Mersin Üniversitesi Eğitim Fakültesi Dergisi - Mersin University Journal of the Faculty of Education

"Mersin Üniversitesi Eğitim Fakültesi Dergisi, 2015; 11(3): 752-763"
DOI: 10.17860/efd.19683

Investigation of Usage Frequency and Department on Usability Issues with the Online Examination System (OSSI)*

Kullanım Sıklığı ve Bölümün Online Sınav Sistemi'nin (OSSİ) Kullanılabilirlik Sorunlarına Etkisinin İncelenmesi

Ali BATTAL**, Kürşat ÇAĞILTAY***

Abstract: Usability is an important term for systems related to learning and teaching because usability issues can negatively affect student achievement. The aim of this study is to investigate the usability of a specific online examination system, OSSI, which is used by a public university for midterm and final exams. Researchers examined mean differences between System Usability Scale (SUS) scores of students from different departments with different usage frequencies. Usability issues were investigated by gathering students' ideas. Both quantitative and qualitative data procedures were employed. Quantitative data were analyzed through analysis of variance, while content analysis was applied to qualitative data. Scores showed no statistical mean difference between usage frequencies but did show a statistically significant mean difference between CEIT and other two departments. Qualitative analysis revealed the importance of screen size, placement of buttons, question formatting, and other screen elements. The results of this study contribute to the design and development of OSSI and other online examination systems.
Keywords: E-assessment, online examination systems, usability

Öz: Öğrenme ve öğretim sürecinde kullanılan sistemlerdeki sorunların öğrenci başarısını olumsuz etkileyebileceğinden dolayı bu sistemlerin kullanılabilirlik yönünden incelenmesi önemlidir. Bu çalışmanının amacı bir devlet üniversitesinde vize ve final sınavlarında kullanılan çevrimiçi sinav sistemi, OSSİ'nin kullanılabilirliğini incelemektedir. Bu amaçla farklı bölümlerden ve sistemi farklı sıklıklarda kullanan öğrencilerin Sistem Kullanılabililik Ölçeği ile elde edilen puanları arasında anlamlı ortalama farkı araştırmacılar tarafından incelenmiştir. Buna ek olarak öğrencilerin sistem hakkındaki görüşleri alınarak sistem kullanılabilirliği ile ilgili diğer hususlar da açığa çıkartılmıştır. Hem nicel hem de nitel veri bu çalışma için toplanmıştır. Nicel veri için varyans analizi yapılırken nitel veri içerik analizi uygulanarak analiz edilmiştir. Çalışma sonunda kullanım sıklıklarında anlamlı ortalama farkı bulunmazken bölümler arasında BÖTE öğrencileri ile diğer iki bölümün öğrencileri arasında anlamlı ortalama farkı bulunmuştur. Ayrıca nitel analiz sonucu ekran büyüklünün önemi, düğmelerin ekranda yerleşimi, soruların formatı ve diğer ekran ile ilgili hususların önemi ortaya çıkmıştır. Bu çalışmanın sonuçları hem OSSİ'nin hem de diğer sistemlerin tasarım ve gelişimine katkıda bulunacaktır.
Anahtar Kelimeler: Çevrimiçi değerlendirme, çevrimiçi sınav sistemleri, kullanılabilirlik

Introduction

Rapid advances in technology have affected the lives of people in many aspects, particularly with regard to education. Traditional assessment and evaluation processes are shifting to online formats more and more, and many institutions around the world are also utilizing online examination systems (Gülbahar, 2013). According to Bennett (2003), dozens of US states have transferred their traditional assessment systems online due to the distinct advantages. de Siqueira, Martínez-Sáez, Sevilla-Pavón, and Gimeno-Sanz (2011) created a web-based assessment system for foreign language exams, PAULEX, and their pilot study indicated a high

*Part of this study was orally presented at the 2014 AECT International Convention in Jacksonville, Florida.

**Res. Asst., Middle East Technical University, Faculty of Education, Ankara, Turkey, email: albattal@metu.edu.tr

***Prof. Dr., Middle East Technical University, Faculty of Education, Ankara, Turkey, email: kursat@metu.edu.tr

Gönderi tarihi:23.02.2015 - Kabul tarihi: 19.10.2015

degree of acceptance by users. Dermo (2009) used another system and found that students appreciated the use of online assessment and agreed that it contributed positively to the learning and teaching process (as cited in Gülbahar, 2013).

One public institute in Turkey, OSYM (Measuring, Selection and Placement Center), aims to offer online assessment and evaluation in the near future in addition to traditional formats (OSYM, 2013). MEB (the Turkish Ministry of Education) began offering online assessment options for driving licenses about two years ago (MEB, 2013). Other web-based online examination systems have been developed for specific projects (Karahoca, Karahoca, & Günoğlu, 2009; Karakaya, 2002; Torkul, Kibar, & Taşcı, 2004; Tüfekçi, Ekinci, & Köse, 2013).

