Nikou, S.A. and Economides, A.A. (2019), "A comparative study between a computer-based and a mobile-based assessment: Usability and user experience", *Interactive Technology and Smart Education*, Vol. 16 No. 4, pp. 381-391. https://doi.org/10.1108/ITSE-01-2019-0003

Comparing usability and user experience between a computer-based and a mobile-based assessment

Stavros A. Nikou,

School of Education, University of Strathclyde, Glasgow, UK

Anastasios A. Economides

Interdepartmental Programme of Postgraduate Studies in

Information Systems, University of Macedonia, Egnatia Street 156, 546 36,

Thessaloniki, Greece.

Abstract

Purpose - Beyond the traditional paper-based format, assessment can also be delivered in online environments via desktop computers or mobile devices. While Computer-Based Assessment (CBA) is being extensively used, Mobile-Based Assessment (MBA) is a relatively new practice in many educational institutions. The current study compares the usability and user experience between computer-based and mobile-based assessments in the context of a University course.

Design/methodology/approach - The study follows a between groups design. The participants were one-hundred and ten first-year undergraduate students from a European University. Students in the experimental group participated in the assessment using mobile devices while students in the control group participated using desktop computers. After the assessment, students self-reported their experiences with CBA or MBA respectively. The instruments used were the User Experience Questionnaire and the System Usability Scale.

Findings – Attractiveness and novelty were reported significantly higher in the experimental group (MBA), while no significant differences were found between the two groups in terms of efficiency, perspicuity, dependability and stimulation. The overall score for the System Usability was not found to significantly differ between the two conditions.

Practical implications - Implications for educational practitioners regarding the potential of using mobile devices as an alternative assessment media are discussed.

Originality/value – By comparatively investigating the usability and user experience of computer-based and mobile-based assessment, the study makes a contribution towards the adoption of computer- and mobile-based assessment in the context of tertiary education.

Keywords Usability, User experience, computer-based assessment, mobile-based assessment
Paper type Research Paper

1. Introduction

The proliferation of desktop computers and mobile devices in education introduced a wide range of computer-based and mobile-based instructional methods including computer-based and mobile-based testing and assessment. Computer-based testing and assessment offer several advantages to academics and practitioners like real-time feedback and scoring, automated score processing and analysis, improved security, cost and time reduction (Terzis and Economides, 2011). Moreover, mobile-based testing and assessment can be implemented anytime and anywhere eliminating the need for specialized computer labs, can offer extended capabilities such as personalized instruction, context-awareness and ubiquity and can facilitate formative, self- and peer- assessment as well (Authors, 2017). According to a recent literature review (Authors, 2016), many studies on computer-based and mobile-based assessments provide evidence for a positive impact of computer - and mobile-based assessments on student learning attitudes, motivation and performance.

However, the introduction of computer- and mobile-based testing and assessment in education is developing rather slowly due to barriers such as technical infrastructure, content and curriculum, professional development, organization and leadership (Lucas, 2018; Deutsch *et al*., 2012). User adoption is a critical factor for the successful implementation of any information system. According to the Technology Acceptance Model (TAM) (Davis, 1989), an information system is adopted by users (behavioral intention to use) when the system is easy to use (high perceived ease of use) and it is considered useful (high perceived usefulness). Usability and user experience are two main determinants for successful adoption of an information system. Usability is defined as “the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (ISO 9241-11:2018, 2018). User experience is defined as “a person’s perceptions and responses that result from the use and/or anticipated use of a product, system, or service” (ISO 9241-210:2010, 2010).

Gaining a better understanding of the usability and user experience of computer- and mobile-based testing and assessment can help the design of better assessment experiences for students, facilitating their further adoption in education. Usability evaluation of computer-based instruction is not new and challenges in computer-based instruction and assessment have been successfully addressed to a large extend (Bartram, 2008). However, usability evaluation of mobile learning applications is an active area of research (Kumar and Mohite, 2018). Most studies focus on hardware and operating systems characteristics or single specialized apps (Coursaris and Kim, 2011). Moreover, there is still little comparative research focusing on computer- and mobile-based assessment usability aspects. The current study aims at comparing computer- and mobile-based assessments on usability and user experience metrics. The study findings can facilitate a manageable and interchangeable usage of computer- and mobile-based assessments in educational institutions.

