Development of an Attitudes Scale toward Online Assessment

3 authors, including:

Emrah Gül
Hakkari University
8 PUBLICATIONS **5** CITATIONS

SEE PROFILE

Ömay Çokluk Bökeoğlu
Ankara University
24 PUBLICATIONS **142** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Yükseköğretimde Öğrenci Başarılarının Sınıflandırılmasında Yapay Sinir Ağları ve Lojistik Regresyon Yöntemlerinin Kullanılması View project

INTE 2014

Development of an Attitudes Scale toward Online Assessment

Emrah GÜL[a] [*] Ömay ÇOKLUK[b], Çilem DOĞAN GÜL[c]

[a]Ankara University, Faculty of Educational Sciences, Department of Measurement and Evaluation, Ankara, Turkey
[b] Ankara University, Faculty of Educational Sciences, Department of Measurement and Evaluation, Ankara, Turkey
[c]Ankara University, Faculty of Educational Sciences, Department of Measurement and Evaluation, Ankara, Turkey

Abstract

Measurement and evaluation applications via Internet have an important position in education and psychology. Online assessment provides considerable opportunities particularly in teaching and education to monitor students' academic achievements when combined with well designed learning goals. The purpose of the study is to develop "Attitudes toward Online Assessment Scale". As a result of the literature review, an initial form that included testing items was obtained and those items were submitted to expert view. In the light of expert views, scale items were rearranged and the initial form was given to a total of 231 pre-service teachers included in online assessment applications (113 from Department of Science and Technology; 118 from Computer Education and Educational Technology). Exploratory and confirmatory factor analyses were performed to determine constructional validity of the scale. As a result of exploratory factor analysis, it was seen that it was a 3-factor-scale consisting of "contribution of online assessment applications to teaching and education", "liking online assessment applications" and "use of online assessment applications" and a total of 26 items. Factor loadings of the scale items ranged from 0.45 to 0.78; item total correlations from 0.20 to 0.76; reliability coefficients from 0.62 to 0.91. What's more, the 3-factor construction obtained from the exploratory factor analyses was tested by confirmatory factor analyses and it was seen that the tested model was confirmed with high fit indices. The findings obtained from the study showed that the "An Attitudes Scale Toward Online Assessment " was valid and reliable.

* Corresponding author. Emrah GÜL. Tel: (0 312) 363 33 50
 E-mail address: egul@ankara.edu.tr

1. Introduction

Education is undoubtedly the most crucial process in shaping and improving societies. Effective and fruitful utilization of the process is critical for educational goal attainment. The process requires proper planning, implementation and evaluation. It is known that different methods used to increase students' academic achievements make significant contributions to the elevation of education to desired levels. Inarguably, the concept of attitude, which increases academic achievement and constitutes an important part of learning-teaching processes, has been one of the most significant study fields of education and psychology for years.

An attitude is a cognitive, affective and behavioral response to any psychological objects or attitude objects in the environment or to social issues or incidents, organized by individuals based on their experience, motivation and knowledge. Smith (1968) defines attitude as "a tendency that regularly constitutes people's thoughts, emotions and behaviors toward a psychological object". The psychological object means any objects that make sense to an individual or that an individual is aware of whereas an attitude object is a thing to which one has an attitude. An attitude is a phenomenon that is learned, shapes people's behaviors and causes bias in decision making process. Attitudes are formed as a result of experiences and learning process (Tavşancıl, 2006).

In the age of technology, computers have become attitude objects. Therefore, students' attitudes toward both computers and online applications are crucially important. When educational literature development is considered, online applications have also led to significant developments in education, as in many fields, and have been used by many researchers. Research on online applications particularly in measurement and evaluation applications, an indispensible part of teaching, has shown the significance of assessment (Robles and Braathen, 2002; Wen and Tsai, 2006).

Online assessment can be used for the evaluation of academic achievement. Online assessment significantly contributes to learning when assessed because it provides students with prompt feedback on learned information or concepts. In consideration of practices in traditional classes, generally a few students, when asked a question by teachers, have the chance of answering that question and it is not possible to assess whether others have understood the topics covered during the lesson or not. When students are asked questions online, teachers might give instant feedback, seeing each student's question while teaching and thus teaching becomes much more successful (Robles and Braathen, 2002). We may suggest that the number of applications in higher education where online assessment is effectively used is rather restricted although online assessment methods have the above mentioned advantages (Wen and Tsai, 2006). The number of educational software is increasing every day and such software is offered to students' access. It is essential to pay attention to technological developments also in measurement and evaluation since technology has become integrated with education. Offering online tests for evaluation of students' performances, and familiarization and guidance will make both test applications more economical and reporting process faster. What's more, confidentiality of tests will be controlled more.