OSSI (Online Sinav Sistemi) is an online examination system developed and used by a public university in Turkey. It has been used for almost six years by students to take quizzes, midterms, and final examinations for courses such as Computer I-II through a web-based platform. OSSI provides opportunities for instructors to prepare multiple choice, open-ended, and true/false questions, enables them to receive files from students and enables instructors to evaluate student performance. Tümer, Şahin, and Aktürk (2008) conducted a study related to this system and found a significant difference between male and female students' ideas about certain aspects. However, that study did not undertake an examination of the whole system in terms of usability.

The aim of the current study is to analyze whether students' OSSI usability scores are affected by their department or system usage frequency. Results will reveal not only necessary advancements for OSSI but also the advantages of online examination systems that should be accommodated better by this system. Such advancements will make the system more usable, meaning that it can be used easily and effectively by students to fulfill their needs at home and school (Shackel, 1991).

Literature review

Online examination systems

Before defining online examination system, there is a need to clarify the meaning of assessment. Gikandi, Morrow, and Davis (2011) defined it as a kind of educational activity in which a learner's achievement and progress are measured; they stressed the importance of assessment for formal higher education. Online examination systems assess via web-based platforms and have specific purposes. They can be defined as an online platform in which all traditional tasks are conducted on the Internet, such as preparing and taking exams or announcing the results (Tüfekçi et al., 2013). Online exam systems should address the following issues:

- The exam system should be accessible on common web browsers.
- Questions should be presented in a non-distracting way.
- The exam system should be easy to use, even for users with low level ICT skills.
- The system should support different kinds of report to student and instructor.

Karahoca et al. (2009) have identified many online examination systems in the literature that offer a range of operational functions affecting efficacy, speed, and usability, such as question flexibility (multiple choice, open-ended, true-false), security, random question and answer orders for each user, and exam time and place restrictions. These differences may affect users' abilities and preferences regarding a system. Common advantages exist for online examination systems over traditional exams (Barbera, 2009, as cited in Gülbahar, 2013; Bennett, 2003; Whittington, Bull, & Danson, 2000):

- High quality question design (more variety than multiple choice), including high interaction with instant feedback
- Graphics and multimedia
- Individualization to match student characteristics and abilities with the help of a question bank
- Ability to repeat as frequently as desired to aid student learning (student-centered)

- Automatically scheduled and timed tests (registered via a module)
- Enhanced reporting software supporting both assessment and evaluation
- Cheating reduction through randomization of questions and use of question databases
- Fast grading not prone to human error
- Quality reports based on specific question performance
- Integration with existing information systems
- One central and easily updated copy of the test
- No extra costs, such as printing
- Distance learning support
- No limit in terms of place and time for users

Usability of online examination systems

Usability has been described as whether a system is good enough to please users in terms of fulfilling their needs. Systems should have interface characteristics that are pleasant and easy to use, learn, and remember with few errors (Nielsen, 1993). A usable system "should be used easily and effectively by a specific range of users whom either should be trained or user support should be given in order to overcome the specified range of tasks, within their specified range of environmental scenarios" (Shackel, 1991, p. 24). According to its definition, usability affects people when they access a system in terms of performing goals and tasks.

Usability issues are very important for learning and teaching systems because they affect students' achievement (Gülbahar, Kalelioğlu, & Madran, 2008). Problems that occur due to usability issues in those systems may not only decrease student achievement but also lead to difficulties during evaluation (Crowther, Keller, & Waddoups, 2004, as cited in Gülbahar et al., 2008). Further, a usable online examination system can save time for both instructors and students. Therefore, it is important to investigate usability problems in an effort to correct them where possible in an online examination system used widely by both students and instructors.

Not many usability studies exist in the literature related to online examination systems. Torkul et al. (2004) declared usability an important issue in their online examination system and added that such systems need to have an interface that is easy to use and learn. Gülbahar et al. (2008) conducted a usability study of a higher education web-based assessment and evaluation tool, WebQuest. Tests were administered to five students and five academicians through a user-based approach. Overall, the system was found to be usable with a success rate of 75% for tasks completed by both groups. However, users reported having difficulty accessing and using the materials. In another study, Karahoca et al. (2009) analyzed a web-based exam system in terms of learnability, controllability, design, and satisfaction issues identified by two groups of users, the Information Technologies and Construction Technology departments at a vocational high school. The researchers only found a statistically significant difference with regard to learnability between these two departments. Tüfekçi et al. (2013) developed a web-based exam system and performed a usability study for desktop and mobile versions. All participants easily and successfully completed tasks such as accessing an exam, displaying an exam's questions, proceeding to the next question, and submitting answers. Most participants had high satisfaction with the system and did not feel test anxiety since they took exams when and where they wanted.

