# Review of Educational Software Evaluation Tools

**Res. Asst. Ali GERİŞ**
**Asst. Prof. Ahmet ARSLAN**
**Marmara University, Atatürk Faculty of Education**
**Istanbul, Turkey**
ali.geris@marmara.edu.tr, aarslan@marmara.edu.tr

## Abstract

The purpose of this study is to review existing evaluation tools for educational software and to provide insight and guidance for future projects in this context. In total, 42 studies have been taken into consideration as a result of conducting searches using keywords such as "Educational software evaluation", Software evaluation", "Educational software evaluation forms/scales" in the following databases: "Education Resources Information Centre (ERIC)", "Marmara University e-Library", "Science Direct". Whereas, a set of criteria has been determined for the selection of evaluation tools to be included within the scope of the study and 29 of the initial 42 items have been selected for review. There is an increase in the number of evaluation tools between the years 2006 and 2010. However, there is no sufficient number of evaluation tools targeting "educational games" and this is a remarkable result. It was concluded that reliability and validity studies are a very important part of developing educational software evaluation tools and this is a matter that should be considered in future studies.

**Keywords:** Educational software, evaluation, forms, scales.

**Conference Topic:** Assessment Software Tools

## Introduction

Computer software that is prepared for educational purposes is called "Educational Software"(Kelleci, 2010). Computer-based learning environments and educational software has been used in primary, secondary and high school levels, in addition to institutions that offer distance education (Ateş, 2011). Educational software should be developed in accordance with standards based on teaching-learning principles; in addition the requirements of the target educational program where it will be implemented (Güzeller & Korkmaz, 2007). Kara (2009) has stated that should not be expected any positive results from teaching activities which had been done with low-level software.

The software prepared without considering the elements required for the development of educational software may show negative results against teaching process, students, teachers and teaching purposes (Tankut, 2008). Considering potential negative consequences, we against necessity that particularly in educational software be evaluated as very broad and comprehensive and be considered that evaluations before using this software (Gül & Yeşilyurt, 2011).

Educational software evaluation is to examine effectiveness of the learning software opportunities giving students (Kelly, 2008). But Çeliköz and Erişen (2010) is defined

educational software evaluation as to determine that relationship between software current level and expected level in particular quality. Based on these statements, educational software to satisfy the learning needs of students, to provide experience in accordance with the defined learning objectives and to achieve these so effectively are thought to be very important on evaluation (Ateş, 2011).

Cohen (1983) stated that on the literature had not been yet common decision about evaluation of educational software. Accordingly many studies has been emphasized in with different elements and it has been made evaluation on these elements. Şimşek (1998) is arranged evaluation criteria in his book as sophistication, resolving, objectivity, flexibility, sensitivity and compliance with conditions. Çeliköz and Erişen (2010) said that compatibility, interoperability, effectiveness and efficiency is a quality indicator in educational software. Aslan (2006) has been concluded that in his thesis, educational software collected in two main areas is called "Design Review" and "Evaluation of Wholeness".

It has been seen that evaluation forms or scales based on different criteria as in the evaluation stage in many different studies. When Ateş (2011) considering the educational software evaluation forms scope of they had created six dimensions, Gülbahar and Tınmaz (2006) have created their evaluation forms scope of four dimensions. In the literature it is possible to reach many different size and scope of the established educational software evaluation forms. Due to the lack of a research study on to examine this evaluation forms, it has been seen as an issue that to examine this evaluations forms can be accessed on literature.

## Purpose

Under the light of the literature review and in consideration of the aforementioned problems, the purpose of study has been defined as "Review of educational software evaluation tools in Turkey". In accordance with this main purpose, the following questions have been formulated;

1. How is the distribution of educational software evaluation tools in accordance with their respective year of development?

2. What types of educational software evaluation tools exist?

3. How are the development processes of educational software evaluation tools?

4. How is the status of development work on educational software evaluation tools in terms of relevant validity and reliability studies conducted?

5. What kind of scoring systems do educational software evaluation tools use?

6. From a technical standpoint, what types of software tools have been employed for the development of software evaluation tools?

