Journal of Research on Technology in Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ujrt20

Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review

Prokopia Vlachogianni & Nikolaos Tselios

To cite this article: Prokopia Vlachogianni & Nikolaos Tselios (2021): Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review, Journal of Research on Technology in Education, DOI: 10.1080/15391523.2020.1867938

To link to this article: https://doi.org/10.1080/15391523.2020.1867938

Published online: 23 Feb 2021.

JOURNAL OF RESEARCH ON TECHNOLOGY IN EDUCATION
https://doi.org/10.1080/15391523.2020.1867938

Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review

Prokopia Vlachogianni and Nikolaos Tselios

University of Patras–Patras Campus Rion, Rio Patra, Patra, Greece

ABSTRACT

This article presents the findings of a systematic review of perceived usability of educational technology systems. The research was conducted after studying, organizing, and analyzing the results of 104 research papers evaluating perceived usability of educational technologies using the System Usability Scale (SUS). The results were organized on the basis of (a) the usability score obtained when using the SUS, (b) the type of educational technology used, (c) the subject being learned, (d) the level of education, (e) the type of participant, (f) the age, and (g) the number of participants in each survey. Statistical analysis in all surveys ($N = 170$) demonstrated a good level of usability but with some issues ($M = 70.09$, $SD = 12.98$). The categories of Internet platforms ($M = 66.25$, $SD = 12.42$), university websites ($M = 63.82$, $SD = 16.52$) and affective tutoring systems (ATS) ($M = 68.87$, $SD = 7.30$) seem to have a good usability level according to SUS, preceded by mobile applications ($M = 73.62$, $SD = 13.49$) and multimedia ($M = 76.43$, $SD = 9.45$). Moreover, SUS scores were not found to be significantly related with participants' age ($r = 0.017$, $p = 0.931$, ns), stage of education ($p = 0.539$, ns), or the type of participants ($p = 0.639$, ns). Furthermore, the subject being learned ($p = 0.038$, s) and the number of participants in each survey ($r = -0.259$, $p = 0.001$, s) seem to relate to the obtained SUS scores. A slight, statistically insignificant improvement is noted in the perceived usability over the years ($p = 0.182$, ns). The findings of this review will serve as a useful reference guide for educational technology designers, practitioners, and teachers.

ARTICLE HISTORY
Received 10 August 2020
Revised 16 December 2020
Accepted 19 December 2020

KEYWORDS
Systematic review; usability; System Usability Scale (SUS); learning management systems (LMS); mobile applications; educational technology; e-learning; affective tutoring systems (ATS); university websites; multimedia

Introduction

Educational technology is nowadays an integral part of the learning process. It is present at all levels of education, from preschool to universities, and even in informal settings in different types and forms. Teachers in preschool and primary school use robotic kits, simulations, multimedia, and games. Universities frequently use learning management systems (LMS), mobile applications, and affective tutoring systems; even some corporations use virtual reality (VR) and augmented reality (AR) for learning purposes (e.g., to train their employees). The exponential increase in technology usage in education makes imperative research on its effectiveness in order to manage the available resources as efficiently as possible.

The contribution of information and communication technologies (ICTs) in education is very significant (Livingstone, 2012; De Witte & Rogge, 2014; Piper et al., 2015; Petko et al., 2017; Dushayeva, 2019; Altanopoulou et al., 2015; Safar et al., 2016). Most of them are highly engaging, providing learners the opportunity to learn at their own pace and sometimes at their own

CONTACT Nikolaos Tselios nitse@ece.upatras.gr University of Patras–Patras Campus Rion, Rio Patra, Patra 26504, Greece.

learning style. However, technology usage does not necessarily lead to improved learning outcomes. To achieve an optimal adoption and use of a learning environment, certain dimensions need to be seriously considered to provide a suitable learning experience.

One of the already-mentioned dimensions is the usability of these technological systems for educational purposes and the way they are perceived by learners and other stakeholders to whom they are addressed. According to Mayes and Fowler (1999), educational effectiveness is greatly affected by perceived usability. Although many studies explore aspects such as student achievement (Schacter, 1999; Youssef & Dahmani, 2008), uses (Lei & Zhao, 2007), learning styles (Manochehr, 2006) and assessment (Wang et al., 2006; Pelgrum, 2001), obstacles to the integration of ICT in education (Pelgrum, 2001; Bingimlas, 2009; Yildirim, 2007; Wee & Zaitun, 2006), and subjects being learned (Cohen & Nycz, 2006; Williams, 2005), the field of educational technology lacks a widespread culture of usability, as reflected in the available research data.

Usability is one of the key factors for successful technology adoption. Research data have shown that the perceived usability of technological systems greatly determines the learning experience. However, adoption of usability studies in the field of educational technology was not observed until recently, despite the important role usability plays in the effectiveness of the educational technology systems (Orfanou et al., 2015).

A widely accepted definition of usability comes from the International Organization for Standardization (1998), which emphasizes three dimensions: effectiveness (the ability of users to complete their work qualitatively using the system), efficiency (resources spent on tasks), and satisfaction (subjective user reactions to system use) (ISO 9241-11). In addition, it has been shown that the three dimensions of usability (effectiveness, efficiency, and satisfaction) are weakly correlated (Frøkjaer et al., 2000). At first, the primary focus was on the two objective dimensions of effectiveness and efficiency, but this proved to be insufficient. Perceived usability (satisfaction) had to also be evaluated (Lewis, 2018).

Perceived usability has a significant impact on both the learning experience of the students and the learning outcome and thus on academic performance. If the user's interface is easy to use, the learners will use the system more often. If, on the contrary, the system is considered difficult to use, learners spend more time learning the system than the content (Ardito et al., 2006). In addition, a flawed interaction design for an educational technology could impede its adoption (Tselios et al., 2008). A usability evaluation is an assessment of the overall effectiveness of the system in the learning process (Thuseethan et al., 2014). Also, usability can result in improved learning experiences for learners (Tselios et al., 2008).

Thus, a need emerges for a standardized tool to assess the perceived usability of educational technology systems. Accordingly, each technological system is evaluated and classified. To evaluate perceived usability, Brooke (1996) developed the System Usability Scale (SUS), a reliable, zero-cost psychometric tool used worldwide with high validity and reliability. It comprises 10 alternating positive- and negative-formulated statements for which a respondent gives a subjective evaluation of a system's usability. The SUS's special attributes have rendered it an ideal tool for evaluating educational technology systems in the present research. First, the instrument is technology agnostic (Bangor et al., 2008; Sauro, 2013; Revythi & Tselios, 2019), which means that it can be used for evaluating any technological product, such as websites, mobile applications, and learning management systems. As Brooke (2013) points out, many of the current technologies did not even exist when he invented the SUS scale. Second, it is the most widely adopted tool for usability evaluation (Tullis & Albert, 2008). Third, it has proved to be reliable even with small sample size (Tullis & Stetson, 2004; Sauro, 2013).

Most of the questionnaire data are difficult to interpret unless there is a basis for comparison and reference. Moreover, usability can only be defined by reference to specific contexts and not as an absolute concept. Thus, one consequence of the specialized nature of usability is that it is very difficult to make comparisons of usability in different systems (Brooke, 1996). A few researchers (Bangor et al., 2008; Lewis & Sauro, 2009; Kortum & Bangor, 2013) have published some products' usability benchmarks for SUS. However, there is not a similar study for educational technology's usability, to the best of our knowledge. The research reported here attempts to provide a frame of reference of usability benchmarks among five major categories of educational technology. Thus, awareness is raised of the importance of usability in educational technology, and all stakeholders now have access to a useful, concentrated reference guide that takes into account many variables (e.g., age, subject being learned, educational stage) for developing or improving products.

Literature review

During the last two decades, research data has validated the reliability of the SUS scale. Consistent scale reliability and validity are confirmed by the literature, with Cronbach's alpha found repeatedly greater than 0.80 and in most of the research papers even above 0.90 (Chu et al., 2019; Bangor et al., 2008; Orfanou et al., 2015; Pal & Vanijja, 2020; Amariei, 2020; Kortum et al., 2020; Finstad, 2010).

Regarding the relation between SUS rating and participants' age, a slightly significant negative correlation is noticed in almost half of the research papers: Younger users tend to rate the usability of the SUS scale higher. Bangor et al. (2008) examined the possible associations between age and SUS score in a sample of 213 participants. They found a significant negative correlation between SUS score and age ($r = -0.203$, $p = 0.03$, s). Granić and Ćukušić (2011) found a significant negative correlation between SUS score and age ($r = -0.467$); ratings of younger participants are higher on the SUS scale than ratings of older ones. Harrati et al. (2016) evaluating the Moodle platform and García-Peñalvo et al. (2019) evaluating the WYRED platform, respectively, found that the younger age group in their studies gave the highest SUS ratings. For a dataset of 769 students, Orfanou et al. (2015) found a small, nonsignificant difference at the 0.05 level, with a negative correlation between SUS score and age ($r = -0.061$, $p = 0.09$). Binyamin et al. (2016) found no correlation between SUS score and age ($r = -0.223$, $p = 0.119$, ns). A recent study (Mujinga et al., 2018) revealed a statistically significant relationship ($p = 0.090$, $r = 0.036$, s) between age and SUS score, in which older users scored higher than younger users.

