e-international journal of educational research
Volume: 4 Issue: 2 Spring 2013 pp. 29-47

Adaptation of Attitudes toward Educational Measurement Inventory (ATEMI) to Turkish

Res.Asis.Ceyhun Ozan
Atatürk University-Turkey
ozanceyhun08@gmail.com

Assoc.Prof.Dr.Erdoğan Köse
Atatürk University-Turkey
erdogan63@gmail.com

Abstract

The aim of the study is to adapt Attitude toward Educational Measurement Inventory (ATEMI) to Turkish and report the validity and reliability studies. Pre-service teachers (n= 490) enrolled in a measurement and evaluation course in the Atatürk University in Turkey completed ATEMI. Results of language equivalency showed that the correlation coefficients between Turkish and English forms ranged between .70 to .93. Results of exploratory factor analysis showed that the 31 items loaded on three factors. The total variance explained was 47.4% and factor loadings ranged between .31 to .83. Comfirmatory factor analyses indicated that a three factors structure of the ATEMI provided a good fit to the observed data. The internal consistency reliability coefficient of the scale was found as .92 and the test-retest reliability coefficient was found as .78. Based upon the results of the validity and reliability studies, it was determined that the adapted Turkish form of the inventory is a valid and reliable instrument that can be used to determine the attitudes of the pre-service teachers towards measurement and evaluation in education. As the inventory is applied on the pre-service teachers enrolled to "Measurement and Evaluation" course in teacher training programs in the beginning of the school term, it can be important for the instructors to determine the negative attitudes of the pre-service teachers towards the course and to take precautions and develop training strategies for this situation.

Keywords: Education, Measurement and evaluation, Pre-service teacher, Attitude

INTRODUCTION

Measurement and evaluation is the primary element of any education process and pursuit of quality in education. The main purpose of assessment and evaluation is to assess and to improve the quality of the education (Linn & Grönlund, 1995). Measurement and evaluation, which is an integral part of primary and secondary schools in open educational systems, is unquestionably one of the substantial qualities that the pre-service teachers should have in the teacher training institutions. According to Mertler (2003) and Plake (1993), the primary responsibility of the teachers is the evaluation of in which level the students' learning is. The assessment of classroom activities takes too much time for the teachers and has an important place in their lives. In this respect, teachers' positive attitudes towards measurement and evaluation to make reliable evaluations are one of the important factors (Alkharusi, Kazem, & Al-Musawi, 2011; Popham, 2006). However, some studies carried out on the pre-service teachers that have taken measurement and evaluation course show that pre-service teachers have negative attitudes towards it because of the reasons such as its not meeting their

professional needs, containing mathematical processes and being a really hard course (Alkharusi, 2009; Bryant & Barnes, 1997; Kottke, 2000; VanZile-Tamsen & Boes, 1997).

Teacher training system has been reconstructed with the assign of all the teacher training institutions to universities in 1982, in Turkey. Within this period, measurement and evaluation course has taken place as part of teacher training. With the restructuring in 1997, measurement and evaluation course has started to be taught with the name of 'Instructional Planning and Evaluation' by combining curriculum development and implementation course. However, this implementation was abandoned with the restructuring in 2006 and it has taken its place in curriculums with the name of 'Measurement and Evaluation' course up until today.

Today, restructuring and arrangements about teacher training syllabus have been discussed, and the importance and necessity of teaching profession courses have been emphasized. Erden (1995) stated that pre-service teachers are supposed to study more, attend eagerly and be successful in their courses if they are appealed with or have positive attitudes toward the teacher certificate courses. Similarly, Ekici (2008) stated that pre-service teachers with higher attitudes toward the teacher certificate courses will be more eager in learning and more successful, and they will love their profession when they start teaching. Besides, according to Koçak and Önen (2011), it is a waste of time to search solutions for the problems, encountered during teacher training process, without knowing the attitudes of pre-service teachers. Any arrangement, which may provide positive values about their profession for pre-service teachers who are going to start teaching soon, should be carried out throughout their education process. Also, Karaca (2006) stated that as education is an important tool to change attitudes, knowing what attitudes their students have toward their courses can be a significant factor to improve the quality of the education.

Several measurement instruments have been developed to measure the pre-service teachers' attitudes toward teaching profession courses in the teacher training programs in Turkey. Yüksel (2009) developed 'Educational Courses Inventory' to study the attitudes of the pre-service teachers toward teaching profession courses; attitude scales were developed to research the attitudes toward 'Instructional Planning and Assessment' course by Karaca (2006), 'Introduction to Education' course by Koçak and Önen (2011), 'Curriculum and Instruction' by Oğuz (2012), 'Lessons of School Experience' course by Kılınç and Salman (2007); however no attitude inventory hasn't been developed for 'Measurement and Evaluation' course. In this study, two of the three dimensions of the inventory whose linguistic equivalence, validity and reliability studies are wanted to be done aims to measure the attitudes of pre-service teachers toward 'assessment and evaluation' course in the teacher training program. In this sense, it is thought that it would be important to bring the inventory in Turkish culture. Another dimension of the inventory aims to measure the general attitude in education toward measurement and evaluation.

When analyzing the inventories developed or adapted in assessment and evaluation course in Turkey, several inventories have been developed to measure the pre-service teachers' perceived levels of efficacy of measurement and evaluation

Yaman and Karamustafaoğlu (2011), assessment preferences Gülbahar and Büyüköztürk, (2008), measurement and evaluation techniques (Balcı and Tekkaya, 2000); however, there has not been any inventories aimed to measure the pre-service teachers' attitudes toward measurement and evaluation. It is thought that creating the Turkish form of the inventory would contribute to the studies which intend to improve academic success of the students. In this sense, the aim of the study is to adapt 'Attitude toward Educational Measurement Inventory', developed by Bryant and Barned in 1997, to Turkish and does the validity and reliability studies.

