VIEWS OF INSTRUCTORS AND STUDENTS REGARDING ONLINE MEASUREMENT AND EVALUATION APPLICATIONS

Gonca Usta[1]; Yurdagül Günal[2]; Müge Uluman[3]

[1]Cumhuriyet University, Faculty of Education, Department of Educational Sciences, Sivas
[2]KaradenizTechnical University, Fatih Faculty of Education, Department of Educational Sciences, Trabzon
[3]Marmara University, Faculty of Education, Department of Educational Sciences, İstanbul.
[2]yurdagulgunal@hotmail.com

ABSTRACT

The purpose of this article was to describe online measurement applications in Turkey by using opinions of instructors as practitioners and students about online measurement. The study group of this qualitative research consisted of 8 instructors offering courses in distance education and 46 students of distance education from Karadeniz Technical University and the University of Cumhuriyet. Answers of lecturers and students were often found to be in parallel. This study finds out that one outstanding difference between participants' answers that students found online measurement practices reliable whereas instructors disagreed. Students' opinions about online measuring applications seem more positive than instructors. Our recommendations include taking measures to solve the problems of internet connection, utilizing different types of materials besides multiple-choice items, and preparing of items considering item writing rules regardless of the type of items.

Keywords: online measurement, online evaluation, instructors, students, internet connection

INTRODUCTION

Developments in technology also lead to different applications in the field of assessment and evaluation which is an integral part of education. According to Vendlinski, Underdahl, Simpson, and Stevens have stated that technology offers novel measurement and evaluation methods which provide a wide range of efficient sources for effective teaching and learning for educators (as cited in Buzzetto-More & Alade, 2006). Online measurement is a process used for measuring a particular aspect purpose by means of a computer connected to the Internet. The wide concept of online measurement and evaluation can be referred as a process consisting of digital technology activities used in evaluation (The Joint Information Systems Committee [JISC], 2007). According to Semerci and Bektaş (2005), Internet-based measurement is a basic element that needs to develop along with instruction process rather than as a stage located at the end of it as it first comes to mind. Online measurement can be seen as an element promoting measurement as an integral part of learning (Russell, Elton, Swinglehurst & Greenhalgh, 2006). Therefore, Russel, et al. (2006) defines online measurement as measurements which facilitate high-quality feedback between teachers and students and contributes to learning that supports collaborative learning.

Online measurement is used to keep track of the process due to its ability to perform measurement quickly. Following the development of the student, it is intended to provide effective feedback for teachers, students, and parents. Such an evaluation is not only a useful tool for teachers to follow student performance but also very useful for students for monitoring their level of development, easiness of application and flexibility of time and location (Trillo, Illari, Lopez & Barisaboa, 2007). Unlike paper and pencil tests, online measurement application provides convenience

in terms of assessment. Besides, it allows giving fast and effective feedback. Comparison of paper and pencil tests with Inter-based evaluation yields results for the latter (Olsen, Maynos, Slawson & Ho as cited in Kumalar & Pürtaş, n.d.). Besides all these, the online testing system reduces unnecessary paper waste and is more environmentally friendly. In this respect, it can be said to be more useful than paper-and-pencil tests. On the other hand, the reliability of online measurement tests is a controversial issue since examinees cannot be controlled during testing (Rowe, 2008). Therefore, despite providing such advantages as being practical, it is better reflecting student performance and creating a stress-free test environment, a question such as "Can we trust the results of the online evaluation system?" comes to mind.

Online measurement has also disadvantaged. In order to overcome these advantages, the same features to be included in a pen-and-paper test are sought in the online measuring system. Therefore, in order to ensure the validity of online measurement, certain requirements must always be observed. According to Semerci and Bektaş (2005), these requirements are as follows (1) For ensuring validity in measurement, reliability must be in its place as a basic requirement; (2) Students must be prevented from cheating; (3) It must be made sure that the person who is taking the test is the right person; (4) It is necessary to ensure content validity, which is the balance of item subject; (5) Examination question must be constantly analysed to ensure repeatability; (6) Testing must be done in a suitable environment free from drawbacks such as noise and extreme heat and cold; (7) Uninterruptible power supply must be created on the client computers.

