The Toronto Empathy Questionnaire: Evaluation of Psychometric Properties among Turkish University Students

Tarık Totan*

Tayfun Doğan**

Fatma Sapmaz***

Suggested Citation:

Abstract

Problem statement: Today, it is widely accepted that empathy is a multidimensional factor that facilitates human relations. The common idea that empathy comprises more than one component has created diversity in the assessment of the said factor; many researchers have developed empathy scales that include different dimensions. However, uni-dimensional assessments minimize differences between assessments and develop an accepted core assessment tool.

Purpose of Study: The Toronto Empathy Questionnaire (TEQ) is a self-report style, uni-dimensional, 16-item, five-point Likert type scale developed to assess the empathy levels of individuals. The objective of this study is to adapt the TEQ into Turkish and to analyze its psychometric properties in a sample of Turkish university students. *Methods:* Study participants included 698 university students from Ege and Sakarya University. In the research, the Emphatic Tendency Scale and the Basic Empathy Scale were used as data collection tools along with the TEQ. In the adaptation of the questionnaire, a linguistic equivalence study was performed first. The psychometric properties of the TEQ were analyzed through item analysis, exploratory and confirmatory factor

* Correpsonding Author:Ph.D., Abant İzzet Baysal University, Faculty of Education, Department of Guidance and Counseling, totan@ibu.edu.tr.

** Ph.D., Sakarya University, Faculty of Education, Department of Guidance and Counseling, tayfun@tayfundogan.net.

*** Ph.D., Sakarya University, Faculty of Education, Department of Guidance and Counseling, sapmazfatma@gmail.com.

analysis, criterion-related validity, internal consistency and test-retest methods.

Results: As a result of the linguistic equivalency study, a positively significant correlation was found between the original form and the Turkish form of the questionnaire. The exploratory and confirmatory factor analysis results demonstrated that the questionnaire had a uni-dimensional structure. Within the scope of the criterion-related validity, positively significant correlations were found between the TEQ, Emphatic Tendency Scale and Basic Empathy Scale. The TEQ's internal consistency coefficient and test-retest reliability coefficient were .79 and .73 respectively. The findings of this study showed that the Turkish form of the TEQ was a valid and reliable assessment tool to assess the empathy levels of university students.

Key words: Empathy, Toronto Empathy Questionnaire (TEQ), adaptation, psychometric properties

Humankind, as a social entity, is in the position of communicating with others at every stage of life. However, the problems of establishing healthy interpersonal communication come into the forefront. Interpersonal relations have gradually gotten more complicated in conjunction with urbanization, technological progresses, changes in industry and complicating community structure (Bayam, Şimşek, & Dilbaz, 1995). Therefore, as a significant element of healthy communication, the ability to empathize is more important today. Given the related literature, the number of studies on communication and interpersonal relations has increased gradually and the attention is directed toward the concept of empathy.

Although research on empathy started at the end of the 19th century with the German definitions of "einfühlung," the process gained speed as American experimental psychologists translated the term from German to English in the early 20th century (Wispé, 1990). There are different definitions and functions of empathy in various psychological consultation theories (Marcia, 1990). However, if the matter in question is empathy, the first name that comes to mind is Carl R. Rogers, who conducted studies on empathy throughout his life and discussed empathy as an indispensable element of the psychological consultancy process (Dökmen, 1987). According to Rogers, empathy is that "a person puts himself/herself in other's place and sees events from his/her point of view, understands and feels his/her emotions and ideas accurately and communicates it to him/her" (Rogers, 1983; Dökmen, 1988). It is remarkable that Rogers emphasizes two dimensions of empathy in his definition (1983) cognitive and emotional. These dimensions are also emphasized in other empathy-focused studies (Stephan & Finlay, 1999; Engeler & Yargıç, 2007).

Emotional empathy is defined as a process of understanding other individuals' emotions and responding to and sharing such emotions. On the other hand, cognitive empathy is defined as the ability to perceive other individuals' feelings and understanding their emotions and ideas (Yüksel, 2004). As seen in the definitions, the

emotional dimension of empathy involves —in the simplest term— a person's appropriate responses to others' emotional responses such as sadness and anger. In other words, a person may feel sorry for sad people or treat them with tenderness and affection as a response to their sadness. Nonetheless, the fact that a person has a similar emotion does not depend on knowing the source of another's emotional response. That is to say, a person may feel an emotion similar to that of others even if he/she just sees or knows that others suffer without understanding why they suffer. On the contrary, the cognitive dimension of empathy is oriented to thinking and understanding and covers a cognition-based process. In this dimension, a person understands the situation or the emotions and ideas of others (Spreng, Kinnon, Mar, & Levine, 2009). İkiz (2006) says of the cognitive aspect of empathy, "a person understands of what others feel" (37). According to him, the emotional aspect of empathy includes "a person's feeling of what others feel" (37).

As a result, in the literature, some researchers highlight the cognitive aspect (Gallup & Platek, 2002) while others underline the emotional aspects (Mehrabian & Epstein, 1972). However, most agree that empathy comprises both cognitive and emotional components (Pecukonis, 1990; Shamay-Tsoory, Tomer, Goldsher, Berger, & Aharon-Peretz, 2004; Dadds et al., 2008). Due to the fact that the levels of awareness and understanding of people with a high level of emotional and cognitive empathy increase, it will be easier to establish a healthy and satisfactory communication (Dökmen, 1987, 2000).

Today, it is widely accepted that empathy is a multidimensional factor that facilitates human relations. The common idea that empathy comprises more than one component has created diversity in the assessment of the faculty and many researchers have developed empathy scales that include different dimensions (e.g., Hogan, 1969; Mehrabian & Epstein, 1972; Davis, 1980, 1983; Elliott et al., 1982; Özbay & Şahin, 2000; Lawrence, 2004; Wakabayashi et al., 2006; Muncer & Ling, 2006; Jolliffe & Farrington, 2006; Engeler & Yargıç, 2007; Dadds et al., 2008; Bora & Baysan, 2009; Kaya & Siyez, 2010). Based on the diversity of multidimensional empathy scales and the marked differences between the results of such scales, Spreng et al. (2009) developed a uni-dimensional assessment tool. The objective of developing the scale in question was not to return from multidimensional assessments to uni-dimensional assessments but to minimize the differences between assessments and develop a core assessment tool (Spreng et al., 2009). In line with this objective, Spreng et al. (2009) reviewed widely accepted empathy scales and developed the Toronto Empathy Questionnaire. The objective of this study is to adapt the TEQ into Turkish and to analyze its psychometric properties in a sample of Turkish university students.

