods of differential item functioning (DIF). *Ottawa: National Defense Headquarters, June*, 1-57.

Saldırganlık Ölçeği Kısa Formu Türkçe Versiyonunun Psikometrik Özelliklerinin İncelenmesi: Cinsiyet ve Yaş için Ölçme Eşdeğerliği ve Değişen Madde Fonksiyonu

Giriş

Saldırganlık; biyolojik, psikolojik, sosyal ve kültürel faktörlerin bir arada etkili olduğu çok boyutlu bir yapıda gelişmektedir (Vitoratou, Ntzoufras, Smyrnis, ve Stefanis, 2009). Birçok kuramcı ve uygulamacı saldırganlık davranışını ve ilişkili olduğu diğer davranışları açıklamaya çalışmaktadır (Chang, Schwartz, Dodge, ve McBride-Chang, 2003; Sexton ve diğerleri, 2019).

Saldırganlığı ölçmek için çeşitli ölçme araçları geliştirilmiş olup (Buss ve Perry, 1992; Kang, Lim, Suh, Gang, ve Pedersen, 2020; Orpinas ve Frankowski, 2001; Palmstierna ve Wistedt, 1987; Raine ve diğerleri, 2006), bunlar arasında en sık kullanılan ölçme aracı Buss ve Perry (1992) tarafından geliştirilen Buss-Perry Saldırganlık Ölçeğidir. Ölçek fiziksel saldırganlık, sözel saldırganlık, öfke ve düşmanlık alt boyutlarından oluşan 29 maddelik bir ölçektir. Ölçeğin psikometrik özellikleri farklı yöntem ve örneklerle test edilmiş ve sonuçlar dört faktörlü yapıyı doğrulamıştır (Bernstein ve Gesn, 1997; García-León ve diğerleri, 2002; Gerevich, Bácskai, ve Czobor 2007; Harris, 1997; Reyna et al. 2011). Bununla birlikte çalışmaların birçoğunda bazı maddeler ölçekten çıkarıldığında daha iyi uyum değerleri ve faktör yükleri elde edilmiştir (Bernstein ve Gesn, 1997; Gerevich ve diğerleri, 2007; Harris, 1995;).

Bryant ve Smith (2001) daha rafine bir ölçme aracı yaratmak için Buss ve Perry Saldırganlık ölçeğinin en iyi çalışan 12 maddesini belirlemiştir. Bu yeni kısa form dört faktörün her birinde üç madde olacak şekilde kısaltılmıştır. Kısa form hem Arjantin, Hollanda ve Çin gibi farklı ülkelerde (Reyna et al, 2011; Maxwell (2007) hem de tutuklular ve ruh sağlığı bozukluğu olanlar gibi farklı örneklerde (Buffington-Vollum, 2005; Diamond ve Magaletta, 2006; Diamond, Wang) test edilmiştir, sonuçlar ölçeğin iç tutarlılığının ve uyum değerlerinin yüksek olduğunu göstermektedir.

Saldırganlık Ölçeği'nin kısa formunun yeterli psikometrik özelliklere sahip olmasının yanı sıra, ölçek diğer ruh sağlığı değişkenleri ile de yüksek ilişki göstermektedir. Yapılan çalışmalar saldırganlığın kolektif narsizim (De Zavala, Cichocka, Eidelson, ve Jayawickreme, 2009), ruminasyon (Borders, Earleywine, ve Jajodia, 2010), kibirlilik (Carver, Sinclair, ve Johnson; 2010), anksiyete ve alkol kullanımıyla (Johnson, Carver, ve Joormann, 2013) ilişkili olduğunu ortaya koymuştur.

Türkiye'de de Buss-Perry Saldırganlık Ölçeği'nin 29 maddelik formu lise, üniversite öğrencileri ve madde bağımlıları ile yapılan araştırmalarda kullanılmıştır. Elde edilen araştırma bulguları ölçeğin geçerli ve güvenilir bir araç olduğunu göstermektedir (Evren, Çınar, Güleç, Çelik, ve Evren, 2011; Madran, 2012; Önen, 2016). Bununla birlikte ölçeğin kısa formunun psikometrik özellikleri Türkiye'de çalışılmamıştır

Saldırganlık ölçeğinin kısa formu farklı kültürlerde, cinsiyette, klinik ve klinik olmayan gruplarda kullanılabilir. Ölçek hem psikolojik iyi oluş hem de psikolojik sorunlarla korelasyon göstermektedir. Ölçeğin bu özellikleri dikkate alındığında kısa formunun dilimize kazandırılması önem taşımaktadır. Bu çalışmanın amacı, Saldırganlık Ölçeği Kısa Formu'nun Türkçedeki psikometrik özelliklerini ergen ve yetişkinler ile kadınlar ve erkekler için test etmektir.

Yöntem

Çalışma 15-18 yaşları arasındaki 778 ergen ve 19-44 yaşları arasındaki 1067 yetişkin katılımcı ile gerçekleştirilmiştir. Ergen çalışması için ölçeğin kısa formu beş farklı lisede öğrenim gören toplamda 778 öğrenciye uygulanmıştır. İlk uygulamada 383 öğrenciyle çalışılmış, elde edilen veri üzerinde Paralel Analiz (PA) ve Açıklayıcı Faktör Analizi (AFA) yapılmıştır. İkinci uygulamada 395 öğrenciye ulaşılmış ve bu veriler üzerinde Doğrulamalı Faktör Analizi (DFA) yapılmıştır. Yetişkin uygulamaları

İçin toplamda 1067 yetişkinle çalışılmıştır. Katılımcılar Aydın Adnan Menderes Üniversitesi, Ege Üniversitesi ve Ankara Üniversitesinde pedagojik formasyon eğitimi alan kişiler ile halk eğitim merkezlerinde kurslara katılan kursiyerlerden oluşmaktadır. Veriler yetişkinlerden iki farklı uygulamayla toplanmıştır. İlk uygulamada 648 katılımcıya erişilmiş, bu katılımcılardan alınan veriler PA ve AFA için kullanılmıştır. İkinci uygulamada toplanan 391 veri ile ise DFA yapılmıştır.