The paper is structured as follows: in the Methodology section we describe the instruments used, the participants and the procedure. In the analysis section we describe the user tests. In the conclusions section we present the results of the user

tests, we draw conclusions and their implications for education practitioners and also the limitations of the study and future work.

2. Methodology

2.1 Participants

The participants were 110 first-year undergraduate students, 51 males (46%) and 59 females (54%), enrolled in an introductory informatics course, in the Department of Economic Sciences of a European University. The average age of the students was 20.2 (SD=1.99). All students had had the same exposure to information technologies. Based on a self-reported questionnaire about computer efficacy and mobile-devices efficacy, that they were asked to fill in in advance (Kenny *et al*., 2012), students exhibited a similar average level of computer and mobile devices skills. Successful participation in the assessment was accounted towards the final course grade. Student participation was voluntary.

2.2 Instruments

2.2.1 Usability

Among the many potential survey instruments that exist to measure usability, we have chosen the System Usability Scale (SUS) (Brooke, 1996) because of its versatility (Kortum and Sorber, 2015), validity and reliability (Bangor *et al.,* 2008), wide acceptance and easy administration.

The questionnaire consists of 10 items that are answered using a 5-step Likert scale ranging from “strongly disagree” to “strongly agree”, resulting in a single score between 0 and 100 (in 2.5 points increments) where higher scores indicate better usability. Table 1 depicts the 10 items of the SUS:

Table 1. *System Usability Scale (SUS)*

a/a	Item
1	I think that I would like to use this system frequently.
2	I found the system unnecessarily complex.
3	I thought the system was easy to use
4	I think that I would need the support of a technical person to be able to use this system
5	I found the various functions in this system were well integrated
6	I thought there was too much inconsistency in this system
7	I would imagine that most people would learn to use this system very quickly
8	I found the system very cumbersome to use
9	I felt very confident using the system
10	I needed to learn a lot of things before I could get going with this system

2.2.2 User Experience

A widely used tool for usability assessment is the User Experience Questionnaire (UEQ) (Laugwitz *et al.,* 2008). It assess the comprehensive impression of user experience in a convenient and quick way. It has also been shown to be a reliable and psychometrically validated instrument (Laugwitz *et al.,* 2008). It contains 6 scales with 26 items (pair of words: an adjective and its antonym) to determine the following aspects of the system, as table 2 demonstrates. Users evaluated their preference between each pair of words using a 7-point scale. Efficiency, perspicuity and dependability are often referred as pragmatic quality aspects (goal-directed), while stimulation and novelty are called hedonic quality aspects (not goal-directed) (Schrepp, 2018).

Table 2. User Experience Questionnaire (UEQ)

Scale	Subscales
Attractiveness	annoying / enjoyable, good /bad, unlikable / pleasing, unpleasant / pleasant, attractive/ unattractive, friendly / unfriendly
Perspicuity	not understandable / understandable, easy to learn /difficult to learn, complicated / easy, clear / confusing
Efficiency	fast / slow, inefficient / efficient, impractical / practical, organized / cluttered
Dependability	unpredictable / predictable, obstructive / supportive, secure / not secure, meets expectations / does not meet expectations
Stimulation	valuable / inferior, boring / exiting, not interesting / interesting, motivating / demotivating
Novelty	creative / dull, inventive / conventional, usual / leading edge, conservative / innovative

2.2.3 Research Design

The participants were randomly assigned into two groups. The control group (55 students) participated in the assessment using the desktop computers in the University Computing Center and the experimental group (55 students) participated in the assessment using their mobile devices. The assessment comprised 30 multiple choice questions, assessing factual knowledge from both the theory module (general ICT concepts) and practice module (use of office productivity software) of the course. The interfaces of both the computer-based assessment (CBA) and the mobile-based assessment (MBA) were very simple. The user had to log into the system. Each question (along with its four possible answer choices and the “OK” and “Next” buttons) was presented in a separate computer or mobile device screen. The student had to choose the right answer, confirm his choice with the “OK” button and then move to the next question by pressing the “Next” button. The text was in the native language of the students and the duration was 30 min. The maximum score, if all questions were answered correctly was 30. The interface was kept as simple as possible to avoid possible destructions and the aesthetics of both delivery modes were kept identical. After the assessment, students were asked to fill in the System Usability Scale (SUS) and User Experience Questionnaire (UEQ).