Generally speaking, when online test applications are historically considered, it is seen that such applications are not new and can be traced back to the 1970s (Walter, Daeschener, Hoffman and Lyons, 1970). As a result of the online test application studies by The US Army and clinical psychologists, some advantages such as unbiased scorers and controlled test validity have been mentioned. Until 1985, research showed that especially when psychological tests were applied online, there were no significant differences between the test scores obtained from paper and pencil forms and online forms to measure intelligence and abilities and high correlations were observed between the forms (Mills, Potenza, Fremer, Ward, 2002; Russel, Goldberg, O'Connor, 2003).

Use of online assessment applications in general ability and intelligence tests has advantages, especially in drawing shapes, visual aid utilization and so on. In their study, Pellegrino, Hunt, Abate and Farr (1987) suggested people could be offered stable and mobile objects on account of computer technology and also variables such as response time in answering test items could be measured.

Attitudes, of course, are good predictors of academic achievement. Attitudes of students toward online assessment are critical as they enable them to use online applications effectively. In recent years, it has been observed that technology was largely employed in most educational studies. However, it is also clear that the number of studies to explore attitudes of students toward online assessment is not high (Dermo, 2009). Generally, in technology supported research, attitudes toward computers and technology are measured and academic achievement is attempted to be predicted by those variables, but attitudes of students toward online assessment applications might be negative although their attitudes toward computers are positive (Bindak and Çelik, 2006; Ergün, 2002).

PISA Report (2010) of Computer Based Assessment of Student Skills in Science, it examined whether score means obtained from a paper and pencil test and an online test version of student skills in science significantly varied. Data analysis showed that there were no significant differences between the test scores of the students obtained from the paper and pencil form and the online form of the test, but the male participants got higher scores in the online version of the test when compared to the female ones. However, it was suggested that the high performances of the male participants in the online version of the test could not be predicted by those variables as there were no significant differences between motivation, computer familiarization and enjoyment and the test performances.

The discussions above show that it is essential to examine attitudes and views of participant students in online assessment applications because attitudes toward assessment influence the face validity of assessment (Anastasi, 1982). If students do not trust tests they have taken, then they will not give reliable reactions to test stimulants (Domino and Domino, 2006). The attitude variables observed in the above mentioned studies are not sufficient to predict online applications. The main purpose of the research is to develop a reliable, valid attitude scale for online assessment.

1. "Attitudes toward Online Assessment Scale (AOAS)" Development Steps

1.1. Step I: *Item Writing*

For the AOAS development, the literature was reviewed first. In the review, assessment studies were taken into account and a 52-item item pool was started. It was particularly regarded that the scale items would be about cognitive, affective and behavioral components of attitudes and would include statements in present tense and first person narrative perspective in general (Erkuş, 2003). In wording, the items that would cause similar responses from individuals with different attitudes were considered as they were known to be insufficient in attitude measurement. The numbers of positive and negative expressions were converged as much as possible. Out of 52 items, 26 were arranged to serve cognitive components, 14 to serve affective components and 12 to serve behavioral components and 29 of them were positively constructed while 23 items were negatively constructed.

The initial form consisting of a total of 52 items was rearranged pursuant to expert view. The scale was assessed by three experts of measurement and evaluation according to suitability for attitude components. The scale was finalized by extracting the items which did not give good fit in factors and did not have the features of attitude components. Then, the scale was examined by a linguist in terms of language and wording. After the necessary changes, the number of the scale items was lowered to 41 and was ready for the pre-test application. The participants were asked to mark their attitudes on a 5 point Likert type scale with the following options: "I totally agree", "I agree", "I am uncertain", "I disagree" "I totally disagree".