METHOD

Sample: In the study, three different study groups' views in terms of linguistic equivalence, validity and reliability studies of the inventory were sought. These three study groups are formed from pre-service students who registered in the Kazım Karabekir Education Faculty of Atatürk University in 2011-2012 spring term. All participants were requested to fill the voluntary participation form, and they were informed about the nature and purpose of the research project. They were made aware that they could withdraw from the project at any time. The researchers have paid attention to form the groups with pre-service teachers who take measurement and evaluation course because the inventory includes items related to this course. To do the linguistic equivalence study of the inventory, the first study group is formed of 80 students getting formal education in the department of foreign language education and also registed with the course of measurement and evaluation. In the second study group which is formed to determine the construct validity of the inventory, there are 360 pre-service teachers from several different departments. The demographic information about the pre-service teachers in the second study group is indicated in Table 1. And the third study group, in which the reliability of the inventory is determined with test-retest method, is formed of 50 pre-service teachers who get formal education in elementary social sciences department.

Table 1: The Data Related to the Second Study Group

	f	%
Sex		
Female	224	62.2
Male	133	36.9
Not reported	3	0.8
Department		
Mathematics	134	37.2
Social Sciences	74	20.6
German Language	30	8.3
French Language	27	7.5
Pre-school	42	11.7
Visual Arts	34	9.4
Music	19	5.3
Total	360	100

Instrument: 'Attitude toward Educational Measurement Inventory' (ATEMI) developed by Byrant and Barnes (1997) to measure the attitudes of pre-service teachers toward measurement and evaluation in education is used after getting permissions from the developers of the inventory by contacting with them through e-mail. The original inventory is formed of three factors and 31 items; 13 items of relevance, 10 items of affective and 8 items of course. 17 items of the inventory includes negative expressions. The inventory is a five-level likert item and is rank-ordered as 'strongly disagree' (1), 'disagree' (2), 'undecided' (3), 'agree' (4), and 'strongly agree' (5). After inverting the negative items and grading them, the higher the obtained points mean the more positive attitudes of pre-service students have toward assessment and evaluation course. Internal consistency coefficients of Cronbach's alpha related to the sub-dimensions of the inventory and the total points change between .88 and .93, and the correlation coefficients related to the test-retest method applied after three weeks change .60 and .74. The original and final translated items are presented in the Appendix 1.

Procedures: In the first stage, linguistic equivalence studies were done. First of all, the English form of the inventory was translated into Turkish, and then it was gone through by three assistant professors who expert in education and the fluent in two languages, and the corrections considered necessary were done. Afterwards, the Turkish form was analyzed for meaning and grammar structure and necessary corrections were done and Turkish test form was obtained. The Turkish and English forms of the inventory were applied on 80 third year students from the department of English language education every with three weeks interals and the scores between the two inventories were calculated with Pearson Product-Moment Correlation coefficient. The construct validity of the inventory was analyzed with exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Exploratory factor analysis is used to explore the factor construct predicated on the relations between variables. In confirmatory factor analysis which analyzes model-data fit, hypothesis of the relations between variables is tested (Kline, 2000; Tabachnick & Fidell, 2007). Chi-Square Goodness, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Residual (SRMR), Normed Fit Index (NFI), Non-Normed Fit Index (NNFI), Comparative Fit Index (CFI), Goodness of Fit Index (GFI) and Adjusted Goodness of Fit Index (AGFI) were used to evaluate the validity of the model created in CFA. Item total correlations of each item in the inventory were calculated to determine how sufficient they are in distinguishing the pre-service students in terms of their attitudes toward measurement and evaluation in education. According to total points, independent samples t-test was used for the significance between item points of lower-group with 27% and upper-group with 27%. The correlation coefficient obtained from test-retest method and Cronbach's Alpha internal consistency coefficient was calculated to determine the reliability of the inventory.

FINDINGS

Linguistic equivalence

In the linguistic study of ATEMI, the original form of the inventory (in English) and Turkish form of the inventory have been applied on 80 pre-service teacher of English language teaching and the correlation between the points obtained from these two inventories has been taken as linguistic equivalence coefficient of the inventory. In this application, the English forms have been given first to the pre-service teachers and then the Turkish forms. The findings related to the correlation coefficients between the points obtained from the Turkish and English forms of the inventory are indicated in Table 2.

Table 2: Data about Linguistic Equivalence of ATEMI According to Items

Items	r	Items	r	Items	r
i1	.70	i12	.71	i22	.72
i2	.84	i13	.83	i23	.71
i3	.83	i14	.70	i24	.71
i4	.83	i15	.71	i25	.73
i5	.93	i16	.73	i26	.73
i6	.84	i17	.71	i27	.70
i7	.71	i18	.73	i28	.73
i8	.82	i19	.73	i29	.74
i9	.71	i20	.71	i30	.72
i10	.72	i21	.71	i31	.76
i11	.82	i22	.72		

When analyzing Table 2, it is seen that the correlation coefficients between the points obtained from the original and Turkish forms of the inventory are between .70 and .93. These results indicate that the consistency between the English and Turkish forms of the inventory is high. Thus, it has been acknowledged that the linguistic equivalence of the English and Turkish forms is provided for each item in the inventory.

Construct Validity

***Exploratory Factor Analysis (EFA)*:** In the study, exploratory factor analysis (AFA) was used to reveal the structure of the original form of ATEMI on Turkish students. EFA aims to reach from a great number of variables (items) to definable limited number of significant constructs that those variables can explain together. The key criterion in evaluation the results of factor analysis is factor loading that can be interpreted as correlations between variables and factors. Higher factor loadings are considered to be an indicator that the variable can be under the mentioned factor (Büyüköztürk, 2010). Item's factor loading value is usually preferred to be 0.45 or higher. However, it is seen that factor loading value is approved to be until 0.30 for a limited number of items in applications (Büyüköztürk, 2010; Hair, Black, Babin,

Anderson, 2010; Kline, 2000; Tabachnick & Fidell, 2007). In this study, .30 is accepted as a criterion to find factor loading values adequate. In EFA firstly the correlation matrix of all the items was analyzed and then it was looked through if there was a considerable amount of significant correlations, and it was seen that there were significant relations appropriate to make factor analysis. Then, Kaiser-Meyer-Olkin (KMO) and Bartlett`s test of sphericity were done for sampling adequacy. According to Kaiser (1974) and Pallant (2005) KMO value should be greater than .60 and the Bartlett`s test of sphericity should be significant. It can be said that as the KMO values is closer to 1, the relations between variables will be distinct and the factor analysis will give reliable results (Field, 2009). In this study, it was found that KMO Sampling Adequacy coefficient is .91, $\chi 2$ value of Bartlett`s test of sphericity is 5113,267 ($p<.001$). These values show that the data is appropriate for factor analysis.