7. What kind of practitioner groups have been considered as users in the existing educational software evaluation tools?

8. How are the number of dimensions distributed in the existing educational software evaluation tools?

9. How are the number of items distributed in the existing educational software evaluation tools?

## Method

### Research Model

In this study, content analysis method was used for the purpose of reviewing educational software evaluation tools and to provide insight and guidance for future projects through recommendations. Content analysis method is a scientific approach where oral, written and other materials are evaluated objectively and handled through systematic review (Sert, Kurtoğlu, Akıncı, & Seferoğlu, 2012). In the content analysis method, there are four main phases as follows: the processing of qualitative research data obtained from collections of resources, coding of said data, detection of themes and the organization and interpretation of results (Yıldırım & Şimşek, 2006). This here study is based on the fundamentals of the content analysis method and results have been achieved in this manner.

### Survey and Selection Criteria

Some criteria have been determined for the selection of the evaluation tools to be used within the scope of the study. Accordingly, a survey has been conducted using keywords such as "Educational software evaluation", "Web based software evaluation", Software evaluation", "Educational software evaluation forms/scales" in the following databases: "Education Resources Information Centre (ERIC)", "Marmara University e-Library", "Science Direct", "Google Scholar" and "National Thesis Centre".

Studies collected individually by the authors according to specified search criteria were reviewed collaboratively. As a result of this review some studies were eliminated due to repetitions. In total, 42 studies have been reached as a result of this process. Out of the 42 studies, 11 have been left out of the scope of the study due to the reason of not originally including the evaluation form itself, whereas 2 have been left out for not having been developed in the Turkish language. As a results of the examination, it has been determined that five evaluation tools were actually based on secondary sources. But these five evaluation tools were still taken into review because the primary sources themselves could not be accessed by the authors. 29 educational evaluation tools have therefore been selected for this study. Evaluation processes followed certain criteria in an effort to maintain reliability and validity throughout the study.

## Findings

The distribution of years of 29 studies included in the review was found to be as follows: one study in 1996, two in 1998, one in 1999, one in 2002, two in 2004, two in 2005, one in 2006, four in 2007, four in 2008, four in 2010, one in 2011, three in 2013 and in 2014. Whereas 13.8 % of studies were conducted between 1996 and 2000, 20.7 % of them were conducted between 2001 and 2005, 44.8 % of them conducted between 2006 and 2010 and 20.7 % of them were conducted 2011 and 2015.

Two evaluation tools were defined as "Surveys" (6.9 %), sixteen evaluation tools were defined as "Forms" (55.1 %), four evaluation tools were defined as "Criteria Lists" (13.8 %) and evaluation tools were defined as "Scales" (24.2 %) by the authors in 29 studies.

In 29 evaluation tools, one of them (3.5 %) were prepared for "Educational Games", eighteen of them (62.0 %) were prepared for "Educational Software" and ten of them (34.5 %) were prepared for "Web Based Educational Software".

Due to the lack of a specific description, 17.2 % of evaluation tools (5) were defined as "Unspecified". "Secondary Source" was used to define five evaluation tools (17.2 %), which had no description about the processes of preparation for the fact that the main sources included in these studies could not be reached. Ten of them (34.5 %) was prepared on the basis of sheer literature reviews. And also 17.2% of evaluation tools (5) were prepared on the basis of both literature reviews and expert opinion. Four evaluation tools (13.9 %) were created by authors according to their own study purposes and these tools were defined by authors as "Adaptation" types of tools.

Four of 29 evaluation tools (13,8 %) display "3-Point Likert" rating type, nineteen of them (65.5%) have "5-Point Likert" rating type and one of them (3.5 %) has "7-Point Likert" rating type. 10.3 % of evaluation tools (3) have two different rating types so they were defined as "Mixed" type. In addition to this two of them (6.9 %) have "Positive / Negative" rating types. Positive / Negative include in "Yes/No" and "Agree / Disagree" rating types.