This article presents the findings of a systematic review of usability of educational technology systems. The research was conducted after studying, organizing, and analyzing the results of 104 research papers evaluating the perceived usability of various educational technologies at different settings using the System Usability Scale (SUS). The SUS scale is a psychometric tool that shows very high validity and reliability for usability evaluation (Orfanou et al., 2015). However, it seems that there is no universal consensus, as there are lots of questionnaires for usability evaluation, and it seems only that in recent years have these begun to be used for ICT in education. This literature review has been guided by the following research questions:

1. What is the level of usability of different types of technology used in education as derived from the System Usability Scale (SUS)?
2. Is there a significant difference in usability among these types of technology?
3. Is there a relationship between the age, subject being learned, level of education to which the technology is addressed, stakeholder, type of educational technology used, number of participants in each sample, and SUS score?
4. At what level of education is each evaluated type/category of educational technology used most?
5. Is there a significant difference in perceived usability for the same type of technological system over time?

Methodology

The goal of this research is twofold. First is to obtain and summarize results from previous research. Second is to build a framework and to provide a quantitative research-based benchmark as a frame of reference for usability of educational technology to all stakeholders (e.g., developers, administrators, educators). It is argued that such a study is significant since there is not any integrated systematic usability evaluation approach to contribute to design and development of the technological systems that are used in educational settings.

Moreover, this review intends to fill a large research gap by providing a "guidebook" to helping design and evaluate technological systems in education, aiming at resource management optimization. An added value is the categorization of educational technology into types, highlighting the different levels of usability among different types of technological systems. This research places emphasis on the need for more learner-centered educational technology systems by addressing the importance of their usability. Thus, it will have an impact on the improvement of both currently existing and future technological systems, as well as the selection of the most appropriate and user-friendly ones for each case.

First, the article presents the findings of a systematic review of usability evaluation of educational technology systems. The research was conducted after studying, organizing, and analyzing the results of 104 research papers evaluating the perceived usability of various educational technologies at different settings evaluated with the System Usability Scale (SUS). The methodology adopted follows the guidelines proposed by Kitchenham (2004).

The search described in this study was conducted using the following keywords at Google Scholar: SUS, usability, educational technology. The research papers used dated from 2006 to 2019. The collection of the data started on October 2019 and ended in January 2020. Where there were inadequate data, we contacted the researchers who provided the corresponding article. The following inclusion and exclusion criteria were adopted.

Inclusion criteria:

1. Research papers that evaluate usability for educational technology systems with SUS.
2. Each educational technology system should be a system used for educational purposes.

Exclusion criteria:

1. The research paper was not written in English.
2. The SUS scale was not adopted.
3. The SUS scale was not used to evaluate software for educational purposes.

The search procedure was conducted as follows: The keywords "SUS, usability, educational technology" were searched for at Google Scholar, with 3020 results. After eliminating the duplicates, 1899 results remained. Following this, the authors read the full research papers, or read the abstract if the full paper was not available. The research papers that met the eligibility criteria were aggregated into Microsoft Excel. One hundred and four research papers were retained that satisfied both the inclusion and exclusion criteria for further analysis. A second screening of the research papers was conducted for verification of the collected data. A few studies were excluded if there were not enough participants in some analyses .

The results of the literature review were organized in terms of (a) the usability score as derived from the SUS scale, (b) the type of educational technology used, (c) the number of participants in each sample, (d) the subject being learned, (e) the age, (f) the type of participants, and (g) the level of education to which the technology is addressed.

In the present study, our findings were initially grouped according to the type of educational technology used and evaluated on the SUS scale. The classification into generic categories is based

Table 1. Overview of the examined papers.

Internet platforms (LMS, MOOC, wiki, etc.)
Al-Omar (2018), Alqahtani (2019), Alshammari et al. (2016), Al-Sumaty and Umar (2018), Ayad and Rigas (2010), Binyamin et al. (2016), Blecken et al. (2010), Chaudy (2015), Christoph et al. (2017), de la Guía et al. (2012), Diwakar and Noronha (2018), Erdoğan et al. (2017), Erdoğmuş et al. (2015), García-Peñalvo et al. (2019), Granić and Ćukušić (2011), Gutiérrez-Carreón et al. (2015), Harrati et al. (2016), Harrati et al. (2017), Ivanović et al. (2018), Kaewsaiha (2019), Katsanos et al. (2012), Lehong et al. (2019), Luo et al. (2014), Marco et al. (2013), Orfanou et al. (2015), Pirker et al. (2019), Protopsaltis et al. (2013), Qaiduzzaman et al. (2018), Rahimi et al. (2015), Ras and Maquil (2011), Rizzardini et al. (2013), Rosato et al. (2007), Shi et al. (2013), Srimarong and Achalakul (2017), Thuseethan et al. (2014), Tsai and Yen (2013), Tsironis et al. (2016), Vertesi et al. (2018), Vincenti et al. (2017), Wesiak et al. (2015), Wu et al. (2009), Xenos et al. (2017)
Mobile applications
Arain et al. (2016), Armstrong and Wilkinson (2016), Botella et al. (2018), Davids et al. (2011), De Paolis et al. (2019), Escamilla et al. (2018), Ganapathy et al. (2016), Garcia-Ruiz et al. (2017), Hakala and Myllymäki (2014), Martín-Gutiérrez et al. (2015), Mathew (2012), Mirzaei (2016), Mustapa et al. (2018), Nicolaidou et al. (2019), Order (2015), Pombo and Marques (2018), Pombo and Marques (2019), Ponte et al. (2019), Pugoy et al. (2016), Spachos et al. (2014), Wang et al. (2008), Wang et al. (2010), Yağmur and Çakır (2016), Zbick et al. (2015)
University websites
Alnasser et al. (2017), Benaida et al. (2018), Demir and Parraci (2018), NaifJabli and Demir (2018), Şengel (2013), Sengel (2014)
Multimedia
Davids et al. (2014), Garcia-Ruiz et al. (2019), Hsieh and Lin (2006), Joshi et al. (2013), Kardong-Edgren et al. (2019), Lin et al. (2011), Lin (2018), Lin et al. (2012), Martin-Gonzalez et al. (2016), Odriozola et al. (2012), Schmidt et al. (2019), Sudarmilah and Siregar (2019), Tsai et al. (2018), Wismer et al. (2018)
Affective tutoring systems (ATS)
Feidakis et al. (2014), Lin et al. (2015), Lin et al. (2014), Lin et al. (2018), Lin et al. (2012), Lin et al. (2014), Ma (2017), Sedrakyan et al. (2017), Su et al. (2014)
Research papers reporting studies that do not fall into the above categories of educational technology systems
Alamer et al. (2015), Bahingawan et al. (2018), Barradas et al. (2019), Bureš et al. (2017), Christoph et al. (2017), Diehl et al. (2015), Estrada et al. (2019), Leow et al. (2016), Peters et al. (2019)

on previous researchers' work (Kortum & Bangor, 2013; Bangor et al., 2008; Sauro, 2011a). The types of educational technology fall into five categories: (a) Internet platforms, (b) mobile applications, (c) university websites, (d) multimedia, and (e) affective tutoring systems (ATS) (see Table 1). This classification is roughly based on Luo and Lei's work (Luo & Lei, 2012). Specifically, the Internet platforms category contains learning management systems (LMS), content management systems (CMS), learning content management systems (LCMS), massive open online courses (MOOCs), wikis, and cloud applications, and the multimedia category contains e-books, AR-VR books, and all kinds of visualization. Studies reported in 95 of the 104 collected papers fell into the five major categories of educational technology systems that occurred. The other nine papers were reporting studies related to other types of educational technology.

Regarding the learning discipline, five categories emerged, as follows: (a) natural sciences (physics, chemistry, biology), (b) foreign languages, (c) informatics, (d) skills, and (e) medicine. As far as the type of participant is concerned, five more categories occurred: (a) university students, (b) primary education students, (c) developers, (d) secondary education students, and (e) teachers.

A statistical analysis was conducted using IBM SPSS Statistics v25.0. Possible correlations were examined between various variables and SUS score regarding the research questions. We also took into consideration possible variations of the SUS score that may exist due to different sample sizes by providing weighted means.

Results

Perceived usability levels of educational technology systems

Statistical analysis in all surveys ($N = 170$) derived from the total of 104 research papers related to technology education systems revealed a good level of usability but with some issues ($M = 70.09$, $SD = 12.98$). This result is expected and is in line with Bangor et al. (2009) and Sauro

Table 2. Mean SUS score of types of educational technology.

Category	*N*	Mean SUS score	*SD*
Internet platforms (LMS, MOOC, wiki, etc.)	77	66.25	12.42
Mobile applications	33	73.62	13.49
University websites	12	63.82	16.52
Multimedia	21	76.43	9.45
Affective tutoring systems (ATS)	13	68.87	7.30

Table 3. Comparison across SUS scores for each category examined.

	Platforms	Mobile applications	University *Websites*	Multimedia
Platforms				
Mobile applications	0.006**			
University websites	0.549	0.049*		
Multimedia	0.001**	0.409	0.028*	
Affective tutoring systems (ATS)	0.463	0.238	0.345	0.019*

**Correlation significant at the 0.01 level (two-tailed).
*Correlation significant at the 0.05 level (two-tailed).

(2013). Possible variations in perceived usability scores due to the sample size of each survey were also considered. To that end, four surveys were excluded because they did not provide number of participants. Weighted SUS mean score ($M = 63.30$, $SD = 16.18$).