Table 3: Factors of Scale and Factor Loadings of Items

Item Number	Relevance	Affective	Course
3	.826		
2	.726		
5	.726		
4	.652		
1	.623		
6	.584		
9	.465		
12	.441		
8	.440		
13	.434		
7	.397		
11	.363		
10	.321		
16		.827	
17		.788	
19		.630	
18		.518	
14		.468	
20		.459	
15		.418	
21		.418	
23		.327	
22		.312	
24			.834
25			.822
27			.641
30			.547
26			.545
31			.442
29			.418
28			.372
Explained Variance Total= % 47.4	% 32.069	% 9.430	% 5.901

The results of principal component analysis technique and oblique rotation factor analysis are limited to 3 factors in EFA because the original form of Attitudes toward

Educational Measurement Inventory has three factors. If there is orthogonality between the factors of an inventory, varimax rotation technique is used; and if there is a constant relation sequence, oblique rotation is generally used (Gorsuch, 1974; Tabachnick & Fidell, 2007). Oblique rotation technique has been used in this study because the factors of the inventory are related. According to results of the analysis, a 3-factor construct explaining 47.4 percent of the total variance and taking part in the sub-dimensions of the original form altogether with the items has been emerged. Besides, it has not been necessary to remove any of the items from the inventory because all the factor loadings of the items are higher than .30. The data about the construct of the Turkish form is indicated in Table 3.

As it is seen in Table 3, the Turkish form of ATEMI consists of 3 dimensions as in the original form. The first dimension has been called as 'Interest' and consisted of 13 items. Interest dimension explains 32.07% of total variance and its factor loadings vary between .31 and .83. Interest dimension is intended for the importance and practicality of assessment and evaluation of students and therefore it reflects the pre-service teachers' interests for assessment and evaluation. Items about the emotions of the pre-service teachers, particularly mathematical aspect of them, about assessment and evaluation has been taken part in the second dimension called as 'Affective'. This dimension that is consisted of 10 items explains 9.43% of the total variance and the factor loading of the items vary between .31 and .83. The third and last dimension called 'Course' consists of 8 items and explains 5.90% of the total variance. The items in this dimension are related to assessment and evaluation course found in teacher training programs and the factor loadings vary between .37 and .83. These items reflect the pre-service teachers' pleasure form the assessment and evaluation course and their desires to learn much more about this course. The correlations between the dimensions have also been analyzed besides the factor loadings of the inventory and the variance proportions explained by them. Correlation coefficients about the dimensions are indicated in Table 4.

Table 4: Correlations between Factors

Factors	Relevance	Affective	Course
Relevance			
Affective	.458*		
Course	.700*	.517*	
Total	.875*	.779*	.863*

*p<0.01

As it is seen in Table 4, total points and the correlations between the points of sub-factors vary between 0.78 and 0.88, and these correlation coefficients have been found to be significant at the level of 0.01. The correlations' being high and significance shows that these three sub-factors are the sub-factors of attitudes toward educational measurement.

Confirmatory Factor Analysis **(CFA):** CFA is a method to determine the validity which has been particularly used in the adaptation of the inventories developed in other

cultures and samples. According to Sümer (2000), CFA is an analysis to evaluate how the factors, created of many different variables based on a hypothetic basis, corresponds to the real data. In other words, CFA aims to analyze to what extent the construct that has been determined or conceived before is confirmed with the obtained data. CFA is based on examining a prediction in which specific variables mainly take part in pre-determined factors on a hypothesis, while the factor construct of the data is determined on the basis of factor loadings (weights) without specific expectations or hypothesis testing in exploratory factor analysis.

Many fit indexes are used to determine the conformance compliance of the model tested in CFA. It is suggested that numerous fit indexes should be used to present the conformance of the model as they have strong and weak aspects compared to each other while evaluating the conformance between hypothetical model and real data (Büyüköztürk, Akgün, Özkahveci & Demirel, 2004). The commonly used indexes being talked about are Chi-Square Goodness, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Normed Fit Index (NFI), Non-normed Fit Index (NNFI), Comparative Fit Index (CFI), Goodness of Fit Index (GFI) and Adjusted Goodness of Fit Index (AGFI). The calculated chi-square/df value between 2 and 3, RMSEA value between 0.05 and 0.08, SRMR value between 0.05 and 0.10, NFI and GFI values between 0.90 and 0.95, NNFI and CFI values between 0.95 and 0.97, and AGFI value between 0.85 and 0.90 mean that the model is in an acceptable conformance while the calculated chi-square/df value between 0 and 2, RMSEA value between 0.01 and 1, SRMR value between 0 and 0.05, NFI and GFI values between 0.95 and 1, NNFI and CFI values between 0.97 and 1, and AGFI value between 0.95 and 1 mean that the model shows an elegant conformance (Baumgartner & Hombur, 1996; Byrne & Campbell, 1999; Gefen & Straub, 2000; Hooper, Coughlan & Mullen, 2008; Hu & Bentler, 1999). The fit indexes of the model in the CFA have been analyzed and it is seen that the statistics are not in intended levels (Chi-square/df=4.28; RMSEA=0.095; SRMR=0.097; NFI=0.91; NNFI=0.93; CFI=0.93; GFI=0.87 and AGFI= 0.84). As a result of the analysis of modification indices, correlation levels between the errors of some items have been considered and revision has been made accordingly. The result of adding (releasing) a parameter which has stable modification indices or adding a new parameter shows the decrease in the chi-square value (Sümer, 2000). For this purpose, the correlations between I3 and I2, I25 and I24, I17 and I16, I19 and I18, I6 and I5, I5 and I4, and I13 and I5 have been released. Items released correlations are at the same dimensions. So, releasing correlations between items don`t have an impact on construct validity of scale adversely. The values about the new goodness of fit index obtained as a result of the modifications are indicated in Table 5.