Method

Mixed-method research design was employed in this study. Although both quantitative and qualitative data collection procedures were applied, the main emphasis was on quantitative data. The following research questions were investigated:

- Is there a significant mean difference among OSSI SUS scores of students according to usage frequency?
- Is there a significant mean difference among OSSI SUS scores of students from different departments?

- What are the students' ideas about the usability issues of OSSI?

Instruments

In this study, two instruments were used for data collection. The System Usability Scale (SUS) was employed to collect student usability scores for OSSI. This scale was developed by Bangor, Kortum, and Miller (2008) and adapted to Turkish by Çağıltay (2011). It consists of five positive and five negative statements on a 5-point Likert type scale (1 = strongly disagree to 5 = strongly agree). Possible scores ranged from 0 to 100. The SUS instrument offers several advantages to researchers in terms of ease of use and adaptability. According to Bangor, Kortum and Miller (2009), the SUS provides quick and easy data implementation and administration for facilitators as well as participants. In addition, it can be accessed on the Web from almost any interface, including cell phones, desktops, and tablets. Bangor et al. (2008) conducted a further study in which they applied the SUS to collect data from 2,324 participants with 206 usability tests over a 10 year period. Their results showed that the SUS was highly reliable and useful over a wide range of interface types.

A semi-structured interview was the second instrument for collecting qualitative data. The interview form consisted of a consent form and several open-ended questions. The following procedures were applied to establish the interview questions. After several questions were prepared by the researcher, they were checked by two usability experts. Based on their suggestions, some questions were excluded or revised to best fit the context being investigated. Finally a graduate student examined the final questions in terms of understandability, grammar, and wording.

Participants

In the quantitative part of the study, 370 undergraduate students participated. Their ages ranged from 17 to 29. All the participants had used OSSI at least once to take online exams in their courses. The most common courses were Computer I and Computer II, offered by the faculty of education to preservice teachers to prepare them to teach with ICT. Participants indicated being generally willing to use ICT in their courses. They were enrolled in diverse departments from the faculty of education at a public university in Turkey, including Computer Education and Instructional Technology (CEIT), Elementary Mathematics Education (EME), English Language Teaching (ELT), and Mentally Handicapped Teaching (MHT).

The qualitative part of the study included eight voluntary participants from diverse departments, three female and five male. Out of this group, three students were from ELT, two were from EME, and three were from CEIT. The eight participants from the qualitative part also participated in the quantitative part. Table 1 shows the overall distribution of participants according to their department.

Table 1. *Characteristics of Participants in Departments*

Departments	Frequency(f)		Total
	Female	Male	
CEIT	9	155	164
EME	47	36	83
ELT	27	48	75
MHT	48	-	48

Data analysis and collection procedures

Quantitative data were gathered via an online platform in the fall semester of academic year 2013–2014. A two-way ANOVA was conducted to investigate the mean difference of SUS scores of students from different departments with different usage frequencies. ANOVA was appropriate because it can be used to compare several means and the study had two independent

variables (Field, 2009). IBM SPSS (Statistical Package for Social Sciences) version 20 was used for conducting data analysis. The assumptions of normality and homogeneity of variance were checked in order to evaluate the appropriateness of the data for analysis in the first stage. Descriptive statistics for dependent variables were allocated in the second stage. Then, the results of ANOVA were addressed. Finally, post-hoc analysis was conducted.

In the qualitative part of the study, semi-structured interviews were conducted with eight voluntary participants from different departments. Before the interview, a consent form was given to each interviewee and they were informed about the study by the researchers. After the interviews, the data were transcribed and analyzed by the researchers according to Creswell's (2012) six steps of qualitative analysis. First, the data were prepared and organized. Then, the data were explored and coded, and themes were developed for reducing the data. Next, findings were represented and reported in figures and tables. Finally, the findings were interpreted in terms of the research questions.

Results

Quantitative results

A two-way independent ANOVA was conducted in order to investigate whether a statistically significant mean difference existed among the groups' SUS scores. First, assumptions were checked to evaluate the appropriateness of the data for ANOVA. While normality was not met, homogeneity of variance was met according to Levene's test, $F(6,363) = 1.871$, $p > .05$. According to Field (2009), *F* statistics can be quite robust against violations of normality.

Table 2 shows the descriptive statistics of the dependent variable across department and OSSI usage frequencies. The numbers of groups across different departments and usage frequencies varies. While CEIT students used the system the most, the other departments' were nearly equal in usage frequencies. According to the means of the SUS scores as shown in Table 2, the highest belongs to the low usage CEIT students (*M*: 87.22, *SD*: 11.21) and the lowest belongs to the low usage MHT students (*M*: 69.58, *SD*: 18.10).