SUS benchmark data for types of educational technology

Analysis of the data collected revealed usability levels in each type of educational technology. According to Bangor et al. (2009), a SUS score above 51 is interpreted as "OK" with low marginal acceptability ranges, a SUS score above 72 is considered acceptable with "good" usability levels, and a SUS score above 85 corresponds to "excellent" usability levels.

One hundred fifty-six studies from 95 research papers fell into the five major categories of technology education systems. As it is shown in Table 2, the Internet platforms category comprised 77 surveys and its mean SUS score was 66.25 ($SD = 12.42$). The mobile applications category consisted of 33 surveys and its mean SUS score was 73.62 ($SD = 13.49$). The university websites category consisted of 12 surveys and its mean SUS score was 63.82 ($SD = 16.52$). The multimedia category comprised 21 surveys and its mean SUS score was 76.43 ($SD = 9.45$). The affective tutoring systems (ATS) category comprised 13 surveys and its mean SUS score was 68.87 ($SD = 7.30$).

In this context, the Internet platforms, university websites, and affective tutoring systems (ATS) categories seem to have good usability levels but not without flaws. In addition, the mobile applications and multimedia categories' scores indicate a satisfactory level of usability.

One-way analysis of variance (ANOVA) was conducted, and a statistically significant difference ($p = 0.002$) was found among the categories of educational technology.

As shown in Table 3, an independent samples *t*-test between pairs of the categories revealed statistically significant differences between the mean SUS scores of Internet platforms and mobile applications ($p = 0.006$, s), mobile applications and university websites ($p = 0.049$, s), university websites and multimedia ($p = 0.028$, s), multimedia and affective tutoring systems ($p = 0.019$, s), and Internet platforms and multimedia ($p = 0.001$, s).

The rest of the pairs did not reveal significant differences (see Table 3): platforms and university websites ($p = 0.549$, ns), platforms and affective tutoring systems ($p = 0.463$, ns), mobile

Table 4. Comparison across SUS scores for each subject being learned examined.

	Natural sciences	Foreign languages	Informatics	Skills	Medicine
Natural sciences					
Foreign languages	0.744				
Informatics	0.652	0.422			
Skills	0.532	0.751	0.267		
Medicine	0.016*	0.05*	0.001**	0.323	

**Correlation significant at the 0.01 level (two-tailed).
*Correlation significant at the 0.05 level (two-tailed).

applications and affective tutoring systems ($p = 0.238$, ns), university websites and affective tutoring systems ($p = 0.345$, ns), and mobile applications and multimedia ($p = 0.409$, ns).

The normality distribution was examined for the dependent variable SUS mean. Kolmogorov–Smirnov and Shapiro–Wilk tests revealed a normal distribution.

Age of participants and SUS score

In a dataset of 30 surveys, we investigated the possible correlation between the age of participants and the mean SUS score. No significant correlation was found ($r = 0.017$, $p = 0.931$, ns).

Subject being learned and SUS score

A one-way ANOVA in 67 surveys investigated possible correlation between the subject being learned (e.g., natural sciences, foreign languages, informatics, skills, and medicine) and the mean SUS score. A significant difference was observed ($p = 0.038$, s). As shown in Table 4, an independent samples t-test between pairs of the categories revealed statistically significant differences between the mean SUS scores of natural sciences and medicine ($p = 0.016$, s), foreign languages and medicine ($p = 0.05$, s), and informatics and medicine ($p = 0.001$, s).

The rest of the pairs did not reveal significant differences (see Table 4): natural sciences and foreign languages ($p = 0.744$, ns), natural sciences and informatics ($p = 0.652$, ns), natural sciences and skills ($p = 0.532$, ns), foreign languages and informatics ($p = 0.422$, ns), foreign languages and skills ($p = 0.751$, ns), informatics and skills ($p = 0.267$, ns), and skills and medicine ($p = 0.323$, ns).

Level of education and SUS score

A one-way ANOVA was adopted to investigate possible correlation between the level of education (i.e., primary, secondary, and higher education) and the mean SUS score; 112 surveys were included. No significant difference was observed ($p = 0.539$, ns).

Type of participants and SUS score

A one-way ANOVA investigated possible correlation between the type of participants (e.g., university students, teachers, etc.) and the mean SUS score; 124 surveys were included. No significant difference was observed ($p = 0.639$, ns).

Number of participants and SUS score

A negative correlation between number of participants and SUS score was revealed in a dataset of 166 surveys ($r = 0.259$, $p = 0.001$, s).

The results of the already-mentioned statistical analyses between SUS score and other variables, that is, age of participants, subject being learned, level of education, type of participants, and number of participants, are summarized in Table 5.

Table 5. SUS scores for each of the variables examined.

	Age of participants	Subject being learned	Level of education	Type of participants	Number of participants
SUS score	0.931	0.038*	0.539	0.639	0.001**

**Correlation significant at the 0.01 level (two-tailed).
*Correlation significant at the 0.05 level (two-tailed).

Table 6. Type of educational technology and educational stage.

	N	Primary education	Secondary education	Higher education
Internet platforms	36	2.8%	5.6%	91.7%
Mobile applications	16	12.5%	6.25%	81.25%
University websites	11			100%
Multimedia	10	40%		60%
Affective tutoring systems (ATS)	6	16.7%		83.3%

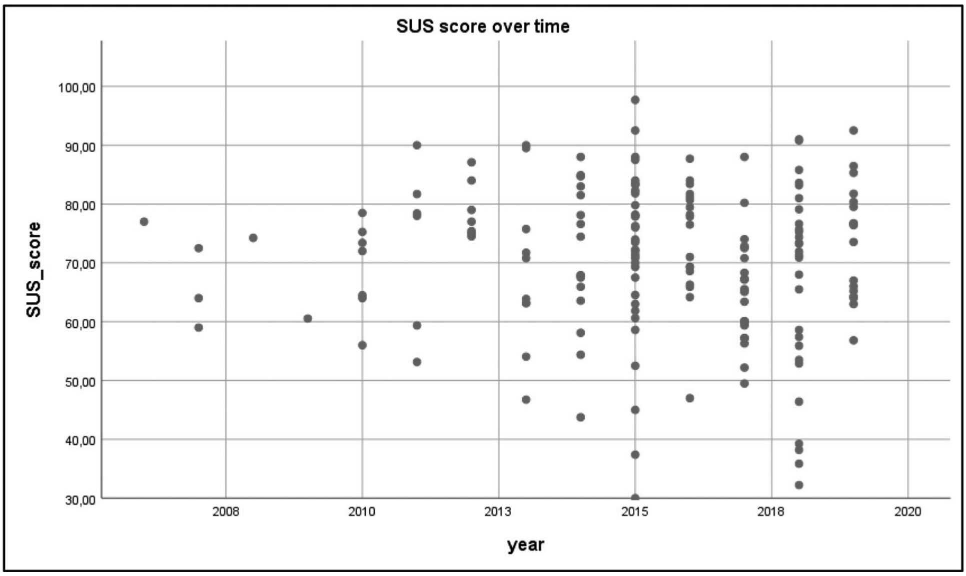

Figure 1. SUS score over time.

Type of educational technology and level of education

In a dataset of 79 surveys, we investigated the usage of each type of educational technology which was evaluated for the perceived usability at the three levels of education (see Table 6). Results show that 91.7% of Internet platforms are used in higher education, and 5.6% in secondary education; 81.25% of mobile applications are used in higher education and 12.5% in primary education. University websites, by definition, are only used in higher education settings. It was found that 60% of multimedia types are used in higher education and the rest for primary education, and that 83.3% of affective tutoring systems are used in higher education and 16.7% in primary education.

SUS scores over time

A one-way repeated-measures ANOVA was conducted to examine perceived usability evolution over time using as reference points the years 2013, 2015, 2017, and 2019 (see Figure 1). The test revealed a nonsignificant improvement in usability over time (Wilks' lambda = .520, $F(3, 7) = 2.154$, $p = 0.182$, ns).

Specifically, for Internet platforms Wilks' lambda = .320, $F(3, 3) = 2$, $p = 0.276$, ns; for mobile applications, Wilks' lambda = .073, $F(2, 1) = 6$, $p = 0.270$, ns; for university websites, Wilks' lambda = .040, $F(1, 1) = 24$, $p = 0.128$, ns; for multimedia systems, Wilks' lambda = .178, $F(2, 2) = 4$, $p = 0.178$, ns), and for affective tutoring systems, Wilks' lambda = .902, $F(1, 2) = 0$, $p = 0.688$, ns, so a nonsignificant improvement was observed for each over time.

Discussion

The aim of this systematic review is to provide a summary of the findings regarding the perceived usability's evaluation at educational technology systems using the System Usability Scale (SUS). This article intends to fill a major research gap by providing usability benchmarks for educational technology for all stakeholders. Usability is a rather overlooked factor when it comes to educational technology even though it has a significant impact on both the learning experience and the learning outcome. In other words, our research places usability at the center of educational technology design, user's experience and by extension the learning itself. The community also becomes aware of the usability levels. In addition, this study examines possible correlations of SUS score with gender, age, level of education, subject being learned, and type and number of participants. Finally, the usability of technology systems in education over time is investigated.