Table 5: The Values about the Goodness of Fit Index

Chi-square	df	p	Chi-quare/df	RMSEA	SRMR	NFI	NNFI	CFI	GFI	AGFI
1171.46	423	p<.05	2.78	0.070	0.076	0.94	0.96	0.96	0.90	0.87

When analyzing Table 5, it is seen that chi-square value is significant and the proportion of chi-square value to degree of freedom is between an acceptable range of 2 and 3. When analyzing the values related to other goodness of fit indexes, it can be said that they are all in acceptable ranges, and according to these results model-data fit is provided and the construct validity of the inventory is confirmed. The diagram about the applied confirmatory factor analysis is indicated in Figure 1.

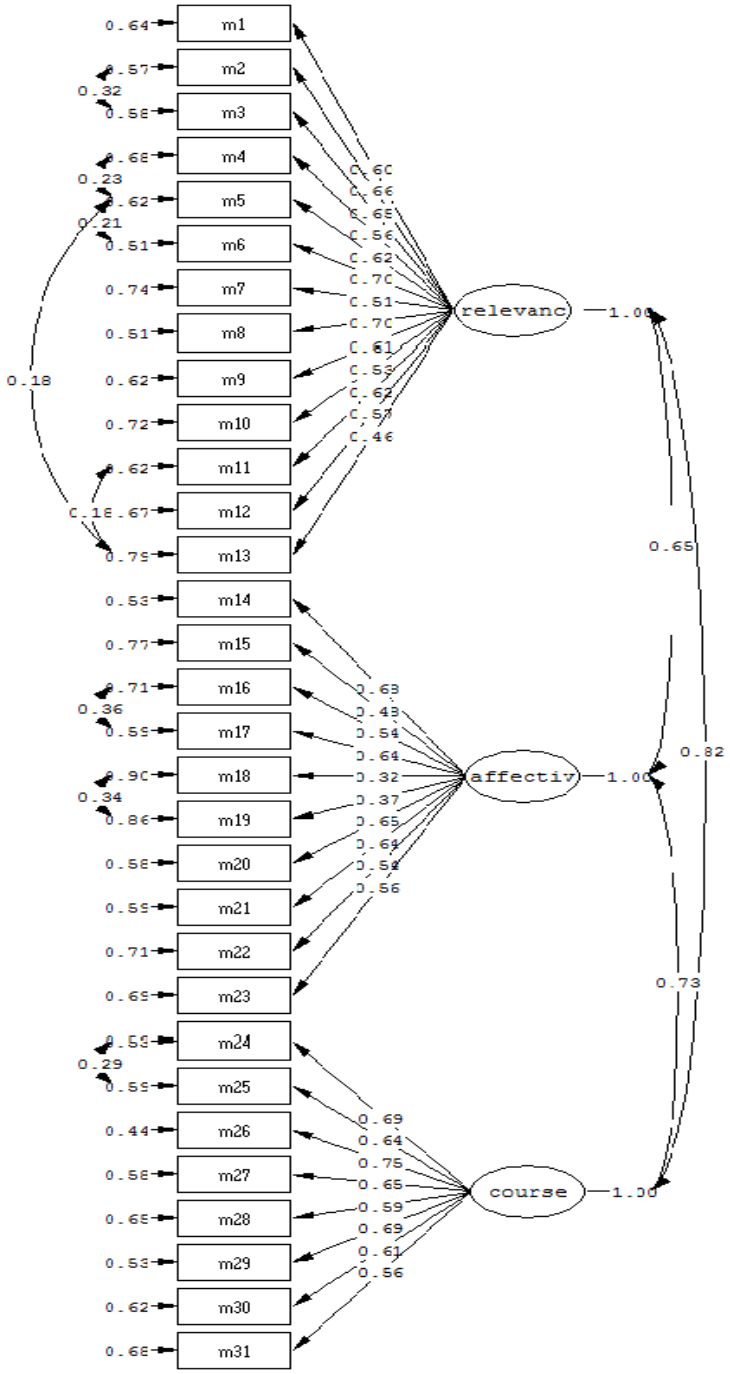

Figure 1: Diagram about the Confirmatory Factor Analysis

The correlation coefficients and the effect of each item on implicit dependent variables are seen in Figure 1. It is seen that the correlation coefficient related to items vary between .43 and .82. According to this obtained evidence, the Turkish version of ATEMI is in one-to-one correspondence with the original version and the 31-item construct consisting of 3 factors is valid hypothetically and statistically.

Reliability

The reliability studies of the inventory were analyzed with two different methods that are internal consistency and stability. The consistency of the inventory was analyzed with internal consistency coefficient of Cronbach's alpha and the stability of the inventory was analyzed with test-retest method. The inventory was applied twice, one after three weeks, on 50 pre-service teachers who did not participate in the construct validity studies for the test-retest method. The data about the internal consistency of the inventory is indicated in Table 6.

Table 6: The Data about the Internal Consistency of the Inventory

Factor	Number of Items	Cronbach Alpha Coefficient
Relevance	13	.88
Affective	10	.83
Course	8	.86
Total	30	.92

According to Cronbach (1951), it can be acceptable if the internal consistency coefficient is between .7 and .8; it is good if it is between .8 and .9; it is perfect if it is between .9 and 1. Therefore, that all of the internal consistency values of the inventory are higher than .83 shows that the reliability of the inventory is good, which means it generates consistent evidence. The evidence obtained through test-retest method was compared to Pearson product-moment correlation. The data about the correlation analysis is indicated in Table 7.