Method

Participants

The research was conducted on a total of 698 university students from three different groups of participants. The first group comprised 33 university students studying at Ege University's Department of English Language and Literature and

participating in the linguistic equivalence study. The second group consisted of 588 university students [357 females (60.7 percent) and 231 males (39.3 percent)] from Sakarya University's Faculty of Education during the 2008-2009 academic year that were chosen by a convenience sampling method among nonprobability sampling techniques. The average age of this group was 20.60 (female $\bar{X}$= 20.22, S= 1.85; male $\bar{X}$= 21.20, S= 1.81). In this group, 269 were freshmen (45.75 percent, female *n*= 196, male *n*= 73), 100 were sophomores (17.01 percent, female *n*= 60, male *n*= 40), 77 were juniors (11.04 percent, female *n*= 39, male *n*= 38) and 142 were seniors (24.15 percent, female *n*= 62, male *n*= 80). The third group consisted of 77 students at Sakarya University's Faculty of Education that participated in the test-retest study.

Data Collection Tools

The Toronto Empathy Questionnaire (TEQ): Developed by Spreng et al. (2009), TEQ is a 16-item (eight items are scored negatively and eight items are scored positively) five-point Likert type scale. During the development of the TEQ, the researchers aimed to assess empathy as an emotional process, contrary to similar scales. To this end, they analyzed earlier assessment tools intended for assessing the empathy skill and created their item pools by determining a total of 142 items from those assessment tools. In their initial studies, Spreng et al. (2009) performed validity and reliability studies by applying 142 items to a group of 200 people. Following the structural validity study, the researchers determined 41 factors with an Eigen value higher than 1 and explaining 75.23 percent of the questionnaire's total variance. Estimating that empathy could be assessed in a single dimension as an emotional process, they restricted their exploratory factor analysis to one single factor. Therefore, they obtained a single factorial structure comprising 16 items, each of which had a factor load higher than .40. In the reliability study, the researchers reported the TEQ's Cronbach-α value as .85. Within the scope of the criterion-related validity studies, they found that the TEQ had a high positive correlation with a similar scale (Empathic Concern by Davis, 1983) and a negative correlation with a dissimilar scale (Autism Quotient by Baron-Cohen & Wheelwright, 2004). In their second study of a different sample, they detected that the questionnaire had similar correlations with the same scales. On the other hand, the third study of 65 university students concluded that the questionnaire's item total correlations varied between .34 and .71. During the same study, researchers discovered the questionnaire's test-retest reliability coefficient was .81. Having completed these validity and reliability studies, they underlined that the TEQ was a short, straight, homogenous and powerful assessment tool to evaluate empathy as an emotional process.

The Empathic Tendency Scale (ETS): ETS was developed by Dökmen (1988) for the purpose of assessing individuals' potential of empathizing in their daily lives. A Likert type scale, it contains 20 items and each question is scored from 1 point to 5 points. The minimum and maximum scores on the scale are 20 and 100, respectively. The total score implies the participants' empathic tendency scores. Higher scores mean higher empathic tendencies and vice versa. The test-retest reliability coefficient of the ETS was .82. The internal consistency reliability coefficient calculated by means

of Cronbach-α method was .72. The correlation between the subscale "understanding emotions" of the Edwards Personal Preference Schedule and the ETS was .68.

The Basic Empathy Scale (BES): BES was developed by Jolliffe & Farrington (2006). The scale's Turkish adaptation and validity and reliability studies were carried out by Topçu, Baker, & Aydın (2009). BES comprises 20 items. There is a five-item Likert-type key for the scale. The TEQ can assess empathy in two sub-dimensions—cognitive and affective. Researchers reported the Cronbach-α reliability coefficient as .83 for the entire scale, .80 for the cognitive sub-dimension and .76 for the affective subscale. The validity of the BES was analyzed by means of a CFA. It concluded that the two-factor structure of the original form was confirmed in the Turkish sample, too.

Procedure

In order to adapt the TEQ, researchers contacted R. Nathan Spreng, one of the developers of the questionnaire, to obtain the necessary permission. Then, the questionnaire was translated into Turkish by four instructors with a good command of English from the field of psychological counseling and guidance. After it had been translated by four different people independently, the translation forms were analyzed by the researchers. The statements that were believed to represent each item best were picked and a single form was created. This form took its final shape following the necessary corrections and discussions. After this stage, high-level correlations were identified between the items of the original form and the translated form. Afterwards, the questionnaire was given to the participants. The data collection tools were applied to the volunteer students during course hours. Applications took approximately 10-15 minutes.

Data Analysis

Prior to being subjected to statistical processes, research data underwent data cleaning (Tabachnick & Fidell, 2007; Osborne & Overbay, 2008). Wrong encodings detected by frequency tables were arranged by looking at raw data. It was determined that the missing values at all parameters were not above 5 percent. The structural validity study employed the exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) methods. EFA identifies the factors under which there are scale items mathematically. CFA is a statistical method aimed at theoretically determining which item is located under which factor before the analysis and allowing for analyzing item-factor association (Child, 2006; Brown, 2006). For EFA and CFA used during the structural validity in study. Schwab (2005a) stated that data collected from a total of 100 participants would be adequate in the principal components analysis as EFA. However, Costello & Osborne (2005) said that, in factor analysis studies, the number of participants is usually determined by the participant item rate, which is generally 10:1 but may decrease to 2:1. If the participant item rate is accepted as 10:1, there should be at least 160 participants for 16 items of the TEQ (16:10= 160). Given the number of the research participants ($n=$ 588), the number is much higher than it is supposed to be (588:16=37). It was determined that, in the research data, the univariate normal distribution ($z= \pm 3.00$)