The competitive advantage of the SUS scale compared to other questionnaires is that it is technologically agnostic (Bangor et al., 2008; Sauro, 2013); this means it can be used for usability evaluation of any product or service without being limited to just one category (e.g., SUPR-Q is only for websites). Therefore, given that the same tool is used across different systems, it is possible to carry out meaningful comparisons.

The usability of the evaluated educational technology systems ($N = 170$) seems to be at a good level but not without flaws ($M = 70.09$, $SD = 12.98$). This finding is in line with prior research papers which state that an average SUS score is about 68 (Sauro, 2011b; Brooke, 2013; Bangor et al., 2009). The weighted mean is 63.30 ($SD = 16.18$).

Analysis of the results showed that Internet platforms ($M = 66.25$, $SD = 12.42$), university websites ($M = 63.82$, $SD = 16.52$), and affective tutoring systems (ATS) ($M = 68.87$, $SD = 7.30$) categories seem to have good usability levels with some issues. In addition, mobile applications ($M = 73.62$, $SD = 13.49$) and multimedia ($M = 76.43$, $SD = 9.45$) scores indicate a satisfactory level of usability. It would be expected that since Internet platforms were established earlier in the field of educational technology, they would have higher ratings on the SUS scale. However, in most cases they are used for educational purposes when users reach higher education. Therefore, usability levels of Internet platforms are below satisfactory levels. As far as the university websites are concerned, these do not directly aim at content learning and their role in the learning process is quite auxiliary. In addition, they are usually a product of student development, and their perceived usability may not be sufficiently taken into account in the design, as demonstrated by the research data.

Kaya et al. (2019) evaluated the perceived usability of widely adopted mobile applications (WhatsApp, Facebook, YouTube, and Mail) in a dataset of 222 participants with the System Usability Scale (SUS). The average score for iOS versions was 79.41 and for Android was 81.2. The corresponding category of mobile application for educational purposes has a mean SUS score of 73.80, well below these. Perhaps the differentiation in perceived usability between the commonly used applications and the applications used for educational purposes can be attributed to an extent to the discovery that repeat users with prior experience of a system tend to rate it up to 11% higher (Sauro, 2011a). This discovery constitutes a significant first indication; however, it is suggested that further studies are required to obtain meaningful conclusions and directions.

Additional analysis between age and the SUS score revealed a nonsignificant, positive correlation ($r = 0.017$, $p = 0.931$, ns). A recent research study (Mujinga et al., 2018) revealed a relation between age and SUS score ($r = 0.036$, $p = 0.090$, s). One possible explanation is that users are more familiar with technology systems, since they are an integral part of everyday life than in the past, closing the "gap" that was observed in previous years. We therefore conclude that the correlation between age and perceived usability yields contradictory results and requires further investigation in order to explore and identify the factors and variables that may act as moderators.

Furthermore, the subject being learned ($p = 0.038$) and number of participants ($r = -0.259$, $p = 0.001$, s) seem to correlate with the SUS score. Further examination regarding the subject being learned revealed that ICTs used in medicine-related subjects are considered more usable than those used in natural sciences ($p = 0.016$, s), foreign languages ($p = 0.05$, s), and informatics ($p = 0.001$, s). This could be attributed to a variety of reasons. The users in specific domains might be more experienced with ICT use in education, or the systems in one category might be designed in a more effective manner, thus fulfilling users' needs.

As far as the stage of education ($p = 0.539$, ns) and the type of participants ($p = 0.639$, ns) are concerned, no statistically significant correlations were found with the SUS score.

Moreover, categories of educational technology systems that were evaluated for usability were further examined to determine which educational levels they corresponded to. Results have shown that in higher education settings the educational technology is most usability-aware, followed at a distance by primary education. It is worth exploring further why the usability evaluation of educational technology systems is at such a low level in secondary education.

Finally, for all categories of educational technology systems we examined, a slight but statistically insignificant improvement is noticed in the perceived usability over the years ($p = 0.182$). Such a result would constitute an encouraging indication, implying a trend of awareness about the importance of perceived usability.

Conclusions

Although SUS was realized in 1996 for the evaluation of perceived usability, its adoption in the design of educational technology systems, took several years. In this review only two surveys, in the years 2008 and 2009, were identified that evaluate perceived usability in education using the SUS. In the following years, SUS adoption increased substantially, indicating the need for considering usability during design and development of educational technology systems.

Our goal was twofold: on the one hand, to provide a comprehensive overview of the current literature, and on the other hand, to explore possible relations between critical attributes with the findings obtained. A systematic review of usability in education was needed for summarizing all the existing information in a thorough and unbiased way. Consequently, this research will lead to more general and compact conclusions that individual studies cannot provide. Subsequently, because quantitative data were derived, it was possible to analyze and investigate possible correlations that individual research projects are not able to detect. To this end, findings were presented regarding SUS scores of the technology systems used in education by analyzing 104 research papers.

Levels of perceived usability among different types of educational technology are now acknowledged. Internet platforms, university websites, and affective tutoring systems seem to have a good usability level according to the SUS scale, with some issues. Mobile applications and multimedia have a satisfactory level of perceived usability according to the criteria set by Bangor et al. (2009).

Furthermore, differences in perceived usability among the categories of educational technology seem to be statistically significant ($p = 0.002$). Specifically, this applies to Internet platforms and mobile applications ($p = 0.06$), mobile applications and university websites ($p = 0.049$), university websites and multimedia ($p = 0.028$), multimedia and ATS ($p = 0.019$), and Internet platforms and multimedia ($p = 0.001$).

Moreover, concerning the attributes of the SUS, on the one hand, participants' age ($r = 0.017$, $p = 0.931$), level of education ($p = 0.539$), and type of participant ($p = 0.639$) are not significantly associated with the SUS score. On the other hand, subject being learned ($p = 0.038$) and number of participants ($r = -0.259$, $p = 0.001$) are statistically associated with the SUS score.

Regarding the fourth research question, most of the educational technologies that were evaluated for usability aim at higher education. Specifically, 91.7% of Internet platforms, 81.25% of mobile applications, 83.3% of ATS, and 60% of multimedia are used in higher education. As mentioned previously, university websites are used only in higher education. Finally, statistical analysis revealed a nonsignificant improvement in educational technology systems' perceived usability over time, for all five categories of the systems that were examined.

However, the present study is not without limitations. According to Sauro (2013), "SUS might not *always* be the best questionnaire." Although SUS is technology-agnostic, perhaps other instruments, such as SUPR-Q and the SEQ might be shown to be equally suitable. As a result, other questionnaires optimized for each educational technology category might be needed. In addition, researchers may want to measure more specific attributes such as findability or consistency. Such items do not appear in SUS. Thus, the development of appropriate items might be required (Sauro, 2018).

It was also hypothesized that the timing of the research coincides with the time of publication of the research paper. Usability evaluation of websites was conducted only by students, and 7 out of 12 of the studies reported in this article were carried out in 2018. Data for this study were obtained only from Google Scholar. Even though Google Scholar is sensitive enough to be used alone for systematic reviews (Gehanno et al., 2013), further research is recommended by including other databases as well, in the following years.

In the future, it would be interesting to investigate whether the scores reported by different cultures or nations are differentiated. In addition, further studies to investigate possible relations between participants' personality characteristics and SUS score are required. It is also suggested that other factors that may correlate with SUS score should be taken into consideration (e.g., gender, attitude toward technology, delivery method, cultural issues) and examined further.

Notes on contributors

Prokopia Vlachogianni is a PhD student in the Department of Educational Sciences and Early Childhood Education at the University of Patras in Greece. She holds a Master's degree in Natural Sciences, Mathematics and ICT didactics from the University of Patras, a Master's degree in Special Education from the University of Macedonia, and a Bachelor's degree in Primary Education from Aristotle University in Thessaloniki, Greece. Her main research interests are on Educational Technology, e-learning, Usability, Personality Traits and Learning Analytics.

Nikolaos Tselios is an Associate Professor in the Department of Educational Sciences and Early Childhood Education at the University of Patras in Greece. He holds a PhD (2002) in Usability Engineering of Educational Software and a Diploma (1997) from the Electrical and Computer Engineering Department, University of Patras, Greece. His main research interests are Educational Technology, Human Computer Interaction, user interface design and evaluation of educational software, usability evaluation methodologies, e-learning, user/student modelling and intelligent user interfaces. He has over 110 publications in international and national journals and conferences and 2 patents, with at least 2400 known citations and h-index = 24.

ORCID

Nikolaos Tselios http://orcid.org/0000-0002-4454-2499

References

Alamer, R. A., Al-Otaibi, H. M., Al-Khalifa, H. S. (2015, July). L3MS: A lightweight language learning management system using mobile web technologies. In *2015 IEEE 15th International Conference on Advanced Learning Technologies* (pp. 326–327). IEEE. https://doi.org/10.1109/icalt.2015.13

Alnasser, A., Alnabit, N., & Alanazi, W. (2017). Usability evaluation of prototypes designed for a Saudi university website. *International Journal of Computing & Information Sciences*, *13*(1), 15–25. https://doi.org/http://dx.doi.org/10.21700/ijcis.2017.xxx

Al-Omar, K. (2018, February). Evaluating the usability and learnability of the "blackboard" LMS using SUS and data mining. In *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 386–390). IEEE. https://doi.org/10.1109/iccmc.2018.8488038

Altanopoulou, P., Tselios, N., Katsanos, C., Georgoutsou, M., & Panagiotaki, M. A. (2015). Wiki-mediated activities in higher education: Evidence-based analysis of learning effectiveness across three studies. *Journal of Educational Technology & Society*, *18*(4), 511–522.