1.2. Step II: *Pilot Study and Study Group*

The study group of the research consisted of a total of 213 pre-service teachers from Department of Computer Education and Educational Technology (N=113) and Department of Science Teaching (118) of Fırat University and Atatürk University. During the selection of the study group, pre-service teachers' familiarization with online assessment applications was the main concern; in other words, they were expected to have used those applications for a semester at least. In this context, it can be said that the study group consisted of those who constantly used online assessment applications and grasped the nature of online assessment. 135 of the pre-service teachers included in the study group were female (58.4%) while 96 were male (41.5%).

1.3. Step III: *Validity Studies of the AOAS*

In abstract concepts, like attitudes, that could not be directly observed, presenting the obtained scores and the relevant construct or evidence for the existence of a construct which is considered to be relevant to scale scores are important in that the measurement tool serves the intended measurement. Therefore, construct validity study is done to see whether the measurement tool which is developed to measure a construct predicted by scientists actually measures the estimated construct (Erkuş, 2003; Tavşancıl, 2002).

Construct validity is the extent to which a measurement tool shows theoretical construct that is intended to be measured by that tool (Anastasi, 1982). Factor analysis is the most frequently used method for construct validity. Exploratory factor analysis was performed to determine validity of the AOAS.

1.3.1. *Exploratory Factor Analysis (EFA)*

Kaiser-Meyer-Olkin (KMO) test, before factor analysis application, was employed to test factorization fit of the sample size. As a result of the analysis, KMO value of the sample that consisted of 231 participants was found 0.93. In the light of this finding, the sample size was considered "excellent" for factor analysis (Tavşancıl, 2006; Çokluk, Şekercioğlu and Büyüköztürk, 2012). Also, the result of Bartlett Sphericity Test was significant (χ^2: 3413.28, $p < 0.01$). These findings showed that the data set had good fit for factor analysis. In order to show factor construct of the scale, principal components analysis, as the factorization method, and Varimax, one of vertical rotation methods, were employed. 0.32 was taken as the minimum limit for factor loadings while deciding whether the scale items were to be extracted or not.

As a result of exploratory factor analysis, the number of factors was decided in consideration with eigen values and scree plot graphs. It was seen that three factors had eigen values higher than 1 and the split in scree plot was in the third factor and a plateau pattern was seen on the graph followed by that value. Thus, a three-factor restriction was decided. As a result of the second factor analysis, total explained variance was found 53.77%. As a result of EFA, 15 items with factor loadings lower than 0.32 and/or with high factor loadings in more than one factor were extracted and re-analyzed. A difference of .10 between factor loadings was taken as the criteria to explore whether the items had high factor loadings in more than one factor or not (Tavşancıl, 2006). The Attitudes toward Online Assessment Scale is grouped under a total of three factors and the rotated factor loadings are presented in Table 1.

As it is clear from Table 1, the first factor entitled "Contribution of Online Assessment Applications to Teaching and Education" of the "Attitudes toward Online Assessment Scale" consisted of 10 items and the vertically rotated factor loadings by using Varimax method ranged from 0.53 to 0.76. When item-total correlations of the items in this factor are considered, it is seen that they ranged from 0.59 to 0.76. The unique variance explained by this factor was 23.58 %.

"Liking Online Assessment Applications", the second factor of the "Attitudes toward Online Assessment Scale" consisted of 11 items and the vertically rotated factor loadings by using Varimax method ranged from 0.49 to 0.78. When item-total correlations of the items in this factor are considered, it is seen that they ranged from 0.63 to 0.73. The unique variance explained by this factor was 22.39 %.
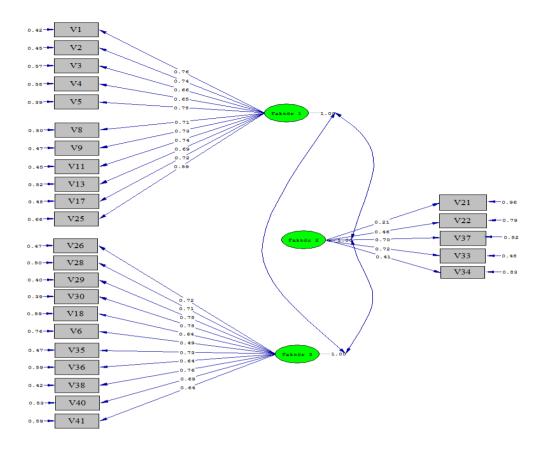
"Use of Online Assessment Applications", the third factor of the "Attitudes toward Online Assessment Scale" consisted of 5 items and the vertically rotated factor loadings by using Varimax method ranged from 0.45 to 0.68. When item-total correlations of the items in this factor are considered, it is seen that they ranged from 0.20 to 0.64. The unique variance explained by this factor was 7.80 %.