The mean SUS scores of all students (*M*: 82.5, *SD*: 14.40) was higher than 70, indicating an acceptable usability level for a system according to Bangor et al. (2008). They equated SUS scores to adjectives to guide practitioner interpretation, and the mean SUS scores for OSSI correspond to "excellent."

Table 2. *Descriptive Statistics of the Dependent Variable across Departments and OSSI Usage Frequencies*

Usage frequency of OSSI	Department	$\bar{X}$	SD	N
Low frequency level (using system < 4 times)	CEIT	87.22	11.21	9
	EME	84.10	14.41	47
	MHT	69.58	18.10	48
	ELT	79.90	11.23	27
High frequency level (using system ≥ 4 times)	CEIT	83.68	12.04	155
	EME	82.30	12.86	36
	ELT	75.31	14.13	48

The SUS score, which is the dependent variable, and usage frequency and department, which are the independent variables, were considered during analysis. Before interpreting the main effects for each independent variable, the interaction effect was calculated between the independent variables and found to be non-significant, $F(2,369) = .20$, $p > .05$, $\eta^2 = 0.001$; therefore, the main effects were then interpreted.

The first independent variable, usage frequency, has two levels, low and high. Low corresponds to participants who used the system fewer than four times, while high corresponds to participants who used it at least four times. No significant mean difference was found between usage frequencies, $F(1,369) = 2.40$, $p>.05$, $\eta^2 = 0.007$. Thus, there was no main effect of usage frequency on SUS scores (see Table 3). Since no statistically significant difference was found, there was no need to conduct post-hoc tests or interpret partial η^2.

Table 3. *Two-way Analysis of Variance Summary Table for Departments and Usage Frequencies on SUS Scores of OSSI*

Source	SS	f	MS	F	p	η^2
Usage Frequency	440.47	1	440.47	2.40	.12	.007
Department	6659.50	3	2219.83	12.07	.00*	.091
Usage Frequency * Department	74.31	2	37.16	.20	.82	.001
Total	76491.23	369				

*$p < .05$

The second independent variable, department, has four groups: CEIT, EME, MHT, and ELT. At least one of the means of the departments was found to be statistically significant after analysis, $F(3,369) = 12.07$, $p < .05$, $\eta^2 = 0.091$. In order to determine which mean was different, a post-hoc test was conducted. Before interpreting the results, alpha (.05) was divided by the number of comparisons in order to control the probability of a Type I error during several pairwise comparisons (Field, 2009). Therefore, alpha was divided by three to reach .017, which was used as the significance level while comparing the means. The reference department was found to be CEIT when comparing the means. In addition, Hochberg's GT2 post-hoc procedure was applied due to the different sample sizes across the groups (Field, 2009). According to the results, there was a significant mean difference between the SUS scores of CEIT ($M = 83.87$) and MHT students ($M = 69.58$) and a significant mean difference between the SUS scores of CEIT ($M = 83.87$) and ELT students ($M = 76.97$), $p < .017$. On the other hand, there was no significant mean difference between the SUS scores of CEIT ($M = 83.87$) and EME students ($M = 83.31$), $p > .017$.

Qualitative results

This section is dedicated to the findings of the qualitative part of the study. The responses of the students are presented in four categories: Familiarity with OSSI, Usability-Related Concerns about OSSI, Features of OSSI, and Usability-Related Issues of OSSI.

Familiarity with OSSI

The results of the qualitative analysis showed that some of the students experienced difficulty getting accustomed to the online exams compared to traditional paper-based exams, even after using the system for several years and several times. One student stated:

> "*Until university from primary school, we have been taking paper based exams, but at university, we are taking some of the exams through this system online. Although I have been using this system for three years, taking online exams makes me feel anxious since I am unaccustomed to them and experiencing difficulty while getting accustomed."*

Usability-related concerns about OSSI

Some students expressed concerns regarding the usability of OSSI. More specifically, they described feeling anxiety due to possible errors that might occur during exams. One student explained:

> *"I wonder, what if the computer crashes during the exam, or if the Internet connection cut off, then what will I do? What happens then? Fortunately, these didn't happen to me. But I have concerns about these issues."*

The absence of detailed feedback in OSSI related to exams was another concern that students mentioned during interviews. One student observed:

"I think that the only missing point in this system is that we are not able to see which questions we did as correct and wrong."

Another problem was not being able to realize questions that were left blank in the exam. Thus, they had difficulty revising them. The final common concern was the persistent screen presence of the time remaining in the exam (see Figure 1). One student indicated that the countdown increased anxiety and caused difficulty concentrating.