and multivariate normal distribution (Mahalonobis D^2) with a normal distribution (Kolmogorov Smirnov $p \geq .05$) and linearity were not outliers (Schwab, 2005b; Tabachnick & Fidell, 2007). Since these criteria were sufficient, the research data used in the adaptation study was adequate for statistical analyses and the validity and reliability studies were performed. Researchers used internal consistency and test-retest methods in the reliability study of the TEQ. In addition, an item analysis was used to determine the questionnaire items' power of representing the questionnaire; researchers also conducted upper and lower 27 percent-group comparisons for the purpose of each item's power of distinctiveness (Büyüköztürk, 2007). Finally, researchers analyzed the TEQ's distinctive validity study on the base of gender. The statistical analyses were carried out by means of IBM PAWS SPSS 18 (SPSS, 2009) and LISREL 8.80 (Jöreskog & Sörbom, 2006) programs.

Results

Linguistic Equivalency

The most significant procedure during the adaptation of an assessment tool from the society in which it was developed to another society is the translation from the source language to the target language (Geisinger, 1994). According to the bilingual pattern used in the linguistic equivalency study (Deniz, 2007), the participants of a linguistic equivalency study should have an excellent command of both languages. Therefore, 33 students at Ege University's Department of English Language and Literature (all of which had an excellent command of both Turkish and English) participated to the linguistic equivalence study conducted in the first stage of the research. When the findings obtained as a result of the analysis were examined, a positive significant correlation ($r= .72$, $p= 000$) was found in the total of the source and target language forms. Furthermore, when the correlations between the items in the source and target languages were examined by the Spearman *rho* formula because of ordinal data (Gravetter & Wallnau, 2007), significant correlations with values varying between .41 and .72 were detected.

Table 1

Correlation Values Between Items in English and Turkish Forms

Items	rho
Item1	.72**
Item 2	.55**
Item 3	.48**
Item 4	.47**
Item 5	.42*
Item 6	.41*
Item 7	.55**
Item 8	.74**
Item 9	.45*
Item10	.61**
Item 11	.43**
Item 12	.68**
Item 13	.48**
Item 14	.54**
Item 15	.59**
Item 16	.72**

$*p \leq .05; **p \leq .001$

The items in the Turkish form of the questionnaire reflected the original form because researchers observed a sufficient correlation between the TEQ's original form and translated form when the correlation coefficients acquired as a result of the linguistic validity study were analyzed.

Item Analysis

An item analysis was performed in order to determine the questionnaire items' power of predicting the total score. According to the results, the values of Items 1, 6 and 9 were below.30. Therefore, those items were omitted from the questionnaire and the analyses were repeated. As a result of the item analysis conducted after omitting the said items, the item total correlations varied between .31 and .55. Upper and lower 27 percent group comparisons were conducted to determine each item's

power of distinctiveness. Accordingly, the differences between the items were statistically significant ($p \leq .001$). These results demonstrated that the questionnaire items' power to represent the questionnaire and distinctiveness were sufficient. The results obtained are given below in detail.

Table 2

Independent Groups t-test Results of Upper and Lower 27% Group Differences and Item Total Correlations of TEQ

Items	Corrected Item-Total Correlations[1]	t values for each items (Upper and lower 27% group)[2]
Item 2	.31	11.162*
Item 3	.37	9.449*
Item 4	.40	13.030*
Item 5	.44	11.835*
Item 7	.44	13.344*
Item 8	.36	10.202*
Item10	.36	9.344*
Item11	.47	13.073*
Item12	.52	15.550*
Item13	.51	16.597*
Item14	.55	13.616*
Item15	.35	12.178*
Item16	.41	12.247*

[1]n= 588, [2]n_1, n_2= 159, sd= 317, *$p \leq .001$

Structural validity

The structural validity of the TEQ was examined using EFA and CFA methods. As a result of the analysis performed by restricting it to a single factor, it was determined that the Kaiser Meier Olkin (KMO) coefficient was .85 and the Barlett χ^2 value was 1519.05 (p= .000). The variance value of the single factor with an Eigen value of 3.933 was 24.58 percent. However, the item factor loadings of Item 1 (.22), Item 6 (.26), and Item 9 (.29) were below .40. Although the factor loadings of these items were determined to be very low, following the first CFA for original model, it was found that all parameter estimation values of the items were positively loaded and the goodness of fit indexes were partially adequate (χ^2= 405.92, df= 104, χ^2/df= 3.91, GFI= .92, NFI= .88, RFI= .86, CFI= .91, RMR= .057, RMSEA= .070). When the corrected item total correlations were examined so the observed confirmation level was assessed to be sufficient, the values of Item 1 (.19), 6 (.20), and 9 (.23) were found to be very low. Therefore, Item 1, 6 and 9 were omitted from the questionnaire. Turkish validity and reliability studies of the TEQ which originally comprised 16 items were carried out on the base of 13 items. As a result of the EFA performed on the remaining items, the KMO value was .85 and the Barlett Sphericity Test χ^2 value was 1350.23 (p= .000); it explained 29.17 percent of the variance in total. A KMO coefficient of .70 and higher is adequate for accepting the Barlett χ^2 analysis as

important (Meyers, Gamst, & Guarino, 2006). As a result, it was determined that the analysis was sufficient enough.