If all these conditions are fulfilled, online measurement can be said to be very convenient, practical and comfortable for both students and teachers. Especially in the changing and developing world, access to information has reached an insuperable speed and technology has improved. As a result, the internet has become the most common means of access to information. When internet access to information is called in the education system, what first comes to mind is distance education. Online evaluations used in distance education help assess students' academic achievement. In addition, as a different form of distance education, in the context of e-learning applications, some psychological testing applications are used for evaluation. Some universities abroad that offering undergraduate, postgraduate, and doctoral study such as University of Southern Queensland, CQ University, and Fayetteville Technical Community University also provide training on the Internet. In Turkey, distance education is realized via the internet at universities such as Sakarya University, METU, and Anadolu University. Hence, some or all of the measurement and evaluation process is carried out on the internet for those receiving training by this means.

Considering the expansion of distance education and performing assessment and evaluation operations via the internet today, there are advantages and disadvantages of the online measuring system. It seems essential to obtain views of those applying this system and also students. The aim of this study is to describe online measurement applications in Turkey by using opinions of instructors, as practitioners, and students about online measurement. To this end, answers are sought for following research questions: (1) How do you find the online measurement applications? (2) How do you find types of the item used in online measurement applications? / What types of items do you use when performing measurement applications online? Why? (3) What are the problems you encounter during online measurement applications?", "What are your solutions for those problems?"

METHODS

This research is designed as a qualitative study which attempts to find out views of instructors and students involved in distance education regarding online measurement applications. A case study is used thereby the subject matter, and scope of the research is discussed by using a certain group of

the population. The study focuses on views of instructors and students as providers and recipients of distance education about online measurement practices, the problems they face and potential solutions for these problems. The study group consists of 8 instructors offering courses in distance education and 46 students of distance education from Karadeniz Technical University and the University of Cumhuriyet.

Study data are collected with semi-structured interview forms that developed by researchers separately for instructors and students according to the purpose of the study. Both of the interview forms are comprised of four open-ended question and relevant probes. The forms are identical except for question two and probe two. The instrument is developed by relevant literature. To ensure compliance of the instrument with the study aim; five assessment and evaluation specialists and two Turkish language specialists revised and commented on it. The form was revised and finalized in the light of the feedback received.

For validity and reliability of the data collection instrument, these studies are carried out by initially, credibility, reliability, and availability. Initially, the concepts such as credibility, transferability, consistency and availability of confirmation are explained for validation of reliability and validity in the study (Mills, 2003). While credibility (internal validity) is carried out to improve internal validity of the study (credibility), analysis of the data that performed on multiple expert opinions. Transferability (external validity) is aimed to improve transferability of the study, stages of the research process that described in detail. In particular, themes and categories are defined so clearly that readers can understand easily. Consistency (internal reliability) is the findings that presented directly away from comments and generalizations to ensure internal validity of the study. Obtained data are analyzed and coded separately, and the general consensus that reached among all codes. Reliability between coders is calculated by using Miles and Huberman's (1994) correspondence percentage. Reliability = Consensus / (Consensus+ Disagreement). Reliability of this study is found to be 97. Availability of confirmation (external reliability) is aimed to ensure the external reliability of the study; the interviews obtained from the participants are kept for specialized examination by those concerned.

Qualitative data are analyzed with descriptive analysis. In this procedure, obtained data are summarized and interpreted based on the themes that previously identified. Descriptive analysis is performed in four stages, namely creating a framework for descriptive analysis, processing data within the thematic framework, defining the findings, and interpretation of the results (Yıldırım & Şimsek, 2005). To test the reliability, obtained data are coded by two researchers independently from each other. Then, coding results are compared. Both researchers are found to have done coding in the same way. In the final stage, raw data and analyzed data are checked by two other researchers (Yıldırım & Şimşek, 2005, Bilgin, 2006).