3. Results

3.1 Usability

The resulting SUS scores for the control group (CBA) was 76.0 and for the experimental group (MBA) was 78.2. Based on the literature (Sauro and Lewis, 2016; Sauro, 2011), the average SUS score across a large number of previous studies (more than 500) is 68. A SUS score above a 68 would be considered above average and anything below 68 is below average. The letter grade for the CBA SUS score of 76 is B and the letter grade for the SUS score of MBA is B+ (Sauro and Lewis, 2016). Based on the study findings, the perceived usability of both the computer-based and mobile-based assessments are considered high. Table 3 shows the descriptive statistics.

Table 3. SUS - Descriptive statistics for the CBA and MBA

mode	N	Mean	Std. Deviation	Std. Error
CBA	55	76.0	7.29	0.98
MBA	55	78.2	8.13	1.09

The overall score for the System Usability was not found to significantly differ between the two conditions. Independent t-test showed that the difference between the CBA and MBA conditions was not statistically significant ($t = -1.51$, $df = 108$, $p > 0.05$)

3.2 User Experience

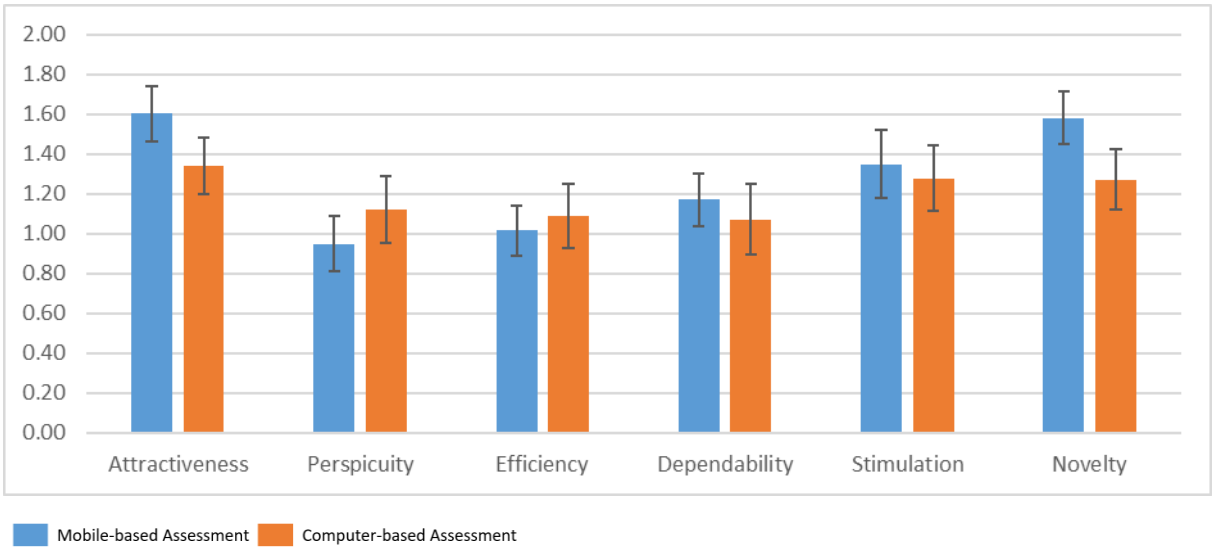
The analysis of the User Experience Questionnaire was based on the Excel tool downloaded from www.ueq-online.org. Table 4 shows the descriptive statistics for the User Experience Questionnaire.

