A comparative study between a computer-based and a mobile-based assessment: Usability and User experience

Stavros A. Nikou, University of Strathclyde, UK
Anastasios A. Economides, University of Macedonia, Greece

Abstract

Purpose - The purpose of the current study is to compare the overall usability and user experience of desktop computers and mobile-devices when used in a summative assessment in the context of a higher education 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 Computer-Based Assessment (CBA) and Mobile-Based Assessment (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 differ between the two conditions.

Practical implications – The usability and user experience issues discussed in this study can inform educators and policy makers for the potential of using mobile devices in online assessment practices, as an alternative to desktop computers.

Originality/value – The study is novel in that it provides quantitative evidence for the usability and user experience of both desktop computers and mobile devices when used in a summative assessment in the context of a higher education course. Study findings can contribute towards the interchangeable usage of desktop computers and mobile devices in assessment practices in higher 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 both instructors and students. Instructors can take advantage of real-time feedback, automated score processing and analysis, improved security, cost and time reduction (Terzis and Economides, 2011). Students experience a positive impact on their learning attitudes, motivation and performance (Authors, 2016). Moreover, mobile-based testing and assessment can be implemented anytime and anywhere, eliminate the need for specialized computer labs, facilitate formative, self- and peer- assessment and offer extended capabilities such as personalization, context-awareness and ubiquity (Authors, 2017). In the context of Bring-Your Own Devices (BYOD) policies, mobile devices can be used alternatively to desktop computers in testing and assessment, minimizing seating time and associative costs.

However, the introduction of 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). Another critical factor for its successful implementation is its adoption by users. According to the Technology Acceptance Model (TAM) (Davis, 1989), an information system can be adopted by users (behavioral intention to use) when the system is considered easy to use (high perceived ease of use) and useful (high perceived usefulness).

Usability and user experience are two main determinants for successful adoption of any information system. Usability is “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 “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).

The current study aims at providing a better understanding of the usability and user experience of desktop computers and mobile devices when used in summative assessment in the context of a higher education course.

The study is organized as follows. The literature review provides a brief overview of comparison studies on usability and user experience for computer and mobile devices. The Methodology section describes the participants, instruments used and procedures. Data analysis section follows with results. Discussions and conclusions section with implications, limitations and future research follows.

2. Literature Review

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). From the other side, usability evaluation of mobile learning applications is an active area of research (Kumar and Mohite, 2018) and it is strongly associated with educational effectiveness and institutional adoption (Vavoula and Sharples, 2009). However, 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 and user experience metrics, with more studies to provide mostly qualitative results.

Early research has shown that students were highly motivated and enjoyed using a mobile application for testing compared to web-based assessment systems (Romero, Ventura and DeBra, 2008). However, studies have reported numerous limitations when designing content for mobile devices (e.g. screen size, scrolling, typing or entering inputs) compared with desktop computers (Guler, Kilic, and Cavus, 2014). These limitations have an impact on system usability and mobile user experience.

For example, studies have reported a significant effect of screen size on usability metrics (effectiveness and efficiency) and perceived usability (Raptis, Tselios, Kjeldskov, and Skov, 2013). Also, users of desktop computer exhibited different web search behavior (search strategies, information seeking, saving results tasks, viewing and clicking) compared to mobile users (Ong, Jarvelin, Sanderson and Scholer, 2017). Web searching and reading may be more difficult to mobile devices comparing to desktop computers (Findlater and McGrenere, 2008).

However, other studies suggested that mobile characteristics negatively affect users' satisfaction only when the task complexity increases (Chae and Kim, 2004). For example, e-books for mobile phones were highly welcomed by the learners (Bidaki, Sanati, and Ghannad, 2013). Gaming with smartphones is preferred due to their touchscreen and portability (Adepu and Adler, 2016). Students who used mobile devices in collaborative environments vs. desktops self-reported a more positive effect in their flow experience than the students that used desktops (Abrantes and Gouveia, 2010). Recent studies have shown that applications on smartphones were judged to be more usable than applications on other platforms (Kortum, and Sorber, 2015). The portability and ease of use mobile characteristics highly influence smartphone usage over desktop computers (Adepu and Adler, 2016).

Usability and user experience research on mobiles vs. desktop computers seem to provide contradictory results depending each time on the specific context of use. Given the widespread adoption of mobile learning and the growing popularity of BYOD policies, an important question is whether mobile devices can provide an equivalent to desktop computers user experience in assessment procedures as well.