Table 7: The Data about Test-Retest Method

Factors	Application	M	SD	r
Relevance	First application	54.72	5.27	.74
	Second application	52.90	6.06	
Affective	First application	37.58	4.34	.71
	Second application	36.08	5.16	
Course	First application	29.78	3.07	.80
	Second application	28.34	4.33	
Total	First application	122.08	10.69	.78
	Second application	117.32	12.01	

When analyzing the evidence obtained through the test-retest study of the inventory, it is seen that the correlation coefficients are .74, .71 and .81, respectively for the three dimensions of the inventory, and it is .78 throughout the inventory. Tavşancıl

(2002) states that the correlation values calculated with test-retest method should at least be .70 to be considered high. According to this, it can be said that the inventory generates consistent results because the inventory's reliability coefficients calculated with test-retest method are high throughout the inventory and in the dimensions of the inventory.

Item Discriminant

Firstly, total correlations of items were calculated to determine how efficient each item in the inventory is to distinguish the pre-service teachers in terms of their attitudes towards assessment and evaluation in education. Secondarily, t-test was used for the significance of the difference between item points of upper-group with 27% and sub-group with 27% according to total points. The results are indicated in Table 8.

Table 8: Item-Total Correlations and t-test for Item Discriminant

Items	Corrected Item-Total Correlations	t	Items	Corrected Item-Total Correlations	t
i1	.550	9.591*	i17	.508	10.319*
i2	.619	11.557*	i18	.348	6.303*
i3	.600	10.158*	i19	.334	5.401*
i4	.453	8.204*	i20	.497	9.370*
i5	.543	9.662*	i21	.535	12.082*
i6	.583	10.791*	i22	.534	9.336*
i7	.401	7.287*	i23	.532	9.879*
i8	.602	11.439*	i24	.591	12.500*
i9	.576	10.876*	i25	.551	10.493*
i10	.458	7.895*	i26	.660	14.181*
i11	.543	10.458*	i27	.545	10.416*
i12	.495	11.273*	i28	.542	10.998*
i13	.437	8.104*	i29	.607	13.743*
i14	.614	11.441*	i30	.545	11.434*
i15	.386	8.297*	i31	.524	9.961*
i16	.387	7.989*			

*p<.001

According to the obtained evidence, it is seen that the adjusted item-total correlations are ranged between .33 and .66, and t (sd=192) value related to the difference of lower and upper groups with 27% determined according to the total points varies between 5.40 (p<.001) and 14.18 (p<.001). .30 is the acceptable boundary for item-total correlations (Nunnaly & Bernstein, 1994). The t-test results showed that the item-total point of upper-group with 27% for all the items is significantly higher (p<0.001) than the same point of lower-group with 27%.

There are 31 items arranged in 3 factors in the Turkish form of Attitude toward Educational Measurement Inventory. The highest point from the inventory which has a

rank order of a five-level likert item and is rank-ordered as 'strongly disagree' (1) and 'strongly agree' (5) is 155, and the lowest point is 31. Reverse scoring should be done because the items 1, 2, 3, 9, 10, 12, 14, 15, 16, 17, 19, 20, 21, 22, 26, 29 and 31 have negative expressions. The higher points got from each sub-dimension of the inventory show that the individual has a positive attitude toward the related sub-inventory. The total points got from the inventory present the general points of the attitude of the individual toward assessment and evaluation in education.

RESULTS AND DISCUSSION

In this study, linguistic equivalence, validity and reliability studies of ATEMI which was developed by Bryant & Barnes (1997) to measure the attitudes of pre-service teachers toward measurement and evaluation in education are done on pre-service teachers in Turkey. In this context, the relation between the Turkish and English forms of Attitude toward Educational Measurement Inventory was firstly calculated to present the linguistic equivalence with the original form of the inventory, and high correlation was found for both sub-factors and the whole inventory. This result shows that ATEMI has adequate linguistic equivalence with the original form of the inventory and the translation process has been completed successfully.

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were used to analyze the construct validity of the inventory. A three factors construct consisting of 31 items, which explain 47.4% of total variance and the items are completely in their dimensions as in the original form, was obtained according to results obtained from exploratory factor analysis (EFA). It was found that the factor loadings of the inventory items vary between the values of .31 and .83. The first dimension was called 'Relevance' and consists of 13 items, the second dimension was called 'Affective' and consists of 10 items, and the third dimension was called 'Course' and consists of 8 items. The 'Relevance' dimension is intended for the importance and practicality of assessment and evaluation and therefore it reflects the interests of the pre-service teachers for assessment and evaluation. The second dimension, 'Affective', reflects the feelings of the pre-service teachers for assessment and evaluation course, especially mathematically. The third dimension called 'Course' reflects the attitudes of the pre-service teachers toward taking an assessment and evaluation course. The items in this dimension are about the pre-service teachers' pleasures of the course and desire to learn much more about assessment and evaluation course. In another factor analysis, CFA, it was analyzed that if the original form of the inventory confirms with this study carried out on Turkish students. Considering the fit index boundaries for CFA, it was determined that the model fits in a sufficient level and the original factor construct of the inventory squares with the Turkish form. Item analysis and comparison of sub and upper groups with 27% were done to present the item discrimination and the degree of items' total point prediction of ATEMI. It was found that the results of adjusted item-total correlations obtained as a result of item analysis vary between .33 and .66. Considering the items with values .30 or higher which have an acceptable adequate feature to measure in terms of discrimination (Büyüköztürk, 2010), it can be said that the consistency of item-total correlations related to inventory is sufficient in

interpretation of item-total correlations. T values of the difference of item points of lower and upper groups with 27% vary between 5.40 (p<.000) and 14,18 (p<.000). These results show that the item-total points of the upper-group with 27% are msignificantly higher (p<0.001) than the same points of the lower-group's with 27% considering all the items.

According to the reliability studies of ATEMI, it was found that the internal consistency correlation of Cronbach's Alpha is 0,93 and it is 0,88, 0,83 and 0,86 for the sub factors, respectively. However, it was calculated that the correlation coefficient obtained with test-retest methods is 0,78 and it is 0,74, 0,71 and 0,80 for the sub-factors, respectively. The obtained evidence indicates that the measurement instrument whose adaptation study has been done is considerably reliable.