Table 4. User Experience Questionnaire - Descriptive statistics for the MBA and CBA

	Mobile-based assessment (MBA)						Computer-based Assessment (CBA)					
mode	N	Mean	Sdt	Conf.	Conf. interval		N	Mean	Sdt	Conf.	Conf. interval	
Attractiveness	55	1.60	0.53	0.14	1.46	1.74	55	1.34	0.53	0.14	1.20	1.48
Perspicuity	55	0.95	0.52	0.14	0.81	1.09	55	1.12	0.64	0.17	0.95	1.29
Efficiency	55	1.02	0.48	0.13	0.89	1.14	55	1.09	0.61	0.16	0.93	1.25
Dependability	55	1.17	0.50	0.13	1.04	1.31	55	1.07	0.67	0.18	0.89	1.25
Stimulation	55	1.35	0.64	0.17	1.18	1.52	55	1.28	0.62	0.16	1.11	1.44
Novelty	55	1.58	0.50	0.13	1.45	1.71	55	1.27	0.58	0.15	1.12	1.42

Figure 1 shows the results for the six scales of the UEQ for the two assessment delivery modes.

Figure 1. User Experience Questionnaire for the CBA and MBA



Independent t-tests showed no statistically significant differences for the perspicuity ($p = 0.122$), efficiency ($p = 0.490$), dependability ($p = 0.380$) and stimulation ($p = 0.547$) scales between the CBA and MBA conditions. However, significant statistical differences were found for the attractiveness ($p = 0.011$) and novelty ($p = 0.003$) scales.

4. Conclusions

This study is aiming at comparing the overall usability and user experience between computer-based and mobile-based assessment in the context of higher education.

4.1 Usability

Regarding the System Usability Scale, the results showed that the usability scores of both computer-based and mobile-based assessments were high. Also, no significant difference between the computer- and mobile-based assessments in the overall scores of the usability was found. The high scores in overall usability of the desktop computers and mobile devices can lead to an interchangeable use of these devices in certain assessment tasks. The results agree with previous research on usability ratings between mobile and computer methods (Proaps *et al.,* 2014). Although SUS is basically a unidimensional measure, recent analyses suggest that it could be considered as a bi-dimensional measure also, with factors associated with the constructs of usability (items 1-3, 5-9) and learnability (Items 4, 10). Taking also into consideration the distinguished subscale of learnability, no difference was found between the two systems since both of them are considered easy to learn by the students.

4.2 User Experience

Regarding the User Experience Questionnaire, the results showed a significant difference in the attractiveness and novelty scales of the UEQ. The overall student impression of the mobile-based assessment was higher compared to that of the computer-based assessment. Students found the use of mobile devices for the assessment more attractive, enjoyable and pleasing. This is in-line with previous research that provides evidence for the positive student attitudes towards the use of mobile devices in assessments (Authors, 2018). Also, students found the mobile-based assessment more innovative and creative capturing more their attention. This may be due to the fact that students, even they self-reported that they were not new in using mobile devices for education purposes, they had rather limited experience in using their mobile for assessment purposes. No significant difference for the scales of perspicuity, efficiency, dependability and stimulation were found. Students found both the computer-based and mobile-based assessments easy to get familiar to, understand and use. They could easily navigate, without unnecessary effort and their interaction with both systems was efficient and fast. Students felt in control of the interaction and also felt confident when working with the computers or the mobile devices. They also found both systems to be enjoyable to use. It worth noted that, even no significant differences existed, stimulation was higher for the mobile-based assessment while perspicuity and efficiency was higher for the computer-based assessment.

4.3 Discussions

The measurement of the usability and the investigation of user experiences in technology-supported learning environments are very important for the improvement of the quality and effectiveness of computer-mediated instruction (Crowther, 2004). Moreover, they are needed in order to know what students think about the systems and therefore are important determinants for the technology adoption. Despite the fact that the capabilities of smartphones technologies have improved dramatically, their use in education in country x is rather limited and sometimes it is banned due to organizational, privacy and security issues (Economides and Grousopoulou, 2010). Based on the study findings, the usability of mobile devices, when used in assessments with closed type questions, is perceived to be equivalent to that of desktop computers. Therefore, the current study can inform educators and practitioners that mobile devices can, for certain assessment types, replace desktop computers, eliminating thus the need for dedicated computers laboratories.

However, for a successful adoption of mobile devices in assessment tasks, further evidence is needed regarding the equivalence in exam performance between computer and mobile-based assessments (Authors, 2017). Moreover, more research is needed to investigate the effect of age and gender on the resulting perceived usability and user experience. Also, other factors, such as the learners' context and different assessment tasks (question types), that could influence the interaction with the assessment delivery devices should be considered.