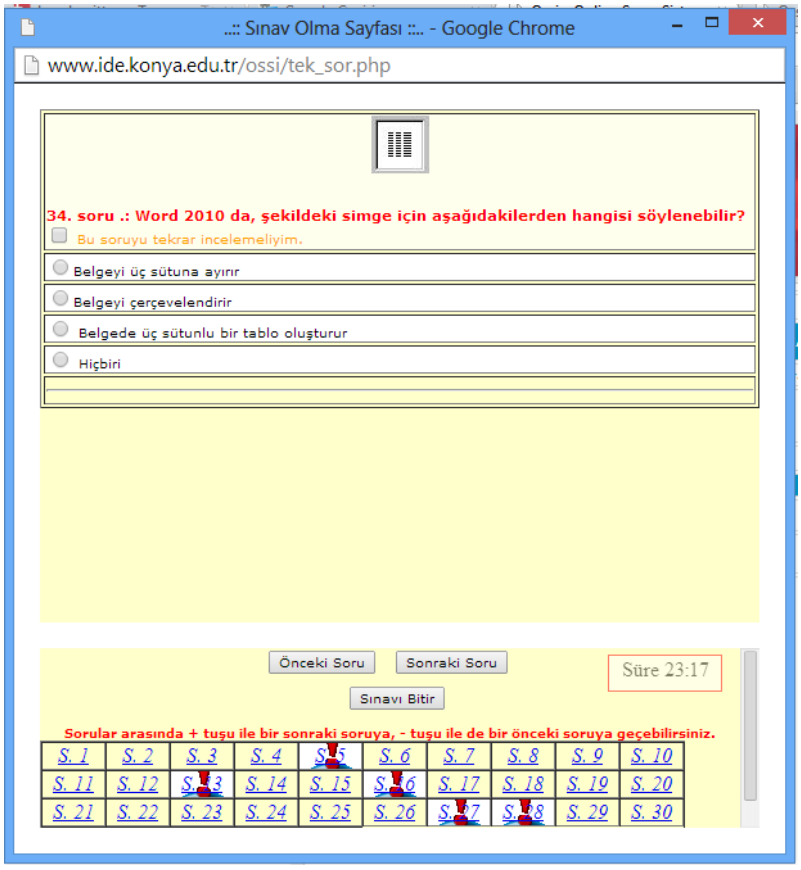


Figure 1: *Question Screen of OSSI*

Features of OSSI

The qualitative results indicated many affordances of online examination systems according to the students:

- easy and stress free exams,
- more reliable and systematic exams,
- fair assessment,
- the ability to take quizzes wherever and whenever,
- no disturbances or distractions during exams,
- more easily understood questions compared to handwritten ones, and
- quick exam results.

Usability-related issues of OSSI

The most commonly accessed part of OSSI is the question screen, so student ideas were gathered about its usability. Responses can be categorized under two themes: positive and negative. For the positive side, most students liked the ability to mark specific questions. One student stated:

"*One of the best benefits of the system is that it enables me to mark questions that I skipped and this way, I can easily turn back to the question and revise it."*

The ability to select choices by clicking anywhere on an answer rather than being forced to click the radio button was another positive note about the system.

As to negative comments about the question page, students reported a major struggle to answer questions that were truncated due to limited screen size. Considering students' constant fight against time, they complained about time wasted scrolling through screens. Moreover, they may have missed parts of questions at the bottom of a page. Students proposed adding the ability to change screen sizes, including an option for full screen.

Students also mentioned the placement of the navigation buttons used for moving between questions and finishing the exam (see Figure 1). These buttons are very close together, and some students accidentally submitted exams after clicking the wrong button. The formatting of the questions was another problem. Some students expressed that the font size was too small, while others complained about the background color. Solving one question on one screen rather than solving all of them on one screen, lack of instruction related to number of questions in the exams were other problems. One student stated:

"*I could not answer all the questions in the exam since I have forgotten the total number of questions due to not being able to see all questions on the screen, and there were no information related to number of questions on the screen.*"

Discussion

The purpose of this study was (a) to investigate mean differences of SUS scores for OSSI among different departments and student usage frequencies and (b) to collect in-depth ideas of students in regard to the usability issues of OSSI. In the first part of the study, a 2x4 ANOVA was conducted in order to answer the first and second research questions. A qualitative approach was applied for the third research question, investigating in-depth usability issues of OSSI.

With regard to the first research question, usage frequency levels had no effect on students' SUS scores. That is, how much time the students used OSSI had no effect on their SUS scores in regard to the usability issues of OSSI. This result could mean that student perceptions about the usability of the system did not change over time. This finding is consistent with the literature; in an evaluation of the usability of web media for educational purposes, it was stated that familiarity of users with the system has no effect on the determination of usability problems (Baş & Tüzün, 2011). Therefore, usability issues for an educational system should be taken into consideration independently.