Educational software evaluation form should be prepare for certain intended user groups. In our research we determined that 13.8 % of evaluation tools (4) had no target groups and we defined this situation as "Unspecified". Five of them (17.2 %) were prepared on the basis of "Primary and Secondary Class Students". Besides, 37.9 % of them had intended users groups of "Experts", i.e. people who are either teachers or academics. As for tools intended for use with "University Students", eight of them (27.6 %) were prepared by experts. One of the tools (3.5 %) was prepared for primary and secondary students, teachers, academics and experts therefore that evaluation tool was defined as "Mixed".

From the perspective of reliability and validity, we determined that two tools (6.9 %) had no reports regarding reliability and validity concerns. Two of them (6.9 %) were concerned with only validity practices while six of them (20.8 %) had only reliability practices. In addition to this, six of them (20.8 %) displayed both reliability and validity practices, however, thirteen of them (44.6 %) stated that reliability and validity practices were omitted from the study.

Nine evaluation tools (31.1 %) had three dimensions while six evaluation tools (20.7 %) had four dimensions. 10.3 % of evaluation tools (3) were separated into five dimensions by authors. In addition, five of them (17.2 %) displayed six or more dimensions. Whereas, 20.7 % of them (6) did not display any dimension over their respective items.

From the perspective of number of items in evaluation tools, we determined that eleven of them (37.9 %) included 0 to 25 items. Similarly eleven of them (37.9 %) of them included 26 to 50 items. Two of them (6.9 %) included 51 to 75 items. In addition to this, 6.9% of them (2) included 76 to 100 items. Three (10.4 %) included 100 or more items in their evaluation forms.

## Results and Discussion

From the review of literature accessible by the authors, it has been observed that educational software evaluation tools in Turkish language are being developed since 1992. The "Software Evaluation Form" developed by Numanoğlu (1992) in this manner is the first example of secondary sources about educational software evaluation tools. The lack of research on the educational evaluation is so remarkable in the early 2000s. The tools used for educational software in these years seem to be either adapted or borrowed from other contexts. Between 2006 and 2010, educational software and educational software evaluation studies seem to have increased intensely and efforts on developing new educational software evaluation tools grown. It is believed that the statement made by Sim, Macfarlane, and Read (2006) regarding how there are insufficient methods on development of educational software evaluation could be main reason that triggered the popularity in this field. However, between the years 2011 and 2015, there was a decrease in developing educational software evaluation tools although technology and computer usage in education kept on increasing.

With the development of Web services, educational environments like object repositories, or concepts such as web-based learning and open courseware are become more and more widespread (Yiğit, Bütüner, & Dertlioğlu, 2008). Due to this situation, it has been observed that web-based educational software evaluation tools comprise 34.5 % of all the reviewed evaluation tools. Ateş (2011) has stated that educational games increase student motivation towards lessons. However, it has been found that there is only one evaluation form regarding educational games. This situation may point towards a lack of educational game area. Evaluation of the quality and content of these educational environment should help to improve these tools effectiveness.

Considering the development processes of educational software evaluation tools, it has been observed that most of them were developed after an initial review of scientific literature. In order to address concerns regarding validity and reliability, it has been observed that authors generally used methods such as consulting expert opinion, conducting statistical factor analyses, or ask a linguist to review the evaluation forms. Educational software evaluation tools show similarity to one another in terms of number of items and number of dimensions. This causes redundancy in items and dimensions of different evaluation tools. On the other hand, it has been seen that there is a general lack of common methods or opinions in the creation of evaluation tools because there are also many different item themes and dimensions present. Another problem is that, there is no common method for establishing reliability and validity in studies.

Future studies performed in the field of educational software evaluation could be helpful for developing tools assessing reliability and validity of the said tools. Also, it is necessary to make up for the lack of evaluation tools targeting educational games. Considering the last 2 or 3 years, there seems to be a decrease in the number of educational evaluation tools developed. However, it is believed that having evaluation and assessment tools that are up to date and that follow new technology closely, this could prove very effective in terms of enhancing teaching and learning processes.

# Brief Biographies of the Authors

**Asst. Prof. Ahmet Arslan**

He is an Assistant Professor at the Computer and Instructional Technologies Department of Marmara University where he teaches in the area of instructional technologies for learning. Dr. Arslan has published numerous papers on web based education, instructional technologies, educational web design, usability and instructional software evaluation. His current research focuses on web based education, educational web design and instructional software evaluation.