RESULTS AND DISCUSSIONS

Findings that obtained from descriptive analysis of the study data collected with semi-structured interview forms are presented in the study. Answers given for questions address to study participants are divided into sub-categories. The question "What can you say when you compare online measurement applications with pen-and-paper tests?" is included in the students' interview form as they are the direct recipient of the application. As for the study, 12 students think that online measuring applications are not different from the other measuring applications, while another 12 students think online measuring is easier than pencil-paper testing. 18 of the students have stated that they could reflect their actual performance better in online measuring applications. 21 students also include some from the abovementioned group of 18 students (K5, K9, K13, and K14). It is reported

that they feel less nervous during online measurement applications. Statement of K5 on this issue is quoted below.

K$_5$: "…For online measurement applications, I attend in my own home with my own computer. Therefore, I do not feel like myself as taking an exam…"

For question "How do you find the question types used in online measurement applications?", Findings obtained from both students and teachers are summarized that as seen in the study, students reported that generally multiple-choice questions are used (f = 35), and they prefer that type of questions to open-ended question type (f = 9). They also have stated that different types of questions must be included (f = 15) and better-qualified questions must be given (f = 21). In addition to this, they found the questions challenging due to the online application (f = 25). K4 has stated her/his views on this subject as following:

K$_4$: "…Using the multiple choice substance rather open-ended substances would be more appropriate. Because it is forcing me to use the keyboard instead of pen-and-paper for replying open-ended items…"

Instructors' views regarding the type of questions used are the study shows that instructors prefer multiple-choice items since assessment does not take long then (f = 5). However, one respondent prefers open-ended questions since they provide an opportunity for a detailed assessment. There are two respondents states that they pay attention to using various types of questions. Respondents of K18 and K22 express their views as following:

K$_1$: "…I am using the multiple-choice question type so that assessment doesn't become time-consuming…"

K$_{22}$: "… I am using the multiple-choice question type due to the high number of students…"

In relation to the type of questions that used in the online measurement applications, comparison of students' and instructors' opinions reveals that both participant groups prefer multiple-choice items. Instructors indicate that they prefer multiple-choice tests because it saves much time in assessment.

In relation question "What are your reasons for preferring the online measurement applications?" views of students and instructors are summarized that the most frequently mentioned reason for preferring online measurement applications are found as the convenience of being able to take the exam even in the case of being in another city (f = 45). In other words, the most important reason is providing flexibility in terms of application space and time. In addition, the opinion that stress was less than in the pen-and-paper exam (f = 21) has come to the fore. In relation to this question, participant K7 has stated her/his views as follows:

K$_7$: "…My biggest reason for preferring is the lack of obligation to take the test in the city where the school is located. Also it reduces exam stress, which can be considered to be one of the reasons…"

Instructors' views regarding why they prefer online measurement applications because as seen in the study the reasons for which instructors prefer online measuring include ease of application (f = 2), short duration of assessment (f = 2) and being less costly (f = 3), which indicate practicality altogether. When two groups of participants are compared regarding the reasons for preferring online measuring, it seems that students can take it stress-free without limitation of space. On the other hand, teachers seem to prefer this since it is a useful tool for practical scoring. Students' and instructors' opinions regarding the problems faced in online measuring applications are summarized that the

probability of electricity cut-off during examinations stands as the biggest problem faced by students (f = 28). The breaking of the internet connection during the examination (f = 21) is reported as another problem. These are followed by mistakes made during saving (f = 15) and challenges in using the keyboard (f = 9). In relation to online measurement applications, K2 has stated her/his views as follows:

K_2:"…Power failure during the test, breaking of the Internet and especially difficulty of using the keyboard made victims both me and my friends."

Instructors' views on this subject that the challenges faced in the implementation process because students do not have good computer literacy skills stand as the most important problem from the instructors' perspective (f = 5). It is followed by other challenges also reported by students such as breaking of the Internet connection (f = 4) and reliability of the application (f = 4). Not being able to identify students' levels was specified as another challenge (f = 2). In this question, K18 and K22 have stated as follows:

K_{18}:"…I have some reservations about the reliability. During the measurement, it is possible that students can achieve online resources related to measurement from another connection. This in turn negatively affects the reliability of the application…"

K_{22}: "…In teacher-made tests, level of the class can be taken into account. However, in online measurement and evaluation, it can be limited to evaluate it sufficiently. Therefore, sometimes unexpected results may occur instead of success…"

In relation with challenges encountered during applications, students emphasized power outages or Internet breaking during the testing also stating that using the keyboard was compelling for them. On the other hand, instructors indicated problems regarding reliability of the application and computer literacy level of students.