For the second research question, significant mean differences were found between students' SUS scores for OSSI when comparing CEIT students to MHT and ELT students, separately. On the other hand, there was no statistically significant mean difference between CEIT and EME students' scores. CEIT students undoubtedly had higher ICT skills than their peers in other departments. Thus, significant differences between CEIT students and others in terms of usability issues are normal. These results are consistent with Karahoca et al. (2009), who found a statistically significant difference in the learnability component of usability between students in Information Technologies and Construction Technology departments at a vocational high school. Students with only basic ICT skills should be considered (Tüfekçi et al., 2013) when designing online examination systems that assess performance to prevent jeopardizing student achievement due to system usability problems (Gülbahar et al., 2008). Moreover, to provide an online examination system that works for all kinds of users, user types should be examined and the system should be designed in an adaptive, effective, and usable way.

The qualitative part of this study provided specific information related to the usability issues of OSSI. These results may improve the usability of OSSI as well as shed light on other online examination systems in terms of usability. The findings drawn from this study suggest

that when designing an online examination system, the following recommendations should be carefully considered:

- Screen size is very important for these systems. The entire text of a question should fit onto the screen, and scrolling should be minimal.
- The placement of buttons on the screen also matters to users. The button for submitting the exam should especially be reconsidered, and user confirmation should be solicited before closing the exam.
- The format of the questions is another topic that demands attention. Since lecturers format their own exam questions, they should revise them at least once in light of these issues.
- Each individual should be able to customize the system, such as moving or removing the exam timer countdown or changing the background color of the question screen.

References

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

Baş, T., & Tüzün, H. (2011). *Evaluation of the usability of educational web media: A case study of GROU.PS*. Paper presented at the 4th International Conference of Education, Research and Innovation (pp. 6738–6742). Madrid, Spain.

Bennett, R. E. (2003). *Online assessment and the comparability of score meaning*. Princeton, NJ: ETS.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston, MA: Pearson Publications.

Çağıltay, K. (2011). *İnsan bilgisayar etkileşimi ve kullanılabilirlik mühendisliği: Teoriden pratiğe*. Ankara, Turkey: ODTÜ Geliştirme Vakfı Yayıncılık.

de Siqueira, J. M., Martínez-Sáez, A., Sevilla-Pavón, A., & Gimeno-Sanz, A. (2011). Developing a web-based system to create, deliver and assess language proficiency within the PAULEX Universitas Project. *Procedia - Social and Behavioral Sciences*, *15*, 662–666. doi:10.1016/j.sbspro.2011.03.160

Field, A. P. (2009). *Discovering statistics using SPSS*. London, England: SAGE.

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, *57* (4), 2333–2351.

Gülbahar, Y., Kalelioğlu, F., & Madran, O. (2008). Usability evaluation of "Web Macerası" as an instructional and evaluation method. *Ankara University Journal of Faculty of Educational Sciences*, *41* (2), 209–236.

Gülbahar, Y. (2013). E-değerlendirme. In K. Çağıltay & Y. Göktaş (Eds.), *Öğretim teknolojilerinin temelleri: Teoriler, araştırmalar ve eğilimler*. Ankara, Turkey: Pegem Akademi.

Karahoca, A., Karahoca, D., & Günoğlu, S. (2009). *Web tabanlı sınav otomasyon sisteminin kullanılabilirlik analizi.* Paper presented at the 4. Ulusal Yazılım Mühendisliği Sempozyumu (pp. 183–190). Istanbul, Turkey.

Karakaya, Z. (2002). *Design, development and implementation of on-line exam system*. Paper presented at the International Open and Distance Education Symposium. Eskişehir, Turkey.

Measuring, Selection, and Placement Center (OSYM). (2013). *Ölçme, Seçme, ve Yerleştirme Merkezi Başkanlığı 2012 Yılı Faaliyet Raporu*. Retrieved from http://www.osym.gov.tr/dosya/1-69840/h/2012-yili-osym-faaliyet-raporu.pdf on February 5th, 2015

Ministry of National Education (MEB). (2013). *Motorlu taşıt sürücü adayları sınavı e-sınav uygulama kılavuzu*. Retrieved from http://www.meb.gov.tr/sinavlar/dokumanlar/

2012/kilavuz/9586_mtsas_klavuz_2012_2012_7_5_16_23_1.pdf on February 5th, 2015

Nielsen, J., (1993). *Usability Engineering*. Cambridge, USA: Academic Press.

Shackel, B. (1991). Usability - context, framework, definition, design and evaluation. In B. Shackel and S. Richardson (Eds.), *Human factors for informatics usability* (pp. 21–37). Cambridge, UK: Cambridge University Press.

Torkul, O., Kibar, A., & Tasci, T. (2004). *Web based examination systems*. Paper presented at the 1th International Conference on Informatics. Cesme, Turkey.

Tüfekçi, A., Ekinci, H., & Köse, U. (2013). Development of an internet-based exam system for mobile environments and evaluation of its usability. *Mevlana International Journal of Education*, *3*(4), 57–74. doi:10.13054/mije.13.59.3.4

Tümer, A. E., Şahin, I., & Aktürk, O. A. (2008). *Online exam system and student opinions about this system*. Paper presented at the 8th International Educational Technology Conference (pp. 322–327). Eskişehir, Turkey.