Alqahtani, A. (2019). Usability testing of Google Cloud Applications: Students' perspective. *Journal of Technology and Science Education*, *9*(3), 326–339. https://doi.org/10.3926/jotse.585

Alshammari, M., Anane, R., & Hendley, R. J. (2016, May). Usability and effectiveness evaluation of adaptivity in e-learning systems. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2984–2991). ACM. https://doi.org/10.1145/2851581.2892395

Al-Sumaty, R. M., & Umar, I. N. (2018, July). Design and evaluation of cloud-based students data management system usability. In *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)* (pp. 1–8). IEEE. https://doi.org/10.1109/icscee.2018.8538428

Amariei, A. L. (2020). *Usability assessment of medical training applications: exploring the dimensionality of the system usability scale* [Bachelor's thesis]. University of Twente.

Arain, A. A., Hussain, Z., Rizvi, W. H., & Vighio, M. S. (2016, July). Evaluating usability of M-learning application in the context of higher education institute. In *International conference on learning and collaboration technologies* (pp. 259–268). Springer. https://doi.org/10.1007/978-3-319-39483-1_24

Ardito, C., Costabile, M. F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T., & Rossano, V. (2006). An approach to usability evaluation of e-learning applications. *Universal Access in the Information Society*, *4*(3), 270–283. https://doi.org/10.1007/s10209-005-0008-6

Armstrong, P., & Wilkinson, B. (2016, November). Preliminary usability testing of ClaMApp: A classroom management app for tablets. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (pp. 654–656). ACM. https://doi.org/10.1145/3010915.3011855

Ayad, K., & Rigas, D. (2010). Comparing virtual classroom, game-based learning and storytelling teachings in e-learning. *International Journal of Education and Information Technologies*, *4*(1), 15–23.

Bahingawan, R. M., De los Reyes, L. M. O., & Llanes, T. R. II. (2018). Usability for gamified CAI (Computer Aided Instruction) chemistry adventure RPG (Role Playing Game) for grade nine chemistry teachers using adobe flash. [Bachelor's thesis, MSU - Iligan Institute of Technology].

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, *4*(3), 114–123. https://uxpajournal.org/wp-content/uploads/sites/8/pdf/JUS_Bangor_May2009.pdf

Barradas, R., Lencastre, J., Soares, S., & Valente, A. (2019). *Usability evaluation of an educational robot for STEM areas* [Paper presentation]. In 11th International Conference on Computer Supported Education, Heraklion, Crete, Greece. May 2–4. https://doi.org/10.5220/0007675102180225

Benaida, M., Namoun, A., & Taleb, A. (2018). Evaluation of the impact of usability in Arabic University Websites: Comparison between Saudi Arabia and the UK. *International Journal of Advanced Computer Science and Applications*, *9*(8), 365–375. https://doi.org/10.14569/IJACSA.2018.090848

Bingimlas, K. A. (2009). Barriers to the successful integration of ICT in teaching and learning environments: A review of the literature. *Eurasia Journal of Mathematics, Science & Technology Education*, *5*(3), 235–245. https://doi.org/10.12973/ejmste/75275

Binyamin, S., Rutter, M., & Smith, S. (2016, November 14–16). *The utilization of system usability scale in learning management systems: A case study of Jeddah Community College* [Paper presentation]. The 9th International Conference of Education, Research and Innovation (ICERI2016), International Academy of Technology, Education and Development, Seville, Spain. https://doi.org/10.21125/iceri.2016.2290

Blecken, A., Bruggemann, D., & Marx, W. (2010, January). Usability evaluation of a learning management system. In *2010 43rd Hawaii International Conference on System Sciences* (pp. 1–9). IEEE. https://doi.org/10.1109/hicss.2010.422

Botella, F., Peñalver, A., & Borrás, F. (2018, September). Evaluating the usability and acceptance of an AR app in learning Chemistry for Secondary Education. In *Proceedings of the XIX International Conference on Human Computer Interaction* (p. 31). ACM. https://doi.org/10.1145/3233824.3233838

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*, *189*(194), 4–7. https://doi.org/10.1201/9781498710411-35

Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies*, *8*(2), 29–40.

Bureš, V., Mikulecká, J., & Ponce, D. (2017). Digital television as a usable platform for enhancement of learning possibilities for the elderly. *SAGE Open*, *7*(2), 1–9. https://doi.org/10.1177/2158244017708817
Chaudy, Y. (2015, April). *Using an assessment engine for creating flexible educational games* [Paper presentation]. UWS Annual Research Conference, Paisley, UK.
Christoph, J., Knell, C., Bosserhoff, A., Naschberger, E., Stürzl, M., Rübner, M., Seuss, H., Ruh, M., Prokosch, H.-U., & Sedlmayr, B. (2017). Usability and suitability of the omics-integrating analysis platform tranSMART for translational research and education. *Applied Clinical Informatics*, *8*(4), 1173–1183. https://doi.org/10.4338/ACI-2017-05-RA-0085
Chu, A., Biancarelli, D., Drainoni, M. L., Liu, J. H., Schneider, J. I., Sullivan, R., & Sheng, A. Y. (2019). Usability of learning moment: Features of an E-learning tool that maximize adoption by students. *The Western Journal of Emergency Medicine*, *21*(1), 78–84. https://doi.org/10.5811/westjem.2019.6.42657
Cohen, E., & Nycz, M. (2006). Learning objects and e-learning: An informing science perspective. *Interdisciplinary Journal of E-Learning and Learning Objects*, *2*(1), 23–34. https://doi.org/10.28945/399
Davids, M. R., Chikte, U. M. E., & Halperin, M. L. (2011). Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. *Advances in Physiology Education*, *35*(3), 295–306. https://doi.org/10.1152/advan.00127.2010
Davids, M. R., Chikte, U. M., & Halperin, M. L. (2014). Effect of improving the usability of an e-learning resource: A randomized trial. *Advances in Physiology Education*, *38*(2), 155–160. https://doi.org/10.1152/advan.00119.2013
de la Guía, E., Lozano, M. D., Penichet, V. R. (2012, September). Co-BrainSystem: Supporting brainstorming to enhance collaborative work in educational environments. In *2012 Federated Conference on Computer Science and Information Systems (FedCSIS)* (pp. 849–855). IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6354388
De Paolis, L. T., De Luca, V., & Paladini, G. I. (2019, June). Touchless navigation in a multimedia application: The effects perceived in an educational context. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics* (pp. 348–367). Springer. https://doi.org/10.1007/978-3-030-25999-0_30
De Witte, K., & Rogge, N. (2014). Does ICT matter for effectiveness and efficiency in mathematics education? *Computers & Education*, *75*, 173–184. https://doi.org/10.1016/j.compedu.2014.02.012
Demir, F., & Parraci, W. (2018). The more complex the less success in online library services: Evaluating the user experience for international students. *Issues and Trends in Educational Technology*, *6*(2), 50–64. https://doi.org/10.2458/azu_itet_v6i2_demir
Diehl, L. A., De Souza, R. M., Gordan, P. A., Esteves, R. Z., & Coelho, I. C. M. (2015). User assessment of "insuOnLine," a game to fight clinical inertia in diabetes: A pilot study. *Games for Health Journal*, *4*(5), 335–343. https://doi.org/10.1089/g4h.2014.0111
Diwakar, A. S., Noronha, S. (2018, December). Usability and usefulness of ADVIcE tool experiment design guidelines for virtual laboratories. In *2018 IEEE Tenth International Conference on Technology for Education (T4E)* (pp. 146–149). IEEE. https://doi.org/10.1109/t4e.2018.00039
Dushayeva, S. J. (2019). The use and effectiveness of methods through "ICT" in English classroom. Вестник науки и образования, *2019*(3–2), 39–41.
Erdoğan, T., Yıldırım, O. G., & Çiğdem, H. (2017). The investigation of the usability of web-based assignment system. *Journal of Theory and Practice in Education*, *13*(1), 1–9.
Erdoğdu, F., Kokoç, M., Pinal, E., Bilgi, Ş., & Murat, Z. (2015). Investigation of an online learning environment in terms of usability. *Participatory Educational Research*, *3*(2), 55–66. https://doi.org/10.17275/per.15.34.2.3
Escamilla, E. F., Ostadalimakhmalbaf, M., Pariafsai, F., Ranka, N., Danesh, M., & Naderi Alizadeh, M. (2018). Impact of using iPad tablets in a construction communication graphics class: evaluation based on system usability scale. *Journal of Educational Technology Systems*, *47*(1), 32–49. https://doi.org/10.1177/0047239518773744
Estrada, G., Dawson, M., & Cárdenas-Haro, J. A. (2019). Investigating issues in computing education: Usability factors for the use of an operating system among African American and Hispanic American High School Students. *International Journal of Information and Communication Technologies in Education*, *8*(1), 5–19. https://doi.org/10.2478/ijicte-2019-0001
Feidakis, M., Caballé, S., Daradoumis, T., Jiménez, D. G., & Conesa, J. (2014). Providing emotion awareness and affective feedback to virtualised collaborative learning scenarios. *International Journal of Continuing Engineering Education and Life-Long Learning*, *24*(2), 141–167. https://doi.org/10.1504/IJCEELL.2014.060154
Finstad, K. (2010). The usability metric for user experience. *Interacting with Computers*, *22*(5), 323–327. https://doi.org/10.1016/j.intcom.2010.04.004
Frøkjaer, E., Hertzum, M., & Hornbaek, K. (2000, April). Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 345–352). ACM. https://doi.org/10.1145/332040.332455
Ganapathy, M., Shuib, M., & Azizan, S. N. (2016). Malaysian ESL students' perceptions on the usability of a mobile application for grammar test: A case study of ESL undergraduates in Universiti Sains Malaysia. *3L the Southeast Asian Journal of English Language Studies*, *22*(1), 127–140. https://doi.org/10.17576/3L-2016-2201-10