Another adaptation study of the original inventory was done by Alkharusi (2011) and it was determined that the inventory is a valid and reliable measurement instrument to determine the attitudes of pre-service teachers in Oman toward measurement and evaluation in education. In this study, the inventory also preserved its 3-factor construct and the items were gathered in the same factors as in the original inventory. Bases upon the results of the validity and reliability studies, it was determined that the adapted Turkish form of the inventory is a valid and reliable instrument that can be used to determine the attitudes of the pre-service teachers toward measurement and evaluation in education. As the inventory is applied on the pre-service teachers taking measurement and evaluation course in teacher training programs in the beginning of the school term, it can be important for the instructors to determine the negative attitudes of the pre-service teachers toward the course and to take precautions and develop training strategies for this situation. Besides, the inventory can be used in studies that researches if the lecture the instructors give in the measurement and evaluation course affects the attitudes of the pre-service teachers toward assessment and evaluation. It is thought that the application of the inventory on different teacher groups in further studies will contribute to the validity and the reliability of the inventory.

References

Alkharusi, H. (2009). Correlates of teacher education students' academic performance in an educational measurement course. International Journal of Learning, 16, 1-15.

Alkharusi, H., Kazem, A., and Al-Musawi, A. (2011). Knowledge, skills, and attitudes of preservice and inservice teachers in educational measurement. Asia-Pacific Journal of Teacher Education, 39 (2), 113-123.

Alkharusi, H. (2011). Validity and reliability of the attitude toward educational measurement inventory. The Asia-Pacific Education Researcher, 20 (3), 612-620.

Balcı, E., and Tekkaya, C. (2000). Ölçme ve değerlendirme tekniklerine yönelik bir ölçeğin geliştirilmesi. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 18, 42-50.

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

Bryant, N. C., and Barnes, L. L. B. (1997). Development and validation of the attitude toward educational measurement inventory. Educational and Psychological Measurement, 57, 870-875.

Büyüköztürk, Ş. (2010). Sosyal Bilimler İçin Veri Analizi El Kitabı. On birinci basım. Ankara: Pegem Akademi Yayıncılık.

Büyüköztürk, Ş., Akgün, Ö. E., Özkahveci, Ö., and Demirel, F. (2004). The validity and reliability study of the Turkish version of the Motivated Strategies for Learning Questionnaire. Educational Sciences: Theory & Practice, 4 (2), 231-239.

Byrne, B. M., and Campbell, T. L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. Journal of Cross-Cultural Psychology, 30 (4), 555-574.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16 (3), 297–334.

Ekici, G. (2008). Öğretmen adaylarının öğretmenlik meslek bilgisi derslerine yönelik tutumları ile öğrenme biçimlerinin değerlendirilmesi. Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi, *5* (1), 111-132.

Erden, M. (1995). Öğretmen adaylarının öğretmenlik sertifikası derslerine yönelik tutumları. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 11, 99-104.

Field, A. (2009). Discovering Statistics Using SPSS. 3rd edition. London: Sage Publications Ltd.

Gefen, D., and Straub, D. W. (2000). The relative importance of perceived ease of use in IS adoption: A study of e-commerce adoption. Journal of the Association for Information Systems, 1 (1), 1-28.

Gorsuch, R. L. (1974). Factor Analysis. Philadelphia: W. B. Saunders Company.

Gülbahar, Y., and Büyüköztürk, Ş. (2008). Değerlendirme tercihleri ölçeğinin Türkçeye uyarlanması. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 35, 148-161.

Hair, J. F. Jr., Black, W. C., Babin, B. J., and Anderson, R. E. (2010). Multivariate Data Analysis. 7th edition. New Jersey: Pearson Prentice Hall.

Hooper, D., Coughlan, J., and Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. Journal of Business Research Methods, 6, 53–60.

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6 (1), 1-55.

Kaiser, H. F. (1974). An index of factorial simplicity. Psychometrika, 39 (1), 31-36.

Karaca, E. (2006). Öğretimde planlama ve değerlendirme dersine yönelik bir tutum ölçeği geliştirme. Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 16, 213-230.

Kılınç, A., and Salman, S. (2007). Okul deneyimi derslerine yönelik tutum ölçeğinin geliştirilmesi. Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi, 27(1), 23-35.

Kline, P. (2000). An Easy Guide to Factor Analysis. New York: Routledge.

Kottke, J. L. (2000). Mathematical proficiency, statistics knowledge, attitudes toward statistics, and measurement course performance. The College Student Journal, 34, 334-347.

Linn, R. L., and Gronlund, N. E. (1995). Measurement and assessment in teaching. 7th edition. Englewood Cliffs, NJ: Merrill.

Mertler, C. A. (2003). Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference? Paper presented at the meeting of the Mid-Western Educational Research Association, Columbus, OH.

Nunnally, J., and Bernstein, I. (1994). Psychometric Theory. New York: McGraw-Hill.

Oğuz, A. (2012). Developing a scale for attitudes towards the curriculum and instruction course. Educational Science: Theory & Practice, 12 (2), 854-860.

Pallant, J. (2005). SPSS Survival Manual: A Step By Step Guide to Data Analysis Using SPSS for Windows Version 12. 2nd edition. Crows Nest NSW, Australia: Allen & Unwin.

Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. Mid-Western Educational Researcher, 6, 21-27.

Popham, W. J. (2006). Needed: A dose of assessment literacy. Educational Leadership, 63 (6), 84-85.

Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. Türk Psikoloji Yazıları, 3 (6), 49-74.

Tabachnick, B. G., and Fidell, L. S. (2007). Using Multivariate Statistics. 5th edition. Boston: Pearson Education Inc.

Tavsancıl, E. (2002). Tutumların Ölçülmesi ve SPSS`le Veri Analizi. Ankara: Nobel Yayıncılık.

VanZile-Tamsen, C., and Boes, S. R. (1997). Graduate students` attitudes and anxiety toward two required courses: Career development and tests and measurement. Paper presented at the meeting of the Georgia Educational Research Association, Atlanta, GA.