Table 3

The Exploratory Factor Analysis Result of the TEQ

Items	h^2	F^1
Item 2	.16	.40
Item 3	.24	.49
Item 4	.26	.51
Item 5	.31	.56
Item 7	.31	.56
Item 8	.24	.49
Item10	.22	.47
Item11	.32	.57
Item12	.40	.63
Item13	.40	.63
Item14	.44	.67
Item15	.21	.45
Item16	.29	.53

F^1= TEQ total

As a result of the EFA performed during the TEQ's structural validity study, it was found that the item factor loads took a value between .40 and .67. Field (2005) expresses that researchers generally expect factor loads to be more than .30 as a result of the factor analysis. However, Hair et al. (2006) state that it should be above .40. As a result of the EFA, the factor loadings of 13 items were sufficient. EFA values were sufficient for 13 items included in the questionnaire's Turkish form and researchers analyzed the verification level of the model using the CFA. Given the first model output, the association of the error covariance belonging to Item 8 and Item 13 was effective in decreasing the chi-square value of the model. Therefore, Item 8 and Item 13 were analyzed; researchers found that they could be accepted as close to each other in terms of meaning. Therefore, the error covariances of these two items were associated.

Table 4

The Goodness of Fit Indexes

Models	χ^2	df	χ^2/df	GFI	NFI	RFI	CFI	IFI	RMR	RMSEA
First Model	265.34	65	4,09	.93	.91	.89	.93	.93	.054	.072
Final Model	234.67	64	3.67	.94	.91	.90	.94	.94	.052	.067

Due to the fact that a noticeable decrease occurred in the chi-square level as a result of the association of the item error covariance, the model following the association was accepted as the final model. The ratio of the chi-square value to the degree of freedom is below 5 in the final model. Moreover, the values belonging to GFI, NFI, RFI, CFI and IFI from the model goodness of fit indexes are more than .90. On the other hand, RMR and RMSEA values are loaded with the values below .08. Researchers (Aron & Aron, 2002; Schumacker & Lomax, 2004; Kline, 2005; Raykov & Marcoulides, 2006; Vieira, 2011) state that, in the structural equivalence model, goodness of fit loaded with .90 or more is a sign of a good fit. Hoe (2008) expresses that RMSEA value below .08 is acceptable as well. The goodness of fit indexes were sufficient, and the diagram belonging to the CFA final model is given below.
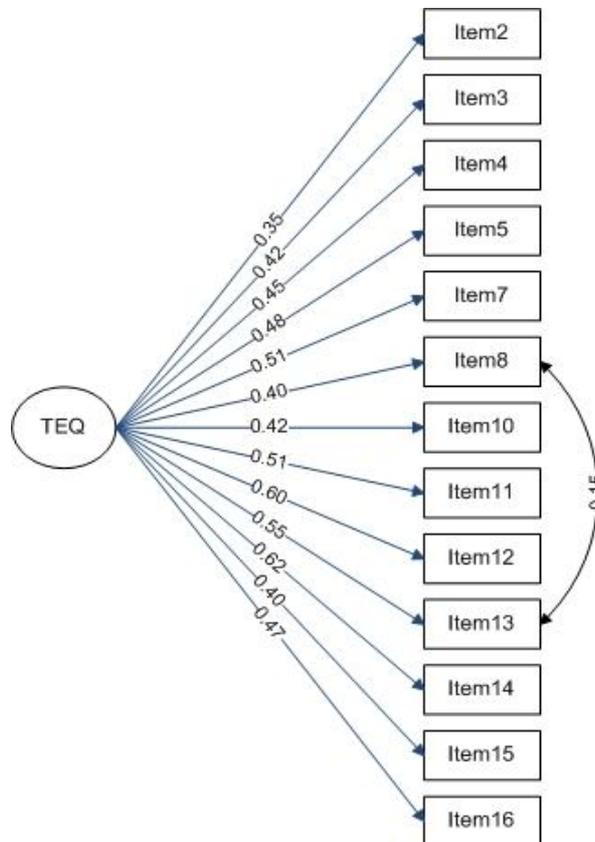

Figure 1. The model output of the TEQ as a result of CFA after item omission.

In the CFA, all the parameter estimations of the final model were positively loaded. Parameter estimations took values between .35 and .62. For the purpose of

analyzing the validity of the said results, the research data was distributed into two groups; one group was 40 percent (*n*= 235) and the other group was 60 percent (*n*= 353). Then a cross validation was performed. As a result of the cross validation, it was found that the parameter estimations belonging to the model in the CFA analyses of both groups were non-zero and positively loaded. Furthermore, it was determined that the goodness of fit of the 60 percent group [χ^2= 229.39, df= 65, χ^2/df= 3.53, GFI= .91, NFI= .87, RFI= .91, CFI= .90, IFI= .90, RMR= .065, RMSEA= .073] and the goodness of fit of the 40 percent group [χ^2= 113.50, df= 65, χ^2/df= 1.75, GFI= .93, NFI= .90, RFI= .88, CFI= .95, IFI= .95, RMR= .056, RMSEA= .056] were at acceptable levels except for first group's NFI and second group's RFI. Therefore, the structural validity of the TEQ's Turkish form was sufficient.

Criterion Related Validity

In order to demonstrate the criterion-related validity of TEQ, BES (Topçu, Baker, & Aydın, 2009), ETS (Dökmen, 1988) and the TEQ were applied to 115 university students. Accordingly, it was determined that the TEQ had a positively significant (*p*≤ .001) correlation of .47 with the cognitive dimension, .59 with the affective dimension, .68 with the entire BES and a positively significant (*p*≤ .001) correlation of .35 with ETS.

Reliability

The reliability of the TEQ was calculated using the test-retest and Cronbach's alpha internal consistency method. For the test-retest study, the questionnaire was applied once every three weeks to 77 students studying at Sakarya University's Faculty of Education, and a correlation of .73 was found between two applications. The Cronbach-α internal consistency coefficient of the questionnaire was .79. According to these results, the TEQ had a sufficient level of reliability.

The Distinctive Validity

Spreng et al. (2009) state that the TEQ showed considerable gender-based differences in their second and third studies. In their first study of 200 university students, the researchers found that the difference between female and male participants was not significant, and that the empathy levels of the female participants (second study $\bar{X}$= 48.93, sd= 6.77; third study $\bar{X}$= 48.93, sd= 6.90) were considerably higher than the empathy levels of the male participants (second study $\bar{X}$= 43.46, sd= 7.79; third study ($\bar{X}$= 43.63, sd= 7.93) in their second study of 79 university students (t_{77}= 3.16, *p*≤ .05, Cohen *d*= .73) and their third study of 65 university students (t_{63}= 2.39, *p*≤ .001, Cohen *d*= .63). The existence of gender-related differences in the data collected within the scope of the research was analyzed by means of *t*-test analysis for independent samples. Table 5 shows the results.