Table 1. Exploratory Factor Analysis Results for the Attitudes toward Online Assessment Scale

Factor I: Contribution of Online Assessment Applications to Teaching and Education	Rotated Factor Loading	Corrected Item-total Correlation
Item 1	.76	0.69**
Item 2	.75	0.64**
Item 3	.73	0.73**
Item 4	.72	0.69**
Item 5	.69	0.76**
Item 6	.69	0.59**
Item 7	.64	0.67**
Item 8	.63	0.63**
Item 9	.61	0.71**
Item 10	.53	0.65**
Eigen Value= 11.40 and Explained Variance= 23.58%		
Factor II: Liking Online Assessment Applications		
Item 11	.78	0.69**
Item 12	.71	0.70**
Item 13	.71	0.63**
Item 14	.69	0.67**
Item 15	.62	0.63**
Item 16	.62	0.66**
Item 17	.61	0.69**
Item 18	.59	0.68**
Item 19	.59	0.72**
Item 20	.56	0.73**
Item 21	.49	0.63**
Eigen Value= 1.67 and Explained Variance= 22.39%		
Factor III: Use of Online Assessment Applications		
Item 22	.68	0.20*
Item 23	.59	0.42**
Item 24	.46	0.64**
Item 25	.45	0.37**
Item 26	.45	0.63**
Eigen Value= 1.45 and Explained Variance=7.80%		
χ^2:3413.28, $p < 0.01$		
KMO: .93		
Total Explained Variance: 53,77%		

  **p< 0.01   *p< 0.05

1.3.2. *Confirmatory Factor Analysis (CFA)*

Confirmatory factor analysis is an analysis which tests whether a pre-defined, restricted construct is confirmed as a model (Brown, 2006). The research examined whether the 3-factor construct assessed by EFA was confirmed by CFA. As a result of CFA, fit indices were found χ^2=747,64 (sd=321, p<.001), (χ^2/sd)=2.32, RMSEA=0.07, GFI=0.81 and AGFI=0,77. Coefficients of item-factor correlations calculated by CFA are presented in Figure 1.

Chi-Square=747.64, df=321, P-value=0.00000, RMSEA=0.076

Fig. 1. CFA Results of the AOAS

When CFA results are considered, the construct that was tested to decide if χ^2/df= 2.32 had excellent fit (Kline, 2005). RMSEA=0.07, another fit index, indicates good fit (Hu and Bentler, 1999; Thompson, 2004). It is suggested in the literature that GFI and AGFI values equal to 1 represent excellent fit (Hooper, Coughlan and Mullen, 2008; Kelloway, 1989; Schumacker and Lomax, 1996; Sümer, 2000). In this study, the following values were found: GFI= 0.81 and AGFI=0.77 and they were accepted as sufficient.

1.4. Step IV: *Reliability Studies of the AOAS*

For reliability of the AOAS, Cronbach alpha coefficient was employed. Cronbach alpha coefficients of the overall test and the factors were found 0.93, 0.91, 0.91 and 0.62 respectively. The estimated Cronbach alpha coefficients showed that the scale had high reliability.

2. Results

The study aimed at developing a reliable, valid data gathering instrument to be employed in the exploration of attitudes of pre-service teachers toward online assessment methods. The scale that initially consisted of 41 items was applied to a group of pre-service teachers after a period of preparation, and reliability and validity analyses were performed on the data obtained from the pilot study. As a result of exploratory factor analysis, the scale items were grouped under the following three factors: "contribution of online assessment applications to teaching and education", "liking online assessment applications" and "use of online assessment applications". On the other hand, 15 items were extracted as they had low factor loadings or high factor loadings in more than one factor.

As a result of the exploratory factor analysis of the scale, the construct clustered under three factors was tested by confirmatory factor analysis and χ^2/df, RMSEA, GFI and AGFI fit indices were evaluated and it was observed that all the indices were sufficient in terms of model fit. Also, the estimated Cronbach alpha coefficients showed the scale had high reliability.