**Res. Asst. Ali Geriş**

He is a Research Assistant at the Computer and Instructional Technologies Department of Marmara University and also Master Degree's candidate own department. His current research focuses on educational web design, instructional software evaluation, educational games and social media on education.

# References

Aslan, Ö. (2006). Lise Matematik Dersinde Kullanılan Eğitsel Yazılım İçeriklerinin Öğrenme Başarısı Yönünden Karşılaştırılması. (Yüksek Lisans Tezi), Marmara Üniversitesi, İstanbul.

Ateş, A. (2011). Eğitsel Yazılım Değerlendirme Ölçeği: Geçerlik ve Güvenirlik Çalışması. Eğitim Teknolojileri Araştırmaları Dergisi, 2(1). Retrieved from http://www.et-ad.net/

Cohen, V. B. (1983). Criteria For The Evaluating Of Microcomputer Coursware. Educational Technology, 23(1), 9-14.

Çeliköz, N., & Erişen, Y. (2010). Eğitimde Bilgisayar Kullanımı. In Ö. Demirel & E. Altun (Eds.), Öğretim Teknolojileri ve Materyal Tasarımı (Vol. 5). Ankara: Pegem A Yayıncılık.

Gül, Ş., & Yeşilyurt, S. (2011). Yapılandırmacı Öğrenme Yaklaşımına Dayalı Bir Ders Yazılımının Hazırlanması ve Değerlendirilmesi. Çukurova Üniversitesi Eğitim Fakültesi Dergisi, 1(40), 19-36.

Gülbahar, Y., & Tınmaz, H. (2006). Implementing Project-Based Learning And E-Portfolio Assessment In an Undergraduate Course. Journal of Research on Technology in Education, 38(3), 309-327.

Güzeller, C., & Korkmaz, Ö. (2007). Bilgisayar Destekli Öğretimde Bir Ders Yazılımı Değerlendirmesi. Kastamonu Eğitim Dergisi, 15(1), 155-168.

Kara, Y. (2009). Biyoloji Öğretimi İçin Hazırlanmış Eğlenceli Eğitim Yazılımı Değerlendirmesi. Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi, 27, 17-30.

Kelleci, Ö. (2010). Bir Eğitsel Yazılım Değerlendirme Formunun Geliştirilmesi ve Uygulanması. (Yüksek Lisans Tezi), Marmara Üniversitesi, İstanbul.

Kelly, G. (2008). A collaborative process for evaluating new educational Technologies. Campus-Wide Information Systems, 25(2), 105-113.

Numanoğlu, M. (1992). MEB BDE projesi bilgisayar destekli öğretim ders yazılımlarında bulunması gereken eğitsel özellikler. (Yüksek Lisans Tezi), Ankara Üniversitesi, Ankara.

Sert, G., Kurtoğlu, M., Akıncı, A., & Seferoğlu, S. S. (2012). Öğretmenlerin Teknoloji Kullanma Durumlarını İnceleyen Araştırmalara Bir Bakış: Bir İçerik Analizi Çalışması. Retrieved from Uşak Üniversitesi:

Sim, G., Macfarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. Computers & Education, 46, 235-248.

Şimşek, N. (1998). Öğretim Amaçlı Bilgisayar Yazılımlarının Değerlendirilmesi: Kavramlar Teknikler, Araçlar ve Uygulama (Vol. İlk Baskı). Ankara: Siyasal Kitabevi.

Tankut, Ü. S. (2008). İlköğretim 7. Sınıf Sosyal Bilgiler Dersinde Bilgisayar Destekli Öğretimin Akademik Başarıya ve Kalıcılığa Etkisi. (Yüksek Lisans Tezi), Çukurova Üniversitesi, Adana.

Yiğit, N., Bütüner, S. Ö., & Dertlioğlu, K. (2008). Öğretim Amaçlı Örütbağ Sitesi Değerlendirme Ölçeği Geliştirme. Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi, 2(2), 38-51.

Yıldırım, A., & Şimşek, H. (2006). Sosyal Bilimlerde Nitel Araştırma Yöntemleri. Ankara: Seçkin Yayıncılık.