Whittington, D., Bull, J., & Danson, M. (2000). *Web-based assessment*: *Two UK initiatives.* Retrieved from http://ausweb.scu.edu.au/aw2k/papers/whittington/paper.html on January 12th, 2015

Uzun Öz

Giriş

Gelişen teknolojinin insan hayatını birçok yönde etkilemesinin bir yansıması olarak eğitim sürecinde kullanılan geleneksel yaklaşımlar da değişmektedir. Bu yansımalardan bir taneside klasik kağıt kalem ile yapılan sınavların çevrimiçi olarak internet üzerinden yapılmaya başlanmasıdır. Çevrimiçi sınav yapan kurum sayısı dünyada arttığı gibi (Gülbahar, 2013), ülkemizde de ÖSYM ve MEB tarafından kullanılmaya başlanmıştır. Çevrim içi sınav sistemleri, sınav sorusu hazırlama, sınav olma ve sonuçları duyurma gibi geleneksel sınava ait tüm işlemlerin internet üzerinden yapıldığı çevrimiçi ortamlar şeklinde tanımlanabilir (Tüfekçi vd., 2013). Çevrimiçi sınav sistemlerin grafik ve çoklu ortam ögelerini bakımından zengin ve kaliteli sorular hazırlanabilmesi, öğrenciler için bireyselleştirme, hızlı notlandırma, insandan kaynaklanan hataları en aza indirme başta olmak üzere bir çok faydası bulunmaktadır.

Nielson (1993) kullanılabilirlik kavramını, var olan bir sistemin kullanıcılarını memnun etmesi ve bu doğrultuda mümkün olan en az hata payı ile kullanılabilir, öğrenilebilir ve kolay hatırlanabilir kullanıcı dostu bir ara yüze sahip olması şeklinde tanımlamıştır. Gülbahar, Kalelioğlu ve Madran (2008) öğrenme ve öğretme sürecinde kullanılan sistemlerde kullanılabilirliğin öğrenci başarısını etkileyebileceğinden dolayı daha önemli olduğunu vurgulamışlardır. Bu sistemlerdeki sorunların öğrencilerin başarısını düşürebileceği gibi aynı zamanda değerlendirme sürecinde de zorluklar oluşturabileceğini belirtmişlerdir. Buna ek olarak kullanılabilir sistemler hem öğrencilere hem de sistemi kullanan eğiticilere zaman kazandıracaktır. Bu nedenle çevrimiçi sınav sistemlerinin kullanılabilirlik yönünden incelenmesi önemlidir.

Bu çalışmanının amacı bir devlet üniversitesinde vize ve final sınavlarında kullanılan çevrimiçi sinav sistemi, OSSİ'nin kullanılabilirliğini incelemektedir. Bu amaçla farklı bölümlerden ve sistemi farklı sıklıklarda kullanan öğrencilerin Sistem Kullanılabililik Ölçeği ile elde edilen puanları arasında anlamlı ortalama farkı araştırmacılar tarafından incelenmişir. Buna ek olarak öğrencilerin sistem hakkında görüşleri alınarak kullanılabilirlik ile ilgili diğer hususlar da açığa çıkartılmıştır.

Yöntem

Araştırma desenlerinden karma araştırma deseni bu çalışma için belirlenmiş olup ağırlık daha çok nicel veriye verilmiştir. Çalışmada nicel veriler Bangor, Kortum ve Miller (2008) tarafından geliştirilen ve Çağıltay (2011) tarafından Türkçeye Sistem Kullanılabilirlik Ölçeği (SKÖ) olarak çevrilen 5'li Likert tipli 10 maddeden oluşan ölçek kullanılmıştır.Çalışmanın katılımcılarını 2013-2014 sonbahar döneminde bir devlet üniversitesinde sistemi kullanmakta olan 370 lisans öğrencisi oluşturmaktadır. Bu öğrenciler Bilgisayar ve Öğretim Teknolojieri Eğitimi (BÖTE), İlköğretim Matematik Öğretmenliği (İMÖ), İngilizce Öğretmenliği (İÖ) ve Zihinsel Engelliler Öğretmenliği (ZEÖ) bölümünde okumaktalar ve sistemi farklı sıklıklarda kullanmışlardır. Nicel veri analizinde iki yönlü varyans analiz kullanılmıştır. Nitel veriler ise çalışmaya katılanlar arasından gönüllü olarak seçilen 8 öğrencinin yarı yapılandırılmış görüşme yoluyla sistem ile ilgili görüşleri alınarak toplanmış ve Creswell'in (2012) altı aşamalı nitel veri analizi tekniği kullanılarak analiz edilmiştir.