Bearing in mind these challenges reported, students and instructors' views regarding potential solutions directed at these challenges are given as seen in study 8, students' first proposal is to provide education to whom will participate in the practice for the first time (f = 27). It is also recommended to hold examinations in collective venues to prevent cheating (f = 18). Students' other recommendations are to improve questions (f = 12) and pose questions that require less keyboard use (f = 12).

Instructors' views on this matter are given that recommendations brought by instructors included realizing the applications in a test centre (f = 5) and using open-ended questions instead of multiple choice ones (f = 3). As an example, part of K21's views is given below.

K_{21}: "...I think that the most important issue comes from exam formats. Open-ended questions should be preferred instead of multiple choice question type. But high number of students, requirement of a long time for assessment and course loads of the teaching staff is a serious obstacle to this situation…"

In the context of recommendations, one point is found to be shared by instructors and students. Both groups of participant believe that examinations must be held in testing centers' views. Apart from this, students bring other recommendations regarding potential challenges they could face during implementation. According to the students, there is no difference between online measuring and other measuring applications and variety is needed. As a response to the question "What do you think of the types of items used in online measurement applications?", they point out that their views show direct proportion with the quality of the items addressed in the application. Main problems encountered in the context of online measuring included breaking of the internet, cutting of electric,

the use of the keyboard, and file saving problems. Students think that online measurement applications are reliable and useful for reducing stress during the exam.

Instructors' interviews demonstrate that online measuring applications are realized through distance education center by which students log in with their identity number and password and answer the items within given period. As for the problems faced during online measurement, they reported that easily addressable problems such as breaking of the Internet connection and lower levels of students' computer literacy sometimes occur. Above all, they expressed their suspicion regarding the reliability of online measurement applications. Lastly, their outstanding recommendation included using of open-ended questions rather than multiple-choice items, realizing the applications in predetermined testing centers within a given period, and conducting of face-to-face measuring applications instead of online measurement.

CONCLUSIONS

The first important finding of this study is that students think online measurement is not different from pencil-and-paper test, and even they could better reflect their performances in online measurement because they are less nervous. Compared to the conventional examinations, online examinations, from the students' point of view, encourage students to acquire deeper knowledge while preparing for the examination and guide them to do research for in-depth knowledge. This finding is supportive of the previous research results (Shen, 2005; Koong & Wu, 2010).

In terms of the types of questions used, both students and teachers prefer for multiple-choice questions. The students prefer this type of questions because it requires less use of keyboards, while instructors emphasize the requirement of shorter time for applying. Comparison of students' success by question types used in online measuring is revealed a slight difference. Students' performance in multiple-choice items seems to be higher than in open-ended questions (Nakayama, Yakamoto & Santiago, 2010).

Concerning the reasons for preferring online measuring applications, students highlight such advantages as taking the test without limitation of space and in a stress-free environment. However, instructors are seen to indicate practicality of implementation as the main reason. Other studies also reveal that students feel more comfortable during online exams and assessment (Dermo, 2009). In Patterson (2006), it is found out that the opportunity of taking online exams from anywhere provides a more relaxing testing atmosphere and reduces students' anxiety.

In the context of the main problems faced during implementation, students often report electric cut-off and internet breaking also followed by the use of the keyboard. Unlike the students, instructors note issues regarding reliability of the application and low level of computer literacy among students. In a study carried out by Gül and Doğan (2011), it is found that reliability and distinctiveness of online measuring applications are higher than pencil-and-paper tests. Furthermore, Nunnaly and Berstein (1994); Chase (1999) have noted that distinctiveness of online measuring applications is higher. Tümer, Şahin and Aktürk (2008) have carried out a study examining whether or views of learners regarding online measurement differ between students who have and do not have a personal computer. As a result, they find out that the students with a personal computer find it easier to use the system for online measurement.