Table 5

Result for the Independent Sample t-test of the Gender-Related Differences in the TEQ

Gender	n	$\bar{x}$	ss	t	df	p	Cohen d
Women	357	40.77	6.46	8.465	586	.000	.57
Men	231	35.97	7.10				

The *t*-test analysis for independent samples determined that, of the research participants, the empathy levels of women ($\bar{x}$= 40.77, s= 6.46) were significantly higher (t_{586}= 8.465, *p*= .000) than the empathy levels of men ($\bar{x}$= 35.97, sd= 7.10). The influence magnitude of this difference is high like the second and third studies by Spreng et al. (2009) because gender is an important distinctive element in empathy. In other words, this research concluded that being female was more significant in high levels of empathy than being male.

Discussion and Conclusion

This study analyzed the psychometric properties of the TEQ in a sample of Turkish university students. In this context, its linguistic validity was examined and a linguistic equivalence study between the original form and the Turkish form was carried out. To demonstrate the structural validity of the TEQ, EFA and CFA were carried out. Within the scope of criterion-related validity, however, correlations between TEQ, ETS and BES were analyzed. On the other hand, reliability of the TEQ was calculated by means of the test-retest and internal consistency methods. Furthermore, an item analysis was performed to demonstrate the questionnaire items' power to represent the total score and distinctiveness. Finally, researchers carried out a gender-based distinctive validity study.

The linguistic equivalency study is of high importance in scale adaptation studies. In the research, after the original form was translated into Turkish and the most suitable statements had been determined, both the original form and the Turkish form were given to a group with a good command of both languages at different times. Then the correlations were analyzed both on the base of each item and on the total score obtained from the data of both applications. Accordingly, all the correlations between the items were positively significant and varied between .41 and .72. In terms of the total score, a relation of was obtained between the original form and the Turkish form. These results are sufficient in terms of linguistic equivalency.

The TEQ items' power predicting the total score and distinctiveness were examined using an item analysis and upper and lower 27 percent group methods. According to the results, Items 1, 6 and 9 were omitted from the questionnaire since their values were below .30 and analyses were performed. As a result of the item

analysis conducted after omitting these items, it was found that the corrected item total correlations varied between .31 and .55, i.e., the item total correlations of all items were above .30. When the groups of 27 percent were analyzed, all the differences between the items were significant. In line with these differences, it was concluded that the questionnaire items' had a distinctive power.

The factorial structure of the original questionnaire comprises a single dimension. In this study, researchers performed an EFA to demonstrate the factorial structure of the TEQ, and a CFA was performed to determine whether the factorial structure of the original form was confirmed in the sample of Turkish university students. As a result of the exploratory factor analysis performed by restricting it to a single factor, a structure having an Eigen value of 3.933 and explaining 29.17 percent of the total variance was obtained. It determined that the factor loads of the questionnaire items varied between .40 and .67. On the other hand, as a result of the CFA, it was concluded that the fit index values were sufficient and the Turkish sample confirmed the factorial structure of the original form.

Researchers analyzed the validity of the TEQ by means of the criterion-related validity as well as the EFA and CFA. Accordingly, this study used the ETS previously developed in the Turkish culture and the BES adapted into Turkish. This research found a correlation of .35 between TEQ and ETS, .68 between the TEQ and BES, .47 between the TEQ and the BES's cognitive subscale and .59 between the TEQ and the BES's affective subscale. The reliability of the scale was analyzed with test-retest and internal consistency (Cronbach-α) methods. Accordingly, the internal consistency reliability coefficient of the TEQ was .79. However, the reliability coefficient applied once in two weeks and calculated by means of the test-retest methods was .73.

Consequently, the results of this study turned the TEQ into a 13-item uni-dimensional assessment tool. These results showed that the TEQ had sufficient validity and reliability in the assessment of Turkish university students' empathy levels. It is thought that the questionnaire can be employed in the psychological consultation process as well as research on human relations and empathy owing to its features such as practical use and assessment and not being time-consuming.

References

Aron, A. & Aron, E. N. (2002). *Statistiscs for behavioral social sciences. A brief course*. New Jersey: Prentice Hall.

Baron-Cohen, S. & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism and normal sex differences. *Journal of Autism and Developmental Disorders, 34*(2), 163-175.

Bayam, G., Şimşek, E. U., & Dilbaz, N. (1995). Üç farklı meslek grubunda empatik beceri düzeylerinin karşılaştırılması [Comparision emphatic ability level between three differents profession groups]. *Kriz Dergisi, 3*(1-2), 205-207.

Bora, E. & Baysan, L. (2009). Empati Ölçeği-Türkçe Formunun üniversite öğrencilerinde psikometrik özellikleri [Psychometric features of Turkish version of empathy quotient in university students]. *Klinik Psikofarmakoloji Bülteni, 19*(1), 39-47.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Büyüköztürk, Ş. (2007*). Sosyal bilimler için veri analizi el kitabı. İstatistiki araştırma deseni SPSS uygulamaları ve yorum* [Data analysis handbook for social sciences. Statistical research design SPSS applications and interpretation]. Ankara: PegemA Yayıncılık.

Child, D. (2006). *The essentials of factor analysis*. London: Continuum International Pub.

Costello, A. B. & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis*. Practical assessment Reseach & Evaluation, 10*(7).

Dadds, M. R. et al. (2008). A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry & Human Devolopment, 39*, 111–122.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multi-dimensional approach. *Journal of Personality & Social Psychology, 44*(1), 113–126.

Deniz, K. Z. (2007). Psikolojik ölçme aracı uyarlama [The adaptation of psychological scales]. *Ankara Üniversitesi Eğitim Bilimleri Dergisi, 40*(1), 1-16.