References

Authors (2018).

Authors (2017).

Authors (2016).

Bangor, A., Kortum, P.T. and Miller, J.T. (2008), “An Empirical Evaluation of the System Usability Scale”, *International Journal of Human-Computer Interaction*, Vol. 24, No.6, pp.574-594.

Bartram, D. (2008), “Available guidelines and standards for psychometric tests and test users”. In Towards a research agenda on computer-based assessment. Challenges and needs for European educational measurement (F. Scheuermann, & A.G.Pereira eds.), pp. 37-48, Luxembourg: European Communities.

Brooke, J. (1996), “SUS: A ‘quick and dirty’ usability scale” in *Usability evaluation in industry* (P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland eds.), pp. 189–194, London/UK: Taylor and Francis.

Coursaris, C.K. and Kim, D.J. (2011), “A meta-analytical review of empirical mobile usability studies”, *Journal of Usability Studies*, vol. 6, pp. 117–171.

Crowther, M.S., Keller, C.C. and Waddoups**,** G.L. (2004), **“**Improving the quality and effectiveness of computer mediated instruction through usability evaluations”, *British Journal of Educational Technology,* Vol. 35, No. 3, pp. 289–303.

Davis, F.D. (1989), “Perceived usefulness, perceived ease of use and user acceptance of information technology”, MIS Quarterly, Vol. 13, No. 3, pp. 319-340.

Deutsch, T., Herrmann, K., Frese, T. and Sandholzer, H. (2012), “Implementing computer-based assessment - A webbased mock examination changes attitudes”, *Computers & Education,* Vol. 58, No. 4.

Economides, A.A. and Grousopoulou, A. (2010), “Mobiles in education: Students' usage, preferences and desires”, *International Journal of Mobile Learning and Organisation*, Vol. 4, No. 3, pp. 235-252.

ISO 9241-11:2018, “Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts”. Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en

ISO 9241-210:2010, “Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems”. Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-1:v1:en

Kenny, R.F., Neste-Kenny, J.M.V., Burton, P.A., Park, C.L. and Qayyum, A. (2012), “Using self-efficacy to assess the readiness of nursing educators and students for mobile learning”, *The International Review of Research in Open and Distance Learning*, Vol.13, No. 3, pp.277–296.

Kortum, P. and Sorber, M. (2015), “Measuring the Usability of Mobile Applications for Phones and Tablets”, *International Journal of Human-Computer Interaction*, Vol. 31, No. 8, pp. 518-529.

Kumar, B.A. and Mohite, P. (2018), “Usability of mobile learning applications: a systematic literature review”, Journal Computer Education, Vol. 5, No. 1, pp. 1–17.

Laugwitz, B., Held, T. and Schrepp, M. (2008), “Construction and evaluation of a user experience questionnaire” in *USAB 2008, LNCS* (A. Holzinger, eds.), vol. 5298, pp. 63–76. Springer, Heidelberg.

Lucas, M. (2018), “External barriers affecting the successful implementation of mobile educational interventions”, *Computers in Human Behavior*, pp. 1-7.

Proaps, A. B., Landers, R.N., Reddock, C.M., Cavanaugh, K.J. and Kantrowitz, T. M. (2014), “Mobile and Computer-based Talent Assessments: Implications of Workload and Usability”, Computer-Human Interaction, 978-1-4503-2474-8/14/04, ACM, Toronto, Canada.

Sauro, J. and Lewis, J. (2016), “Quantifying the user experience: Practical statistics for user research”. Amsterdam; Waltham, MA: Elsevier/Morgan Kaufmann.

Sauro, J. (2011), “SUStisfied? Little-known System Usability Scale facts. User Experience”, *The Magazine of the User Experience Professionals Association*, Vol. 10, No.3. Retrieved from http://uxpamagazine.org/sustified/

Schrepp, M. (2018), “User Experience Questionnaire Handbook version 6 (16.09.2018)”. Retrieved from www.ueq-online.org

Terzis, V. and Economides, A.A. (2011), “The acceptance and use of computer based assessment”, *Computers & Education*, Vol. 56, No.4, pp. 1032-1044.