Yaman, S., and Karamustafaoğlu, S. (2011). Öğretmen adaylarının ölçme ve değerlendirme alanına yönelik yeterlik algı düzeylerinin incelenmesi. Ankara University Journal of Faculty of Educational Sciences, 44 (2), 53-72.

Yüksel, S. (2009). Eğitim fakültesi öğrencilerinin öğretmenlik meslek bilgisi derslerine ilişkin görüşleri. Ankara University Journal of Faculty of Educational Sciences, 42 (1), 435-455.

Appendix 1.

Attitude toward Educational Measurement Inventory
(Eğitimde Ölçme ve Değerlendirmeye Yönelik Tutum Envanteri)

Below are a number of statements about measurement. Read each item carefully and express the extent of agreement between the feeling expressed in each statement and your feeling on a five-point scale. There are no right or wrong answers. Circle the letter which best indicates how closely you agree or disagree with the feeling expressed in each statement.

1 = Totally Disagree
2 = Disagree
3 = Somewhat Agree
4 = Agree
5 = Totally Agree

	Original Version	Turkish Version	
1	Measurement training is not really useful for most teachers.	Ölçme ve değerlendirme eğitimi çoğu öğretmen için aslında çok yararlı değildir.	1 2 3 4 5
2	I see no value in measurement.	Ölçme ve değerlendirmeyi değerli bir şey olarak görmüyorum.	1 2 3 4 5
3	Studying measurement is a waste of time.	Ölçme ve değerlendirme çalışmak vakit kaybıdır.	1 2 3 4 5
4	A good teacher must have training in measurement.	İyi bir öğretmen ölçme ve değerlendirme eğitimi almış olmalıdır.	1 2 3 4 5
5	Most teachers would benefit from taking a measurement course.	Ölçme ve değerlendirme dersi almaları öğretmenler için yararlı olur.	1 2 3 4 5
6	Measurement will be useful to me in my profession.	Ölçme ve değerlendirme mesleğim açısından bana yararlı olacaktır.	1 2 3 4 5
7	I am excited at the prospect of actually using measurement in my field.	Alanımda ölçme ve değerlendirme çalışmaları yapacağımdan dolayı heyecanlıyım.	1 2 3 4 5
8	Measurement is an inseparable aspect of teaching.	Ölçme ve değerlendirme öğretimin ayrılmaz bir parçasıdır.	1 2 3 4 5
9	I have difficulty seeing how measurement relates to my field of study.	Ölçme ve değerlendirmenin eğitim alanımla nasıl bir ilişkisi olduğunu görmekte zorlanıyorum.	1 2 3 4 5
10	Measurement course should not be required for any reason.	Ölçme ve değerlendirme dersi her ne sebeple olursa olsun zorunlu olmamalıdır.	1 2 3 4 5
11	Measurement is an excellent tool for teaching.	Ölçme ve değerlendirme öğretim için mükemmel bir araçtır.	1 2 3 4 5
12	Measurement is not really useful because it is seldom used in a school setting.	Ölçme ve değerlendirme okul ortamında nadiren kullanıldığından çok yararlı değildir.	1 2 3 4 5
13	Measurement is a valid way to evaluate students learning.	Ölçme ve değerlendirme öğrencilerin öğrenmelerini değerlendirmenin geçerli bir yoludur.	1 2 3 4 5

14	Just thinking of taking a measurement course scares me.	Ölçme ve değerlendirme dersi alma düşüncesi bile beni korkutuyor.	1 2 3 4 5
15	I feel insecure in a measurement course.	Ölçme ve değerlendirme dersinde kendime güvenmiyorum.	1 2 3 4 5
16	I feel intimidated when I have to deal with mathematical formulas.	Matematiksel formüllerle uğraşmak zorunda kalacağım için çekiniyorum.	1 2 3 4 5
17	I do not like measurement because I have to deal with math.	Matematikle uğraşmak zorunda olduğum için ölçme ve değerlendirmeyi sevmiyorum.	1 2 3 4 5
18	I feel at ease in math, so I feel comfortable with measurement.	Matematikte iyi olduğum için ölçme ve değerlendirmede kendimi rahat hissediyorum.	1 2 3 4 5
19	Measurement is too mathematically oriented for me to use effectively.	Ölçme ve değerlendirme matematiğe dayalı olmasaydı daha etkin kullanabilirdim.	1 2 3 4 5
20	When I hear the word measurement, I feel frightened.	Ölçme ve değerlendirme kelimelerini duyduğumda korkuyorum.	1 2 3 4 5
21	Measurement is very difficult.	Ölçme ve değerlendirme çok zor.	1 2 3 4 5
22	Being enrolled in a measurement course is an unpleasant experience.	Ölçme ve değerlendirme dersini almak hoş olmayan bir deneyimdir.	1 2 3 4 5
23	I have no antagonistic feeling toward measurement.	Ölçme ve değerlendirme dersine karşı herhangi bir olumsuz duyguya sahip değilim.	1 2 3 4 5
24	I would like to take more measurement courses in the future.	Gelecekte de ölçme ve değerlendirme dersi almak isterim.	1 2 3 4 5
25	I am willing to take more measurement course.	Daha çok ölçme ve değerlendirme dersleri almak isterim.	1 2 3 4 5
26	I wish that I could have avoided taking a measurement course.	Bir ölçme ve değerlendirme dersini almaktan kurtulabilseydim.	1 2 3 4 5
27	I enjoy taking a measurement course because it is fun.	Eğlenceli olduğu için, ölçme ve değerlendirme dersi almaktan mutluyum.	1 2 3 4 5
28	More measurement courses should be required for teachers.	Öğretmenlerin, daha çok ölçme ve değerlendirme dersi almaları gereklidir.	1 2 3 4 5
29	I would only take a measurement course if it were required.	Sadece, zorunlu olursa ölçme ve değerlendirme dersi alırım.	1 2 3 4 5
30	My likes for measurement course outweigh my dislikes.	Ölçme ve değerlendirme dersinde hoşlandıklarım hoşlanmadıklarımdan daha fazladır.	1 2 3 4 5
31	I feel I have been well prepared to be a teacher without taking a measurement course.	Bir ölçme ve değerlendirme dersi almadan da, öğretmenliğe hazır olduğumu hissediyorum.	1 2 3 4 5

e-international journal of educational research
Volume: 4 Issue: 2 Spring 2013 pp. 29-47