Dökmen, Ü. (1987). Empati kurma becerisi ile sosyometrik statü arasındaki ilişki [The relation between empathy ability and sociometric statutes]. *Ankara Üniversitesi Eğitim Bilimleri Fakülte Dergisi, 20*(1-2), 183-207.

Dökmen, Ü. (1988). Empatinin yeni bir modele dayanılarak ölçülmesi ve psikodrama ile geliştirilmesi [Empathy measures based on new methods and its development via psychodrama]. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 21*(1-2), 155-190.

Dökmen, Ü. (2000). *İletişim çatışmaları ve empati* [Communication conflict and empathy]. İstanbul: Sistem Yayınları.

Elliott, R, Filipovich, H., Harrigan, L., Gaynor, J., Reimschuessel, C., & Zapadka, J. K. (1982). Measuring response empathy: The development of a multicomponent rating scale. *Journal of Counseling Psychology, 29*(4), 379-87.

Engeler, A. & Yargıç, L. İ. (2007). Kişiler arası tepkisellik indeksi: Empatinin çok boyutlu ölçümü [Interpersonal Reactivity Index: Measurement of Empathy Multidimensionally]. *New Symposium Journal*, *45*(3), 119-127.

Field, A. (2005). *Discovering statistics using SPSS.* Londra: Sage Publications.

Gallup, G. G. J. & Platek, S. M. (2002). Cognitive empathy presupposes self-awareness: Evidence from phylogeny, ontogeny, neuropsychology and mental illness. *Behavioral & Brain Sciences*, *25*(1), 36–37.

Gravetter, F.J. & Wallnau, L. B. (2007). *Statistics for the behavioral sciences.* (7th Edition). Belmont, CA: Thomson Higher Education.

Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, *6*(4), 304-312.

Hair, J. F., Black, B., Babin, B., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis.* Upper Saddle River: Prentice Hall.

Hoe, S. L. (2008). Issues and procedures in adopting structural equation modeling technique. *Journal of Applied Quantitative Methods*, *3*(1), 76-83.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting & Clinical Psychology, 33*, 307–316.

İkiz, F. E. (2006). *Danışma becerileri eğitiminin danışmanların empatik eğilim, empatik beceri ve tükenmişlik düzeyleri üzerindeki etkisi* [The effect of counseling skills training on counselors' empathic tendency, empathic skill and burnout levels]. Unpublished doctoral dissertation. Dokuz Eylül University, Institute of Education Science, İzmir.

Jolliffe, D. & Farrington D.P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, *29*, 589–611.

Jöreskog, K.G. ve Sörbom, D. (2006). *LISREL 8.80 for Windows [Computer Software].* Lincolnwood, IL: Scientific Software International, In

Kaya, A. & Siyez D. M. (2010). KA-Sİ Çocuk ve ergenler için empatik eğilim ölçeği: Geliştirilmesi geçerlik ve güvenirlik çalışması [Child and adolescent KA-Sİ empathic tendecy scale: Development, validity and reliability study]. *Eğitim ve Bilim*, *35*(156), 110-125.

Kline, R. B. (2005). *Principles and practice of Structural Equations Modeling.* New York: Guilford.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the empathy quotient. *Psychological Medicine, 34*, 911-924.

Marcia, J. (1990). Empathy and psychotherapy. In N. Eisenberg & J. Strayer (Eds.). *Empaty and its development*. Cambridge: Cambridge University Press.

Mehrabian, A. & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*, 525-543.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research. design and interpretation*. California: Sage Publication Inc.

Muncer, S. J. & Ling, J. (2006). Psychometric analysis of the empathy quotient (EQ) scale. *Personality and Individual Diff*erences, *40*, 1111-1119.

Osborne, J. W. & Overbay, A. (2008). Best practices in data cleaning. How outliers and "fringeliers" can increase error rates and decrease the quality and precision of your results. In J. W. Osborne (Ed.). *Best practices in quantitative methods*. California: Sage Publications.

Özbay, Y. & Şahin, M. (2000). Empatik Sınıf Atmosferi Tutum Ölçeği (ESATÖ): Geçerlik ve güvenirlik çalışması [Validity and reliability of Intraclassroom Empathic Attitude Scale (IEAS)]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 19*, 104-113.

Pecukonis, E. V. (1990). A cognitive/affective empathy training program as a function of ego development in aggressive adolescent females. *Adolescence, 25*(97), 59-76.

Raykov, T. & Marcoulides, G. A. (2006). *A first course in Structural Equation Modeling*. Mahwah: Lawrence Erlbaum

Rogers, C. R. (1983). Empatik olmak değeri anlaşılmamış bir varoluş şeklidir [Empathic: An unappreciated way of being]. (Trans. F. Akkoyun). *Ankara Üniversitesitesi Eğitim Bilimleri Fakültesi Dergisi, 16*(1), 103-124.

Schumacker, R. E. & Lomax, R. G. (2004). *A beginner's guide to Structural Equation Modeling*. New Jersey: Lawrence Erlbaum Ass.

Schwab, J. (2005a). *Principal component analysis* [Lecturer Notes]. Retrieved 03.03.2010 on http://www.utexas.edu/courses/schwab/sw388r7/SolvingProblems/SolvingHomeworkProblems.htm

Schwab, J. (2005b). *Principal component analysis: Validation, outliers, and reliability* [Lecturer Notes]. Retrieved 03.03.2010 on http://www.utexas.edu/courses/schwab/ sw388r7/SolvingProblems/SolvingHomeworkProblems.htm

Shamay-Tsoory, S. G., Tomer, R. D., Goldsher, B., Berger, B. D., & Aharon-Peretz, J. (2004). Impairment in cognitive and affective empathy in patients with brain lesions: Anatomical and cognitive correlates. *Journal of Clinical & Experimental Neuropsychology, 26*, 1113–1127.

SPSS (2009). *PAWS Statistics 18 core system user's guide*. Illinois: SPSS Inc.

Spreng, R., N., Kinnon, C. M., Mar, R. A., & Levine, B. (2009). The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment, 91*(1), 62-71.