As a result, the first factor of the AOAS consisted of 10 items, the second 11 items and the third 5 items. The scale was scored in the following way: 1-I totally disagree and 5-I totally agree. Moreover, a total of nine items in the scale were reversely scored.

When the findings of the exploration of validity and reliability of the "Attitudes toward Online Assessment Scale" are collectively considered, we may suggest that the scale is a reliable, valid tool to be employed in studies to explore attitudes of pre-service teachers toward online assessment. In the light of the findings obtained from the research, it can be suggested that the tool developed in the study will cover a considerable need in the literature and be a measurement tool with psychometrically sufficient qualities for further studies.

References

Anastasi, A. (1982). *Psychological testing*. London: Macmillan.

Bindak, R. ve Çelik, H.C. (2006). Öğretmenler için Bilgisayar Tutum Ölçeğinin Güvenirlik Geçerlik Çalışması, *Eurasian Journal of Educational Research-EJER*. 22: 38 - 47.

Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2012). *Sosyal Bilimler için Çok Değişkenli İstatistik: SPSS ve Lisrel Uygulamaları*. Ankara, Pegem Akademi

Dermo J., (2009). E-Assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology,* 40 (2) 203-214

Domino, G. ve Domino, M. L. (2006). *Psychological testing: an introduction*. Cambridge: Cambridge University Press.

Ergün, E. (2002). Üniversite Öğrencilerinin Bilgisayar Destekli Ölçmeden Elde Ettikleri Başarının Kalem-Kâğıt Testi Başarısı, Bilgisayar Kaygısı ve Bilgisayar Tecrübeleri Açısından İncelenmesi. Yayımlanmamış yüksek lisans tezi. Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü, Eskişehir.

Erkuş, A. (2003). Psikometri Üzerine Yazılar, *Türk Psikologlar Derneği Yayınları* No: 24, Ankara.

Hooper, D., Coughlan, J., ve Mullen, M. (2008). Structural equation modeling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, 6 (1), 53-60.

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

Kelloway, K. E. (1989). *Using LISREL for structural equation modeling*: A researcher's guide. London: Sage.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling.* (Second Edition). NY: Guilford Publications, Inc.

Milli Eğitim Bakanlığı. (2011). Fırsatları Artırma Teknolojiyi İyileştirme Hareketi. Web: http://fatihprojesi.meb.gov.tr/tr/icerikincele.php?id=6

Mills, C. N., Potenza, M.T., Fremer, J.J., Ward, W.C. (2001). *Computer Based Testing: Building the Foundation for Future Assessment*. Lawrance Erlbaum Associates, Publishers: Londra

Özgüven, E. (1994). *Psikolojik Testler*, Ankara: Yeni Doğuş Matbaası

Pellegrino, J.W., Hunt, E.B., Abate, R. ve Farr. S. (1987). A Computer-based test battery fort he assessment of static and dynamic spatial reasoning abilities. *Behavior Research Methods. Instruments & Computers*. 19(2), 231-236.

PİSA (2010). *PISA Computer- Based Assessment of Student Skills in Science*. OECD Publications.

Robles B., ve Braathen S. (2002). Online assessment techniques. *Delta Pi Epsilon Journal,* 44(1).

Russel, M., Goldberg, A., O'Connor, K. (2003). Computer Based Test and Validity: a look back into the future. *Assessment in Education*. 10, 279- 293.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. NJ: Lawrence Erlbaum Associates, Inc.

Smith, M. B. (1968), *Attitude Change, İnternational Encylopedia of the Social Sciences*, Crowell Collier and MacMillan.

Sümer, N. (2000). Yapısal eşitlik modelleri. *Türk Psikoloji Yazıları*, 3 (6), 49-74.

Tabachnick, B. G. ve Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Tavşancıl, E. (2006), *Tutumların Ölçülmesi ve SPSS ile Veri Analizi*, Ankara, Nobel Yayın Dağıtım

Tezbaşaran, A. (1997). *Likert tipi ölçek geliştirme kılavuzu*. Ankara: Türk Psikologlar Derneği Yayınları.

Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. (First Edition). Washington: American Psychological Association.

Walter, J., Daeschner, S., Hoffman, R. ve Lyons, T. (1970). Computer-Test Development Center Project. Nihai Rapor. *Teaching Research*, Oregon

Wen, M.L., ve C.-C. Tsai. 2006. University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education* 51: 27 44