García-Peñalvo, F. J., Vázquez-Ingelmo, A., García-Holgado, A., & Seoane-Pardo, A. M. (2019). Analyzing the usability of the WYRED platform with undergraduate students to improve its features. *Universal Access in the Information Society*, *18*(3), 455–468. https://doi.org/10.1007/s10209-019-00672-z

Garcia-Ruiz, M. A., Santana-Mancilla, P. C., & Gaytan-Lugo, L. S. (2017, June). A usability study on low-cost virtual reality technology for visualizing digitized Canadian cultural objects: Implications in education. In *EdMedia + Innovate Learning* (pp. 259–264). Association for the Advancement of Computing in Education (AACE).

Garcia-Ruiz, M. A., Santana-Mancilla, P. C., & Gaytan-Lugo, L. S. (2019, June). A usability study of an interactive auditory display for supporting learning of molecular structure. In *EdMedia + Innovate Learning* (pp. 1400–1405). Association for the Advancement of Computing in Education (AACE).

Gehanno, J. F., Rollin, L., & Darmoni, S. (2013). Is the coverage of Google Scholar enough to be used alone for systematic reviews. *BMC Medical Informatics and Decision Making*, *13*(1), 7–5. https://doi.org/10.1186/1472-6947-13-7

Granić, A., & Ćukušić, M. (2011). Usability testing and expert inspections complemented by educational evaluation: A case study of an e-learning platform. *Journal of Educational Technology & Society*, *14*(2), 107–123.

Gutiérrez-Carreón, G., Daradoumis, T., & Jorba, J. (2015). Integrating learning services in the cloud: An approach that benefits both systems and learning. *Journal of Educational Technology & Society*, *18*(1), 145–157. https://www.ds.unipi.gr/et&s/journals/18_1/13.pdf

Hakala, I., & Myllymäki, M. (2014, December). Video sharing application for educational use: Usability and impacts of participation modes. In *2014 IEEE 12th IEEE International Conference on Emerging eLearning Technologies and Applications (ICETA)* (pp. 343–348). IEEE. https://doi.org/10.1109/iceta.2014.7107608

Harrati, N., Bouchrika, I., & Mahfouf, Z. (2017). Investigating the uptake of educational systems by academics using the technology to performance chain model. *Library Hi Tech*, *35*(4), 629–648. https://doi.org/10.1108/LHT-01-2017-0029

Harrati, N., Bouchrika, I., Tari, A., & Ladjailia, A. (2016). Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis. *Computers in Human Behavior*, *61*(2016), 463–471. https://doi.org/10.1016/j.chb.2016.03.051

Hsieh, M. C., Lin, H. C. K. (2006). Interaction design based on augmented reality technologies for English vocabulary learning. In *Proceedings of the 18th International Conference on Computers in Education* (Vol. 1, pp. 663–666). Asia-Pacific Society for Computers in Education. https://lexitron.nectec.or.th/public/ICCE%202010%20Putrajaya%20Malaysia/ICCE2010%20Main%20Proceedings/c5/Short%20paper/C5SP232.pdf

International Organization for Standardization. (1998). ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs): Part 11: Guidance on usability. https://doi.org/10.3403/01822507

Ivanović, M., Milicević, A. K., Ganzha, M., Bădică, A., Paprzycki, M., Bădică, C. (2018, January). Usability and quality parameters for e-learning environments and systems. In CEUR Workshop Proceedings. CEUR. http://ceur-ws.org/Vol-2217/paper-iva.pdf

Joshi, A., Wilhelm, S., Aguirre, T., Trout, K., & Amadi, C. (2013). An interactive, bilingual touch screen program to promote breastfeeding among Hispanic rural women: usability study. *JMIR Research Protocols*, *2*(2), e47. https://doi.org/10.2196/resprot.2872

Kaewsaiha, P. (2019). Usability of the learning management system and choices of alternative. In *The International Conference on Education, Psychology, and Social Sciences (ICEPS)* (pp. 252–259). Tokyo University of Science. http://www.elic.ssru.ac.th/pongrapee_ka/pluginfile.php/18/mod_page/content/12/Full%20Paper.pdf

Kardong-Edgren, S., Breitkreuz, K., Werb, M., Foreman, S., & Ellertson, A. (2019). Evaluating the usability of a second-generation virtual reality game for refreshing sterile urinary catheterization skills. *Nurse Educator*, *44*(3), 137–141. https://doi.org/10.1097/nne.0000000000000570

Katsanos, C., Tselios, N., & Xenos, M. (2012, October). Perceived usability evaluation of learning management systems: a first step towards standardization of the System Usability Scale in Greek. *In 2012 16th Panhellenic Conference on Informatics* (pp. 302–307). IEEE. https://doi.org/10.1109/pci.2012.38

Kaya, A., Ozturk, R., & Gumussoy, C. A. (2019). Usability measurement of mobile applications with system usability scale (SUS). In *Industrial Engineering in the Big Data Era* (pp. 389–400). Springer. https://doi.org/10.1007/978-3-030-03317-0_32

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele University 33*(2004), 1–26.

Kortum, P., Acemyan, C. Z., & Oswald, F. L. (2020). Is it time to go positive? Assessing the positively worded System Usability Scale (SUS). *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 001872081988155. https://doi.org/10.1177/0018720819881556

Kortum, P. T., & Bangor, A. (2013). Usability ratings for everyday products measured with the System Usability Scale. *International Journal of Human-Computer Interaction*, *29*(2), 67–76. https://doi.org/10.1080/10447318.2012.681221

Lehong, S., van Biljon, J., & Sanders, I. (2019, March). Open-distance electronic learning environments: Supervisors' views on usability. In *2019 Conference on Information Communications Technology and Society (ICTAS)* (pp. 1–7). IEEE. https://doi.org/10.1109/ictas.2019.8703605

Lei, J., & Zhao, Y. (2007). Technology uses and student achievement: A longitudinal study. *Computers & Education*, *49*(2), 284–296. https://doi.org/10.1016/j.compedu.2005.06.013

Leow, M. C., Wang, L. Y. K., Lau, S. H., & Tan, C. K. (2016). Usability of RPG-based Learning Framework. *International Journal of Human-Computer Interaction*, *32*(8), 643–653. https://doi.org/10.1080/10447318.2016.1183863

Lewis, J. R., & Sauro, J. (2009, July). The factor structure of the system usability scale. In *International conference on human centered design* (pp. 94–103). Springer. https://doi.org/10.1007/978-3-642-02806-9_12

Lewis, J. R. (2018). The system usability scale: Past, present, and future. *International Journal of Human–Computer Interaction*, *34*(7), 577–590. https://doi.org/10.1080/10447318.2018.1455307

Lin, C. (2018, August 27–30). Usability evaluation of the game based e-book system on natural science teaching system. In *Innovative Technologies and Learning: First International Conference (ICITL 2018), Portoroz, Slovenia, Proceedings* (Vol. 11003, p. 463). Springer. https://doi.org/10.1007/978-3-319-99737-7_49

Lin, H. C. K., Chao, C. J., & Huang, T. C. (2015). From a perspective on foreign language learning anxiety to develop an affective tutoring system. *Educational Technology Research and Development*, *63*(5), 727–747. https://doi.org/10.1007/s11423-015-9385-6

Lin, H. C. K., Chen, N. S., Sun, R. T., & Tsai, I. H. (2014). Usability of affective interfaces for a digital arts tutoring system. *Behaviour and Information Technology*, *33*(2), 104–115. https://doi.org/10.1080/0144929x.2012.702356

Lin, H. C. K., Hsieh, M. C., Liu, E. Z. F., & Chuang, T. Y. (2012). Interacting with visual poems through AR-based digital artwork. *Turkish Online Journal of Educational Technology-TOJET*, *11*(1), 123–137.

Lin, H. C. K., Hsieh, M. C., Wang, C. H., Sie, Z. Y., & Chang, S. H. (2011). Establishment and usability evaluation of an interactive ar learning system on conservation of fish. *Turkish Online Journal of Educational Technology-TOJET*, *10*(4), 181–187.

Lin, H. C. K., Hsu, W. C., Wang, T. H., Ma, Y. C., & Tsai, M. C. (2018). Development and research of an affective learning system combined with motion-sensing interaction, augmented reality, and mid-air projection. *Journal of Internet Technology*, *19*(6), 1951–1960.

Lin, H. C. K., Wang, C. H., Chao, C. J., & Chien, M. K. (2012). Employing textual and facial emotion recognition to design an affective tutoring system. *Turkish Online Journal of Educational Technology-TOJET*, *11*(4), 418–426.