Stephan, W. G. & Finlay, K. (1999). The role of empathy in improving intergroup relations. *Journal of Social Issues*, *55*, 729-743.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Allyn and Bacon.

Topçu, Ç., Baker, Ö. E. ve Aydın, Y. Ç. (2009). *Turkish adaptation of Basic Empathy Scale (BES)*. The First International Congreess of Educational Research. 1-3.May.2009, Çanakkale, Turkey. Retrieved 03.03.2010 on http://oc.eab.org.tr/egtconf/pdfkitap/ pdf/388.pdf

Vieira, A. L. (2011). *Interactive LISREL in practice. Getting started with a SIMPLIS approach*. New York: Springer.

Wakabayashi, A. et al. (2006). Development of shorts forms of the empathy quotient (EQ-short) and the systemizing quotient (SQ-short). *Personality & Individual Diff*erences, *41*(5), 929-940.

Wispé, L. (1990). History of the concept of empathy. In N. Eisenberg ve J. Strayer (Eds.). *Empathy and its development*. Cambridge: Cambridge University Press.

Yüksel, A. (2004). Empati eğitim programının ilköğretim öğrencilerinin empatik becerilerine etkisi [Empathy education program effects of empathic ability on elementary school student]. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, *17*(2), 341-354.

Toronto Empati Ölçeği: Türk Üniversite Öğrencilerinde Psikometrik Özelliklerinin Değerlendirilmesi
(Özet)

Problem Durumu

Günümüzde empatinin insan ilişkilerini kolaylaştırıcı bir etkiye sahip çok boyutlu bir yeti olduğu görüşü yaygın olarak kabul görmüştür. Empatinin birden çok bileşenden oluştuğu konusundaki yaygın görüş bu yetinin ölçümüne ilişkin çeşitliği de beraberinde getirmiş, pek çok araştırmacı farklı boyutları içeren empati ölçekleri geliştirmişlerdir Çok boyutlu empati ölçeklerinin çeşitliliği ve bu ölçeklerden alınan sonuçlar arasında belirgin farklıklar olmasından yola çıkarak Toronto Empati Ölçeği (TEÖ) özgün formunda tek boyutlu bir ölçme aracı olarak yapılandırılmıştır. Ölçeğin geliştiricileri, TEÖ'ni geliştirilme amaçlarını çok boyutlu ölçümlerden tek boyutlu ölçümlere bir geri dönüş sağlamak olmadığını daha çok ölçümler arasındaki farklılıkları minimum düzeye indirmek ve görüş birliği sağlanmış çekirdek bir ölçme aracı geliştirmek olarak tanımlamaktadırlar.

Araştırmanın Amacı

Bu çalışmada Toronto Empati Ölçeğinin (TEÖ) Türk üniversite öğrencilerinden oluşan örneklemde psikometrik özellikleri incelenmiş; geçerlik ve güvenirlik çalışmaları yapılmıştır. Bu bağlamda ilk olarak dil geçerliliğine bakılmış ve özgün form ile Türkçe form arasındaki dilsel eşdeğerlik çalışması yapılmıştır. TEÖ'nin yapı geçerliğini ortaya koymak üzere betimleyici ve doğrulayıcı faktör analizi yöntemleri kullanılmıştır. Ölçüt bağıntılı geçerlik kapsamında ise TEÖ ile Empatik Eğilim Ölçeği (EEÖ) ve Temel Empati Ölçeği arasındaki ilişkiler incelenmiştir. TEÖ'nin güvenirliği ise test tekrar test ve iç tutarlık yöntemleriyle hesaplanmıştır. Ayrıca ölçek maddelerinin toplam puanı temsil etme gücünü ve ayırt ediciliğini ortaya koyabilmek amacıyla madde analizi yapılmıştır. Son olarak cinsiyete dayalı ayırt edici geçerlik çalışması yürütülmüştür

Araştırmanın Yöntemi

Araştırmada üç farklı grup olmak üzere toplamda 698 üniversite öğrencisi katılımcı olarak yer almıştır. İlk grupta yer alan Ege Üniversitesi, İngilizce Dili ve Edebiyatı bölümünde eğitim alan 33 üniversite öğrencisi ölçeğin özgün ve hedef formuna incelemeye yönelik dil geçerliği çalışmasında yer almıştır. Araştırmanın geçerlik ve güvenirlik çalışmalarının büyük bir kısmının yürütüldüğü ikinci grubu oluşturan katılımcılar 2008-2009 eğitim-öğretim yılı içerisinde Sakarya Üniversitesi Eğitim Fakültesi'nde öğrenim gören olasılıksız örnekleme teknikleri arasında yer alan uygun örnekleme yöntemiyle belirlenen 357'si kadın (%60,7), 231'i erkek (%39,3) toplam 588 üniversite öğrencisidir. Araştırma verisi toplandığı sırada katılımcıların 269'u üniversite birinci sınıfta (%45,7, kadın $n= 196$, erkek $n= 73$), 100'ü ikinci sınıfta (%13,1, kadın $n= 60$, erkek $n= 40$), 77'si üçüncü sınıfta (kadın $n= 39$, erkek $n= 38$) ve 142'si ise son sınıfta (%24,1, kadın $n= 62$, erkek $n= 80$) öğrenimlerine devam etmektedir. Katılımcıların genel yaş ortalaması 20,60 olarak belirlenmiştir (ss= 20,60;

kadın $\bar{x}$= 20,22, ss= 1,85; erkek $\bar{x}$= 21,20, ss= 1,81). Son olarak Sakarya Üniversitesi Eğitim Fakültesine devam 77 öğrenci üzerinden elde edilen veriyle araştırmanın test tekrar test geçerliğine ait çalışmalar yürütülmüştür. Araştırma bulgularının analizleri sırasında betimsel faktör analizi, doğrulayıcı faktör analizi, bağımsız örneklemler için *t*-testi, Pearson Momentler Çarpımı Korelasyon katsayısı ve Spearman *rho* yöntemleri IBM PAWS SPSS 18 ve LISREL 8.80 programları aracılığıyla incelenmiştir.