Eğitimde Ölçme ve Değerlendirmeye Yönelik Tutum Envanterinin Türkçeye Uyarlanması

Arş.Grv.Ceyhun Ozan
Atatürk Üniversitesi-Türkiye
ozanceyhun08@gmail.com

Doç.Dr.Erdoğan Köse
Atatürk Üniversitesi-Türkiye
erdogan63@gmail.com

Genişletilmiş Özet

Problem: Bu çalışmanın amacı, 1997 yılında Bryant ve Barnes tarafından geliştirilen ve orijinal adı "Attitude Toward Educational Measurement Inventory" olan "Eğitimde Ölçme ve Değerlendirmeye Yönelik Tutum Envanteri"nin Türkçeye uyarlama, geçerlik ve güvenirlik çalışmalarını yapmaktır.

Yöntem: Araştırmada, ölçeğin dilsel eşdeğerlik, geçerlik ve güvenirlik çalışmaları için üç farklı çalışma grubu ile çalışılmıştır. Üç çalışma grubu da 2011-2012 eğitim-öğretim yılında Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi`nde öğrenim gören öğretmen adaylarından oluşmaktadır. Envanter, ölçme ve değerlendirme dersi ile ilgili maddeler içerdiğinden çalışma gruplarının bu dersi almakta olan öğretmen adaylarından oluşmasına dikkat edilmiştir. Orijinal envanter, ilgi (relevance) 13 madde, duyuşsal (affective) 10 madde, ve ders (course) 8 madde olmak üzere üç faktör ve 31 maddeden oluşmaktadır. İlk aşamada envanterin dilsel eşdeğerlik çalışmaları yapılmıştır. Öncelikle envanterin İngilizce formu araştırmacı tarafından Türkçeye çevrilmiş, ardından yurt dışında lisansüstü eğitimini yapmış, alana ve her iki dile hâkim 3 öğretim üyesi tarafından incelenmiş ve gerekli görülen düzeltmeler yapılmıştır. Daha sonra Türkçe form anlam ve gramer açısından incelenerek gerekli düzeltmeler yapılmış ve denemelik Türkçe form elde edilmiştir. Envanterin İngilizce ve Türkçe formları, İngilizce öğretmenliği programı I. öğretim 3. sınıfta öğrenim gören 80 öğretmen adayına üç hafta arayla uygulanmıştır. Envanterin yapı geçerliği açımlayıcı faktör analizi ve doğrulayıcı faktör analizi ile incelenmiştir.

Bulgular: Envanterin orijinal formu ile Türkçe formundan elde edilen puanlar arasındaki korelasyon katsayılarının .70 ile .93 arasında olduğu belirlenmiştir. Buna göre envanterde yer alan bütün maddeler için İngilizce ve Türkçe formların dilsel eşdeğerliğinin sağlandığı kabul edilmiştir. Yapılan açımlayıcı faktör analizi sonucunda toplam varyansın %47.4`ünü açıklayan ve maddelerinin tamamıyla orijinal formdaki alt boyutlarında yer aldığı 3 faktörlü bir yapı ortaya çıkmıştır. Ayrıca tüm maddelerin faktör yükleri .30'dan yüksek olduğu için envanterden herhangi bir maddenin çıkarılması gerekmemiştir. Yapılan doğrulayıcı faktör analizi sonuçlarına göre de ki-kare değerinin anlamlı olduğu ve ki-kare değerinin serbestlik derecesine oranının da kabul edilebilir değer aralığı olan 2 ve 3 arasında olduğu belirlenmiştir. Diğer uyum indekslerine ilişkin değerler incelendiğinde de hepsinin kabul edilebilir sınırlar içerisinde yer aldığı ve elde edilen bu sonuçlara göre model-veri uyumunun sağlandığı

e-uluslararası eğitim araştırmaları dergisi
Cilt: 4 Sayı: 2 Bahar 2013 ss. 29-47

ve envanterin yapı geçerliğinin doğrulandığı sonucuna ulaşılmıştır. Envanterin test-tekrar test çalışmasından elde edilen bulgularına göre korelasyon katsayılarının üç alt boyut için sırasıyla .74, .71, ve .80 olduğu envanterin geneli içinde .78 olduğu hesaplanmıştır. Buna göre envanterin içerdiği boyutların ve envanterin genelinin, test-tekrar test yöntemiyle belirlenen güvenirlik katsayılarının yüksek olduğu ve dolayısıyla ölçeğin kararlı sonuçlar verdiği sonucuna ulaşılmıştır.

Sonuç ve Öneriler: Geçerlik ve güvenirlik çalışmalarından elde edilen sonuçlara dayanarak envanterin Türkçeye uyarlanan formunun öğretmen adaylarının eğitimde ölçme ve değerlendirmeye yönelik tutumlarının belirlenmesinde geçerli ve güvenilir bir araç olarak kullanılabileceği sonucuna ulaşılmıştır. Envanterin, öğretmen yetiştirme programlarında ölçme ve değerlendirme dersi alan öğretmen adaylarına dönemin başında uygulanması, öğretmen adaylarının derse karşı olumsuz tutumlarının belirlenmesi, öğretim elemanlarının bu duruma yönelik önlemler alması ve öğretim stratejileri geliştirmesi açısından önemli olabilir. Ayrıca envanter, öğretim elemanlarının verdikleri ölçme ve değerlendirme dersindeki öğretimin, öğretmen adaylarının eğitimde ölçme ve değerlendirmeye yönelik tutumlarını etkileyip etkilemediğine yönelik araştırmalarda da kullanılabilir.

Anahtar kelimeler: Eğitim, Ölçme ve değerlendirme, Öğretmen adayı, Tutum