Lin, H. C. K., Wu, C. H., & Hsueh, Y. P. (2014). The influence of using affective tutoring system in accounting remedial instruction on learning performance and usability. *Computers in Human Behavior*, *41*, 514–522. https://doi.org/10.1016/j.chb.2014.09.052

Livingstone, S. (2012). Critical reflections on the benefits of ICT in education. *Oxford Review of Education*, *38*(1), 9–24. https://doi.org/10.1080/03054985.2011.577938

Luo, H., & Lei, J. (2012). Emerging technologies for interactive learning in the ICT age. In *Educational Stages and Interactive Learning: From Kindergarten to Workplace Training* (pp. 73–91. IGI Global. https://doi.org/10.4018/978-1-4666-0137-6.ch005

Luo, G. H., Liu, E. Z. F., Kuo, H. W., & Yuan, S. M. (2014). Design and implementation of a simulation-based learning system for international trade. *The International Review of Research in Open and Distributed Learning*, *15*(1), 203–226. https://doi.org/10.19173/irrodl.v15i1.1666

Ma, Y. C. (2017, September). The Development of an Affective Tutoring System for Japanese Language Learners. In *International Symposium on Emerging Technologies for Education* (pp. 363–371). Springer. https://doi.org/10.1007/978-3-319-71084-6_41

Manochehr, N. N. (2006). The influence of learning styles on learners in e-learning environments: An empirical study. *Computers in Higher Education Economics Review*, *18*(1), 10–14.

Marco, F. A., Penichet, V. M. R., & Gallud, J. A. (2013). Collaborative e-learning through drag & share in synchronous shared workspaces. *Journal of Universal Computer Science*, *19*(7), 894–911.

Martin-Gonzalez, A., Chi-Poot, A., & Uc-Cetina, V. (2016). Usability evaluation of an augmented reality system for teaching Euclidean vectors. *Innovations in Education and Teaching International*, *53*(6), 627–636. https://doi.org/10.1080/14703297.2015.1108856

Martín-Gutiérrez, J., Fabiani, P., Benesova, W., Meneses, M. D., & Mora, C. E. (2015). Augmented reality to promote collaborative and autonomous learning in higher education. *Computers in Human Behavior*, *51*, 752–761. https://doi.org/10.1016/j.chb.2014.11.093

Mathew, D. A. (2012). *A mobile tablet app for clinical evaluation and medical education: Development and usability evaluation* [Doctoral dissertation]. McMaster University.

Mayes, J. T., & Fowler, C. J. (1999). Learning technology and usability: A framework for understanding courseware. *Interacting with Computers*, *11*(5), 485–497. https://doi.org/10.1016/S0953-5438(98)00065-4

Mirzaei, S. (2016). *Evaluating efficacy and usability of mobile devices for learning new vocabulary items* [Doctoral dissertation]. Flinders University, School of Computer Science, Engineering and Mathematics.

Mujinga, M., Eloff, M. M., & Kroeze, J. H. (2018). System usability scale evaluation of online banking services: A South African study. *South African Journal of Science*, *114*(3/4), 1–8. https://doi.org/10.17159/sajs.2018/20170065

Mustapa, A. M., Nawawi, Z., Ab Ghani, S., Rahman, M. A., Shaadon, Z., & Mustapa, N. S. (2018). *Usability of QiraahBot for extensive Arabic reading activities*. In *SHS Web of Conferences* (Vol. 53). EDP Sciences. https://doi.org/10.1051/shsconf/20185304005

NaifJabli, H., & Demir, F. (2018). The usability of King Khalid University website: Assessing effectiveness, efficiency, and satisfaction. *International Journal of Arts Humanities and Social Sciences*, *3*(7), 10–16.

Nicolaidou, I., Tozzi, F., Kindynis, P., Panayiotou, M., & Antoniades, A. (2019). Development and usability of a gamified app to help children manage stress: An evaluation study. *Italian Journal of Educational Technology*, *27*(2), 105–120.

Odriozola, S., Luis, J., Verbert, K., Duval, E. (2012). Empowering students to reflect on their activity with StepUp!: Two case studies with engineering students. In *Proceedings of ARETL'12 2nd Workshop on Awareness and Reflection* (Vol. 931, pp. 73–86). CEUR Workshop Proceedings.

Order, S. (2015). ICreate': Preliminary usability testing of apps for the music technology classroom. *Journal of University Teaching & Learning Practice*, *12*(4), 8.

Orfanou, K., Tselios, N., & Katsanos, C. (2015). Perceived usability evaluation of learning management systems: Empirical evaluation of the System Usability Scale. *The International Review of Research in Open and Distributed Learning*, *16*(2), 227–246. https://doi.org/10.19173/irrodl.v16i2.1955

Pal, D., & Vanijja, V. (2020). Perceived usability evaluation of Microsoft teams as an online learning platform during COVID-19 using system usability scale and technology acceptance model in India. *Children and Youth Services Review*, *119*, 105535. https://doi.org/10.1016/j.childyouth.2020.105535

Pelgrum, W. J. (2001). Obstacles to the integration of ICT in education: Results from a worldwide educational assessment. *Computers & Education*, *37*(2), 163–178. https://doi.org/10.1016/S0360-1315(01)00045-8

Peters, M., Pechuel, R., Huelsken-Giesler, M., Dütthorn, N., & Hoffmann, B. (2019, March). When learning tools need to handle more than right and wrong addressing the challenges of creating a serious game in the field of nursing education. In *Society for Information Technology & Teacher Education International Conference* (pp. 507–515). Association for the Advancement of Computing in Education (AACE).

Petko, D., Cantieni, A., & Prasse, D. (2017). Perceived quality of educational technology matters: A secondary analysis of students' ICT use, ICT-related attitudes, and PISA 2012 test scores. *Journal of Educational Computing Research*, *54*(8), 1070–1091. https://doi.org/10.1177/0735633116649373

Piper, B., Jepkemei, E., Kwayumba, D., & Kibukho, K. (2015). Kenya's ICT policy in practice: The effectiveness of tablets and e-readers in improving student outcomes. *FIRE: Forum for International Research in Education*, *2*(1), 1–17. https://doi.org/10.18275/fire201502011025

Pirker, J., Holly, M., Almer, H., Gütl, C., & Belcher, J. W. (2019). Virtual reality STEM education from a teacher's perspective. In *iLRN 2019 London, Workshop, Long and Short Paper, and Poster Proceedings: from the Fifth Immersive*. Verlag der Technischen Universität Graz.

Pombo, L., & Marques, M. M. (2018). The EduPARK mobile augmented reality game: Learning value and usability. In *14th International Conference Mobile Learning* (pp. 23–30). IADIS.

Pombo, L., & Marques, M. M. (2019, July). Learning with the augmented reality EduPARK game-like app: Its usability and educational value for primary education. In *Intelligent Computing-Proceedings of the Computing Conference* (pp. 113–125). Springer. https://doi.org/10.1007/978-3-030-22871-2_9

Ponte, R. P., Sanders, L. L. O., Júnior, A. P., Kubrusly, M., & Marçal, E. (2019). Development and usability assessment of a mobile application for neuroanatomy teaching: A case study in Brazil. *Creative Education*, *10*(3), 600–609. https://doi.org/10.4236/ce.2019.103043

Protopsaltis, A., Hainey, T., Borosis, S., Connolly, T., Copado, J., & Hezner, S. (2013). Startup_EU: Using game-based learning and web 2.0 technologies to teach entrepreneurship to secondary education students. *In 7th European Conference on Games Based Learning* (Vol. *1–2*, pp. 484–494).

Pugoy, R. A. D., Ramos, R. C., Figueroa, R. B., Jr., Rivera, M. H. C., Siritarungsri, B., Cheevakasemsook, A., Noimuenwai, P., & Kaewsarn, P. (2016). Augmented reality in nursing education: addressing the limitations of developing a learning material for nurses in the Philippines and Thailand. *International Journal of Open Distance e-Learning*, *2*(1), 11–24.

Qaiduzzaman, K. M., Shahjahan, M., Sobhan, S., Arman, M. S., Taj Noor, M. B., & Rahman, M. (2018). An effective attendance monitoring system with fraud prevention technique for educational institutions. *International Journal of Engineering & Technology*, *7*(3), 1593–1598. https://doi.org/10.14419/ijet.v7i3.13974

Rahimi, A., Embi, M. A., & Rahimi, A. (2015). Evaluation of the e-Learning developed for casemix and clinical coding: Quality of the material and usability of the system. *Argos Special Issue*, *2*, 130–143.