In relation to recommendations for the problems faced, it is seen that instructors and students have a consensus for realizing the applications in a testing center. Responses of instructors and students seem to be in parallel with each other. Yet, one outstanding difference is found between

students and instructors in those students believe online measuring applications are reliable whereas instructors do not. Though students and instructors preferred multiple-choice questions for online measuring applications, they report internet connection and computer literacy levels as the most important problems, respectively. In the particular context of items in the test, the students expectedly indicate that test items should be typed by observing the rules applicable for writing items. Last but not least, the most remarkable recommendation from the instructors' side is for conducting face-to-face measuring applications rather than online measuring.

REFERENCES

Bilgin, N. (2006). *Sosyal Bilimlerde İçerik Analizleri: Teknik lerve Örnek Çalışmalar*. Ankara: Siyasal Kitabevi.

Buzzetto-More, N. A., & Alade, A. J. (2006). Best practices in e-assessment. *Journal of Information Technology Education*, 5, 251-269. Retrieved February 28th 2012 from http://jite.org/documents/Vol5/v5p251-269Buzzetto152.pdf adresindenerişilmiştir.

Chase, C. I. (1999). *Contemporary assessment for educators*. New York: Longman.

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

Gül, E., & Doğan, Ç. (2011). Online değerlendirme güvenilir midir?. *5th International Computer & Instructional Technologies Symposium*. Fırat University, Elazig - Turkey.

Koong, C. S., & Wu, C. Y. (2010). An Interactive Item Sharing Website for Creating and Conducting On-line Testing. *Computers & Education*, 55, 131-144.

Kumalar, M., & Pürtaş, M. (n.d). İlköğretim İçin Bir Online Sınav Sistemi Önerisi. *XIV: Akademik Bilişim Konferansı Bildiri Kitabı.* Pp 129-132.

Miles, H. B., & A. M. Huberman. (1994). *Qualitative Data Analysis.* Sage Publication: Thousand Oaks California.

Mills, G. E. (2003). *Action research: A guide for the teacher researcher*. Upper Saddle River, NJ: Merrill/Prentice Hall.

Nakayama, M., Yamamoto, H., & Santiago, R. (2010). The Role of Essay Tests Assessment in e-Learning: A Japanese Case Study. *Electronic Journal of e-Learning*, 8(2), 173-178.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. 3rd edition. New York: McGraw-Hill.

Olsen, J. B., Maynes, D. M, Slawson, D. A. & Ho, K. (1986). Comparison and Equating of Paper administered, Computer-administered and Computerized Adaptive Tests of Achievement. *The meeting of the American educational Research Association San Francisco*. San Fransisco, USA.

Patterson, D. A. (2006). A Large-Scale, Asynchronous Web-Based MSW Comprehensive Exam Administration: Outcomes and Lessons Learned. *Journal of Social Work Education*, 42(3), 655-668.

Russell, J., L. Elton, D. Swinglehurst, & T. Greenhalgh. (2006). Using the online environment in assessment for learning: A case-study of a Web-based course in primary care. *Assessment and Evaluation in Higher Education*, 31 (4), 465–478.

Rowe, N. C. (2008). Cheating in Online Student Assessment: Beyond Plagiarism. *Online Journal of Distance Learning Administration*, 7 (2).

Semerci, Ç., & Bektaş, C. (2005). İnternet Temelli Ölçmelerin Geçerliğini Sağlamada Yeni Yaklaşımlar. *The Turkish Online Journal Of Educational Technology*, 4 (1), 130-134.

Shen, J. (2005). *Collaborative Examinations in Asynchronous Learning Networks: Field Experiments on Collaborative Learning through Online Assessments* (Doctoral Thesis). USA: New Jersey Institute of Technology.

The Joint Information Systems Committee [JISC]. (2007).

Trillo, R., Ilarri, S., L´opez, J. R., & Brisaboa, N. R. (2007). *Development of an on-lıne assessment system to track the performance of students.* Retrieved May 25th 2015 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.7210&rep=rep1&type=pdf

Tümer, A. E., Şahin, İ., & Aktürk, A. O. (2008). *Online sınav sistemi vebu sistemi leilgiliöğrencigörüşleri. 8th International Educational Technology Conference.* Anadolu University, Eskişehir.

Yıldırım, A., & Şimşek, H. (2005). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri.* Ankara: Seçkin Yayıncılık.