Sonuç ve Tartışma

OSSİ'nin SKÖ ile elde edilen puanların ortalaması (*M*: 82.5, *SS*: 14.40) .70 üstü olduğu için Bangor ve diğerlerinin (2008) kriterine göre genel olarak sistemin kullanılabilir olduğu söylenebilir. Nicel kısımda bağımsız değişkenler iki seviyeli kullanım sıklığı (düşük ve çok kullanım) ile dört seviyeli bölüm olarak belirlenirken, bağımlı değişken ise öğrencilerin SKÖ ile sisteme verdikleri puan olarak belirlenmiştir. Yapılan iki yönlü bağımsız varyans analizi sonucu, kullanım sıklığının iki farklı seviyesi arasında anlamlı bir ortalama farkı bulunmazken, $F(1,369) = 2.40$, $p > .05$, $\eta^2 = 0.007$, bölümlerden en az birisi ile diğerleri arasında anlamlı bir ortalama farkı bulunmuştur, $F(3,369) = 12.07$, $p < .05$, $\eta^2 = 0.091$. Yapılan post-hoc testleri sonucunda BÖTE ($M = 83.87$) ile ZEÖ öğrencileri ($M = 69.58$) ve BÖTE ($M = 83.87$) ile İÖ ($M = 76.97$) öğrencileri arasında anlamlı bir ortalama farkı bulunmuştur. Ancak BÖTE ($M = 83.87$) ile İMÖ ($M = 83.31$) öğrencileri arasında anlamlı bir ortalama farkı bulunamamıştır. Sistemi farklı seviyelerde kullanan öğrencilerin SKÖ sonucu elde edilen puanların ortalaması arasında anlamlı fark olmaması öğrencilerin sistemi ne kadar sık kullanırlarsa kullansınlar kullanılabilirlik ile ilgili görüşlerinin değişmediğini göstermektedir. Bu nedenle kullanılabilirlik ile ilgili hususlar sistem, ilk başta tasarlanırken göz önüne alınmalıdır. BÖTE ile diğer iki bölüm öğrencilerinin SKÖ sonucu elde edilen puanların ortalaması arasında bulunan anlamlı fark ise BÖTE bölümü öğrencilerinin bilgi ve iletişim teknolojilerini diğerlerine göre daha iyi bilmesi şeklinde yorumlanmıştır. Bu nedenle bu tarz eğitim-öğretim sürecinde kullanılacak sistemler tasarlarken sistemi kullanabilecek en düşük BİT becerisine sahip öğrenci profili göz önüne alınmalı ve sistem kullanılabilirlik yönünden o kritere göre tasarlanmalıdır.

Çalışmanın nitel boyutunda ise öğrencilerin sistem ile ilgili görüşleri alınmıştır. Yapılan analiz sonucunda bazı öğrencilerin sisteme alışmalarında yaşadıkları zorluklar kalem ve kağıt ile yapılan sınavlar ile kıyaslandığında daha zor olduğunu belirtmişlerdir. Bir diğer sorunun öğrencilerin sistem kullanılabilirliği konusundaki endişeleridir. Bu endişelerin bazıları sistem dışı internetin gitmesi, bilgisayar donması gibi etkenler olabilirken, sistemden kaynaklı da olabilmektedir. Sistemin öğrencilere ayrıntılı dönüt vermemesi, boş bırakılan soruların fark edilememesi, sürekli ekranda azalan zamanın olması öğrencilerin endişelerini artırdığı belirtilmiştir. Sistemin özellikleri olarak adil değerlendirme, güvenli ve sistematik sınav olma, sınavları istediğin zaman ve mekanda olabilme, hızlı sonuç açıklama şeklinde belirtilmiştir. Son olarak öğrencilerin sınav sorularını gördükleri ekran ile ilgili görüşleri alınmış ve görüşme sonucunda sorulara tekrar dönmek için soruları işaretleyebilme özelliğinin olması, doğru cevabı seçerken şıkkı işaretlemenin yeterli olması olumlu yanları olarak belirtilirken; ekran boyutunun sınırlı olması, sorular arasında gezinme düğmelerinin birbirine çok yakın olması, soruların formatı, ekranda sınavdaki soru sayısı yönergesinin eksikliği olumsuz yanları olarak belirtilmiştir.

Çalışma sonunda çevrimiçi sınav sistemi tasarlanırken şunlara dikkat edilmesi gerektiği sonucuna varılmıştır. a) ekran boyutu soruların tamamını kapsayacak şekilde ayarlanmalı,

kaydırma çubukları en aza indirilmeli, b) ekrandaki düğmelerin yerleşimi iyi ayarlanmalı ve bir birine çok yakın olmamalı, c) sorular düzenlenirken okunabilirlik yönünden gözden geçirilmeli, d) kullanıcılar sistemi kendilerine özgü şekilde kişiselleştirebilmeli.