Ras, E., & Maquil, V. (2011). Preliminary results of a usability study in the domain of technology-based assessment using a tangible tabletop. In Workshop Proceedings of IHM (pp. 3–7). http://valeriemaquil.eu/publications/ras11_assessment_tabletop.pdf

Revythi, A., & Tselios, N. (2019). Extension of technology acceptance model by using system usability scale to assess behavioral intention to use e-learning. *Education and Information Technologies*, *24*(4), 2341–2355. https://doi.org/10.1007/s10639-019-09869-4

Rizzardini, R. H., Amado-Salvatierra, H. R., & Guetl, C. (2013). Cloud-based learning environments: Investigating learning activities experiences from motivation, usability and emotional perspective. In *Proceedings of the 5th International Conference on Computer Supported Education - Volume 1: WCLOUD, (CSEDU 2013)* (pp. 709–716). SciTePress. https://doi.org/10.5220/0004451807090716

Rosato, J., Dodds, C., Laughlin, S. (2007). Usability of course management systems by students. Dept. Computer Information Systems/Computer Science, College of Scholastica, Duluth. https://www.researchgate.net/profile/Jennifer_Rosato/publication/267713654_Usability_of_Course_Management_Systems_by_Students/links/551a86b00cf2f51a6fea5b71/Usability-of-Course-Management-Systems-by-Students.pdf

Safar, A. H., Al-Jafar, A. A., & Al-Yousefi, Z. H. (2016). The effectiveness of using augmented reality apps in teaching the English alphabet to kindergarten children: A case study in the State of Kuwait. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(2), 417–440. https://doi.org/10.12973/eurasia.2017.00624a

Sauro, J. (2011a). *Does prior experience affect perceptions of usability?* https://measuringu.com/prior-exposure/

Sauro, J. (2011b). *Measuring usability with System Usability Scale (SUS).* https://measuringu.com/sus/

Sauro, J. (2013). *10 things to know about System Usability Scale (SUS).* https://measuringu.com/10-things-sus/

Sauro, J. (2018). *Interpreting single items from the SUS.* https://measuringu.com/sus-items/

Schacter, J. (1999). *The impact of education technology on student achievement: What the most current research has to say*. Milken Exchange on Educational Technology.

Schmidt, F., Ohlemacher, J., Hennig, V., Kao, O., & Nordholz, J. (2019). *Case study: Visualizing computer system programming concepts for education* [Paper presentation]. In *47th SEFI Annual Conference*, Budapest.

Sedrakyan, G., Leony, D., Muñoz-Merino, P. J., Kloos, C. D., & Verbert, K. (2017, September). Evaluating student-facing learning dashboards of affective states. In *European Conference on Technology Enhanced Learning* (pp. 224–237). Springer. https://doi.org/10.1007/978-3-319-66610-5_17

Şengel, E. (2013). Usability level of a university web site. *Procedia - Social and Behavioral Sciences*, *106*, 3246–3252. https://doi.org/10.1016/j.sbspro.2013.12.373

Sengel, E. (2014). Discovering how students search a university web site: A comparative usability case study for PC and mobile devices. *Turkish Online Journal of Educational Technology-TOJET*, *13*(4), 12–20.

Shi, L., Awan, M. S. K., & Cristea, A. I. (2013, September). Evaluating system functionality in social personalized adaptive e-learning systems. In *European Conference on Technology Enhanced Learning* (pp. 633–634). Springer. https://doi.org/10.1007/978-3-642-40814-4_87

Spachos, D., Hatzichristou, D., & Bamidis, P. (2014, November). Using mobile applications in continuing medical education. In *2014 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL2014)* (pp. 301–304). IEEE. https://doi.org/10.1109/imctl.2014.7011152

Srimarong, S., & Achalakul, T. (2017, November). Usability evaluation of outcome-based education tool. In *IEEE 9th International Conference on Engineering Education (ICEED)* (pp. 233–237). IEEE. https://doi.org/10.1109/ICEED.2017.8251199

Su, H., Hsieh, Y. C., & Tsai, S. C. (2014). Impacts of affective tutoring system on the academic achievement of primary school students with different cognitive styles–An example of marine education. *Stanisław Juszczyk*, *38*(4), 250–231.

Sudarmilah, E., & Siregar, R. (2019). The usability of "keepin" collect the trash: Virtual reality educational game in android smartphone for children. *International Journal of Engineering and Advanced Technology (IJEAT)*, *8*(4), 944–947.

Thuseethan, S., Achchuthan, S., Kuhanesan, S. (2014). Usability evaluation of learning management systems in Sri Lankan universities. arXiv preprint arXiv:1412.0197. https://arxiv.org/pdf/1412.0197

Tsai, C. H., & Yen, J. C. (2013). The development and evaluation of a Kinect sensor assisted learning system on the spatial visualization skills. *Procedia - Social and Behavioral Sciences*, *103*, 991–998. https://doi.org/10.1016/j.sbspro.2013.10.423

Tsai, M. C., Lin, H. C. K., & Lin, C. (2018). Usability evaluation of the game based e-book system on natural science teaching system. In *International Conference on Innovative Technologies and Learning* (pp. 463–472). Springer. https://doi.org/10.1007/978-3-319-99737-7_49

Tsironis, A., Katsanos, C., Xenos, M., (2016). Comparative usability evaluation of three popular MOOC platforms. In *2016 IEEE Global Engineering Education Conference (EDUCON)* (pp. 608–612). IEEE. https://doi.org/10.1109/educon.2016.7474613

Tselios, N., Avouris, N., & Komis, V. (2008). The effective combination of hybrid usability methods in evaluating educational applications of ICT: Issues and challenges. *Education and Information Technologies*, 13(1), 55–76. https://doi.org/10.1007/s10639-007-9045-5

Tullis, T., & Albert, B. (2008). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics.* Morgan Kaufmann.

Tullis, T. S., & Stetson, J. N. (2004). A comparison of questionnaires for assessing website usability. In Usability Professional Association Conference (Vol. *1*). https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.396.3677&rep=rep1&type=pdf

Vertesi, A., Dogan, H., Stefanidis, A., Ashton, G., Drake, W. (2018). *Usability evaluation of a virtual learning environment: A university case study* [Paper presentation]. 15th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA), October 21–23, Budapest, Hungary. https://doi.org/10.1007/978-3-030-48190-2_9

Vincenti, G., Hilberg, S., Braman, J., Satzinger, M., Cao, L. (2017). Assessing the usability of a novel system for programming education. arXiv preprint arXiv:1711.05649. https://arxiv.org/pdf/1711.05649

Wang, K. H., Wang, T. H., Wang, W. L., & Huang, S. C. (2006). Learning styles and formative assessment strategy: Enhancing student achievement in Web-based learning. *Journal of Computer Assisted Learning*, *22*(3), 207–217. https://doi.org/10.1111/j.1365-2729.2006.00166.x

Wang, A. I., Øfsdahl, T., & Mørch-Storstein, O. K. (2008, April). An evaluation of a mobile game concept for lectures. In *2008 21st Conference on Software Engineering Education and Training* (pp. 197–204). IEEE. https://doi.org/10.1109/cseet.2008.15

Wang, A. I., Wu, B., & Bakken, S. K. (2010, December). Experiences from implementing a face-to-face educational game for iPhone/iPod Touch. In *2010 2nd International IEEE Consumer Electronics Society's Games Innovations Conference* (pp. 1–8). IEEE. https://doi.org/10.1109/icegic.2010.5716895

Wee, M. C., & Zaitun, A. B. (2006). Obstacles towards the use of ICT tools in teaching and learning of information systems in Malaysian Universities. *The International Arab Journal of Information Technology*, *3*(3), 203–209.

Wesiak, G., Höfler, M., Al-Smadi, M., & Gütl, C. (2015). CSCL in non-technological environments: Evaluation of a wiki system with integrated self-and peer assessment. In *PACT 2015: International Psychological Applications Conference and Trends* (pp. 107–111). InPACT.

Williams, P. (2005). Lessons from the future: ICT scenarios and the education of teachers. *Journal of Education for Teaching*, *31*(4), 319–339. https://doi.org/10.1080/02607470500280209

Wismer, A., Reinerman-Jones, L., Teo, G., Willis, S., McCracken, K., & Hackett, M. (2018). Assessing performance and usability of 3D visualization technologies for anatomical training. *MODSIM World*, *2018*(17), 1–11.

Wu, B., Strom, J. E., Wang, A. I., & Kvamme, T. B. (2009, August). XQUEST used in software architecture education. In *2009 International IEEE Consumer Electronics Society's Games Innovations Conference* (pp. 70–77). IEEE. https://doi.org/10.1109/icegic.2009.5293607

Xenos, M., Maratou, V., Ntokas, I., Mettouris, C., & Papadopoulos, G. A. (2017, April). Game-based learning using a 3D virtual world in computer engineering education. In *2017 IEEE Global Engineering Education Conference (EDUCON)*) (pp. 1078–1083). IEEE. https://doi.org/10.1109/EDUCON.2017.7942982

Yağmur, S., & Çakır, M. P. (2016). Usability evaluation of a dynamic geometry software mobile interface through eye tracking. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9753*, 391–402. https://doi.org/10.1007/978-3-319-39483-1_36

Yildirim, S. (2007). Current utilization of ICT in Turkish basic education schools: A review of teacher's ICT use and barriers to integration. *International Journal of Instructional Media*, *34*(2), 171–186.

Youssef, A. B., & Dahmani, M. (2008). The impact of ICT on student performance in higher education: Direct effects, indirect effects and organisational change. *RUSC. Universities and Knowledge Society Journal*, *5*(1), 13. https://doi.org/10.7238/rusc.v5i1.321

Zbick, J., Nake, I., Milrad, M., Jansen, M. (2015, July). A web-based framework to design and deploy mobile learning activities: Evaluating its usability, learnability and acceptance. In *2015 IEEE 15th International Conference on Advanced Learning Technologies* (pp. 88–92). IEEE. https://doi.org/10.1109/icalt.2015.97