Araştırmanın Bulguları

Ölçek uyarlama çalışmalarında dilsel eş değerlik çalışması büyük önem taşımaktadır. Araştırmada özgün formunun Türkçeye çevrilmesi ve en uygun ifadelerin belirlenmesinden sonra her iki dile de hâkim bir gruba özgün form ve Türkçe form farklı zamanlarda uygulanmıştır. Ardından her iki uygulamanın verisi üzerinden hem her madde bazında hem de alınan toplam puan bazında ilişkiler incelenmiştir. Buna göre maddeler arası ilişkilerin tümünün pozitif yönde önemli olduğu .41 ile .72 arasında değiştiği görülmüştür. Toplam puan açısından ise özgün form ile Türkçe form arasında .72 korelasyon elde edilmiştir. Bu sonuçlar dilsel eşdeğerlik açısından yeterli kabul edilebilecek düzeydedir.

TEÖ' nün maddelerinin toplam puanı yordama gücü ve ayırt ediciliği madde analizi ve %27'lik alt-üst gruplar yöntemleriyle incelenmiştir. Elde edilen sonuçlara göre 1, 6 ve 9. maddelerin değerlerinin .30'un altında olduğu saptandığından bu maddeler ölçekten çıkarılarak analizler yapılmıştır. Söz konusu maddeler çıkarıldıktan sonra yapılan madde analizi sonucu madde toplam korelasyonlarının .31 ile .55 arasında olduğu başka bir ifadeyle tüm maddelerin madde toplam korelasyonlarının .30'dan yukarıda olduğu bulunmuştur. %27'lik gruplar incelendiğinde maddeler arasındaki farklılıkların tümünün önemli olduğu görülmüştür. Bu farklılıklar doğrultusunda ölçek maddelerinin ayırt edicilik gücünün yüksek olduğu sonucuna varılmıştır.

Özgün ölçeğin faktör yapısı tek boyuttan oluşmaktadır. Bu araştırmada da TEÖ'nin faktör yapısını ortaya koyabilmek amacıyla betimleyici faktör analizi ve özgün formun faktör yapısının Türk üniversite öğrencilerinden oluşan örneklemde doğrulanıp doğrulanmayacağını ortaya koymak üzere doğrulayıcı faktör analizi yapılmıştır. Tek faktörle sınırlandırılarak yapılan betimleyici faktör analizi sonucunda özdeğeri 3,933 olan ve toplam varyansın % 24,58' ini açıklayan bir yapı elde edilerek ölçek maddelerinin faktör yüklerinin .40 ile .67 arasında değiştiği belirlenmiştir. Doğrulayıcı faktör analizi sonucunda ise uyum indeksi değerlerinin yeterli düzeyde olduğu sonucuna ulaşılmış ve özgün formun faktör yapısının Türk örneklemde doğrulandığı görülmüştür [χ^2= 234,67, df= 64, χ^2/df= 3,67, GFI= ,94, NFI= ,91, RFI= ,90, CFI= ,94, IFI= ,94, RMR= ,052, RMSEA= ,067].

TEÖ'nin geçerliği betimleyici ve doğrulayıcı faktör analizinden başka ölçüt bağıntılı geçerlik yöntemiyle de incelenmiştir. Buna göre daha önce Türk kültüründe geliştirilmiş Empatik Eğilim Ölçeği (EEÖ) ve Türkçeye uyarlanmış Temel Empati Ölçeği kullanılmıştır. TEÖ ile EEÖ arasında .35, Temel Empati Ölçeği ile .68, Temel Empati Ölçeği Bilişsel alt boyutuyla .47, Duyuşsal alt boyutuyla .59 düzeyinde pozitif yönde önemli ilişkiler bulunmuştur. Ölçeğin güvenirliği ise test tekrar test ve iç tutarlık (Cronbach alfa) yöntemleriyle incelenmiştir. Buna göre TEÖ'nin iç tutarlık

güvenirlik katsayısı .79 olarak bulunmuştur. İki hafta arayla gerçekleştirilen, test tekrar test yöntemiyle hesaplanan güvenirlik katsayısı ise .73 olarak bulunmuştur.

Araştırmada TEÖ'nden alınan puanların cinsiyete göre bir farklılık gösterip göstermediği de incelenmiş ve elde edilen bulgulara göre kadın ve erkeklerin empati düzeyleri arasında istatistiksel olarak anlamlı bir farklılık bulunmuştur. Buna göre kadınların empati düzeylerinin erkeklere göre anlamlı derecede daha yüksek olduğu sonucuna ulaşılmıştır.

Özgün formda yer alan 1. Madde "*Birisi heyecanlandığında bende heyecanlanırım*", 6. Madde "*Benden daha az şanslı insanlara karşı duyarlı ve ilgiliyimdir*" ve 9. Madde "*Diğer insanların ruh hallerine uyum sağlarım*" madde toplam korelasyonları ve faktör yükleri .30'dan düşük olduğundan dolayı çıkarılmış ve analizler bu maddeler olmaksızın tekrar yapılmıştır. Bu üç madde dışında TEÖ'nin uyarlanmasında kültürlerarası geçerlilikle ilgili bir problemle karşılaşılmamıştır.

Araştırmanın Sonuçları ve Öneriler

Sonuç olarak TEÖ bu çalışma sonucunda 13 maddelik tek boyutlu bir ölçme aracı haline gelmiştir. Elde edilen tüm bu sonuçlar TEÖ'nin Türk üniversite öğrencilerinin empati düzeylerini ölçmede yeterli geçerlik ve güvenirliğe sahip olduğunu ortaya koymuştur. Ölçek kullanımının ve değerlendirilmesinin kolay ve pratik oluşu, çok zaman almaması gibi özelliklerinden dolayı gerek psikolojik danışma sürecinde gerekse insan ilişkileri ve empati ile ilgili yapılacak araştırmalarda kullanılabileceği düşünülmektedir.

Anahtar Sözcükler: Toronto Empati Ölçeği, ölçek uyarlama, psikometrik özellikler