Veri analizleri için SPSS 25.0 (SPSS Inc.), Factor Analysis 10.10 (Fernando ve Lorenza-Seva, 2017), Lisrel 8.80 (Jöreskog ve Sorbom, 1993) ve jMetrik Version 4.1.1 istatistik paket programları kullanılmıştır. Ölçeğin yapı geçerliği PA, AFA, DFA aracılığıyla test edilmiştir. Ölçeğin yakınsak geçerliği için sürekli öfkeyle, iraksak geçerlik için ise sosyal problem çözmeyle ilişkisine bakılmıştır. Güvenirlik analizleri kapsamında, madde-toplam korelasyonu, test tekrar test güvenirliliği, Cronbach Alpha ve McDonald Omega iç tutarlık değerleri hesaplanmıştır. Ek olarak ölçek maddelerinin en yüksek ve düşük %27 grupta ayrışması t-test ile test edilmiştir. Ölçeği ölçüm değişmezliği erkek-kadın ve ergen-yetişkin örneklerle test edilmiştir.

Sonuç ve Tartışma

Sonuçlar, Saldırganlık Ölçeği Kısa Formu'nun Türkçede dört faktörden oluşan, güvenirliliğe sahip, cinsiyetler arası ölçüm farkı olmayan, ergen ve yetişkinler için saldırganlık ölçümünde kullanılabilecek bir ölçüm aracı olduğunu göstermiştir.

Ölçeğin yapı geçerliğini test etmek için paralel analiz kullanılmıştır. Faktör belirlemede en önemli yöntemlerden birisi olarak görülen PA (Horn, 1965) sonuçları, dört faktörlü yapının doğrulandığını göstermektedir. Ölçek çalışmalarında, yapı ve alt yapıların nasıl ve kaç tane olacağını belirleyebilmek için birden çok yöntem başvurulması önerilmektedir (Erkuş, 2012). Bu doğrultuda çalışma kapsamında PA'nın yanı sıra ölçeğin faktör yapısı AFA ile de incelenmiştir. Elde edilen sonuçlar hem ergen hem de yetişkin gruplarda dört faktörlü orijinal yapıyı doğrulamaktadır. Yamaç eğrisi grafikleri de ergen ve yetişkin gruplar için ölçeğin dört faktörlü yapıya sahip olduğunu desteklemektedir. Ergen ve yetişkin gruplarda uygulanan PA ve AFA sonuçları maddelerin ilgili faktörün altında yüksek değerle gruplandığını göstermektedir.

PA ve AFA'dan sonra, DFA kullanılmıştır. Birinci ve ikinci düzey DFA sonuçları da ölçeğin dört faktörlü yapıda iyi uyum değerlerine sahip olduğunu göstermektedir. Ölçek birinci ve ikinci düzey DFA'ya ek olarak, tek boyutlu DFA ile de test edilmiştir. Sonuçlar dört faktörlü yapısının daha iyi uyum değerlerine sahip olduğunu ortaya koymuştur.

Ölçeğin faktörleri arasındaki korelasyon ergen ve yetişkin örnekleme incelenmiş orta düzey yakın yada orta düzeyde korelasyon gösterdikleri bulunmuştur. Ölçeğin yakınsak ve iraksak geçerliği ergen ve yetişkin grupta test edilmiştir. Saldırganlığın sürekli öfkeyle beklenen yönde pozitif ve orta düzeyde anlamlı ilişkiye sahip olduğu bulunmuştur. Iraksak geçerlik ise sosyal problem çözme ile ilişkisine bakılmış ve ölçeğin iraksak geçerliğe sahip olduğu belirlenmiştir.

Ergen ve yetişkin grup için güvenirlilik analizleri kapsamında hesaplanan madde toplam korelasyonu değerlerinin 0,30'un üzerinde olduğu görülmüştür. Ölçeğin iç tutarlığı için incelenen Cronbach Alpha ve McDonald Omega değerlerinin hem ergen hem de yetişkin grup için yeterli düzeyde olduğu bulunmuştur. Ölçeğin iç tutarlık katsayısının hesaplanmasının dışında güvenirliliği değerlendirmek için ölçeğin test-tekrar test güvenirliliğine bakılmıştır. Ergen ve yetişkin gruplar için ölçeğin test-tekrar test güvenirliliğine sahip olduğu görülmüştür.

Bu çalışmada Buss ve Perry Saldırganlık Ölçeği Kısa Formunun ergen ve yetişkin örnekleme için yeterli düzeyde psikometrik özelliklere sahip olduğu belirlenmiştir. Elde edilen sonuçlar birbirleriyle ve literatürle tutarlılık göstermektedir (Braynt ve Smith, 2001; Hornsveld ve diğerleri, 2009; Maxwell, 2007; Morales-Vives, Codorniu-Raga, ve Andreu Vigil-Colet, 2005). Ölçeğin kısa formu üzerinden test edilen faktör yapısının, ölçeğin uzun formundaki gibi güçlü bir faktör yapısına sahip olduğunu kanıtlamaktadır. Sonuçlar Saldırganlık Ölçeği Kısa Formu'nun geçerli ve güvenilir bir ölçek olduğunu, erkek, kadın, ergen ve yetişkin gruplar için kullanılabileceğini göstermiştir.