The current study aims at comparing computer- and mobile-based assessments on usability and user experience metrics. To achieve this, the following research questions were addressed:

1. Is there any difference in the overall usability between desktop computers and mobile devices in the context of a summative assessment in a higher education course?
2. Is there any difference in the user experience between desktop computers and mobile-devices in the context of a summative assessment in a higher education course?

3. Methodology

3.1 The computer and mobile-based versions of the assessment

In order to answer the aforementioned research questions, we developed a computer-based and mobile-based version of the same assessment.

The study focuses on the overall usability and the total experience in interacting with aspects of the summative assessment, rather than on the user interfaces (pages, buttons, fonts, visuals). User interfaces were kept similar between the two modes (with only the screen sizes to differ). The aesthetics and layouts of both versions, as Figure 1 shows, were kept as simple as possible to avoid possible destructions. However, due to the different delivery conditions, different interactions needed each time to navigate through the two assessment modes (e.g. tapping on the touchscreen instead of clicking the mouse, or holding the mobile device instead of sitting in front of a desktop computer monitor). Both versions of the assessment were developed using the jQuery framework for the user interface and PHP and MySQL for the server backend support.

Figure 1. The computer and mobile-based tests

-here -

The user had to log into the system. Each question was presented in a separate computer or mobile device screen along with its four possible answer choices and the “OK” button. The text was in Greek. The student had to choose the right answer, confirm his choice with the “OK” button and move to the next question. 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.

3.2 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 Greek University. The average age of the students was 20.2 (SD=1.990. All students exhibited a similar average level of computer and mobile devices skills and experience, based on a questionnaire about computer efficacy and mobile-device efficacy [12], that they were asked to fill in in advance. Student participation was voluntary. The majority of the students had already used computers and mobile devices for their own personal study. However, previous use of computer- or mobile devices in summative testing were rather limited in the context of their higher education studies, with the exceptions of a rather sparse usage in filling on-line surveys or questionnaires.

3.3 Procedures

The participants were randomly assigned into two groups. The control group (55 students) participated in the assessment using the desktop computers located in the University Computing Center and the experimental group (55 students) participated in the assessment in a lecture hall with Wi-Fi support, using their mobile devices. The monitor screen sizes ranged from 19 to 21 inches and the mobile devices screen sizes ranged from 3.5 to 5.3 inches. The 30 minutes summative assessment was delivered at the same time to both groups. Proctors invigilated and supervised both conditions. Successful participation in the assessment was accounted towards the final course grade. After the assessment, participating students were asked to fill in the System Usability Scale (SUS) and User Experience Questionnaire (UEQ).

3.4 Instruments

3.4.1 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 was originally created by Brooke at Digital Equipment Corporation, and it allows to evaluate a wide variety of products and services, including hardware, software, mobile devices, websites and applications.

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)*

-here-

3.4.2 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)

-here-

4. Data Analysis and Results

4.1 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. 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 ($t = -1.51$, $df = 108$, $p > 0.05$)

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

-here-

Brooke (1996) suggested that SUS is a unidimensional instrument and questionnaire items better not considered individually. However, to highlight the different perspectives of the questionnaire, Table 4 shows the responses on the individual items of the questionnaire with the mean (raw score), standard deviation and adjusted score for each.

Table 4. SUS scores for the computer and mobile-based assessments

-here-

4.2 User Experience

The analysis of the User Experience Questionnaire was based on the Excel tool downloaded from www.ueq-online.org. Table 5 shows the descriptive statistics for the User Experience Questionnaire.

Table 5. User Experience Questionnaire - Descriptive statistics for the MBA and CBA

Figure 2 shows the results for the six scales of the UEQ for the two assessment delivery modes.

Figure 2. User Experience Questionnaire for the CBA and MBA

-here-

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.

5. Discussions and Conclusions

The current study aims at comparing computer- and mobile-based assessments on usability and user experience metrics. The study is novel in that it provides quantitative evidence for usability and user experience metrics of using desktop computers and mobile devices when used in a summative assessment in the context of a higher education course.

Regarding the System Usability Scale, the results showed that the usability scores of both computer-based and mobile-based assessments were high. In addition, no significant difference between the computer- and mobile-based assessments in the overall scores of the usability was found. The results agree with previous research on usability ratings between mobile and computer methods (Proaps *et al.,* 2014). The similar scores in overall usability for the desktop computers and mobile devices may suggest an interchangeable use of these devices in certain assessment tasks. This may lead to minimize computer seat time (and the associated costs) since on-line assessments can be implemented in regular lecture rooms with Wi-Fi support, instead of dedicated computer rooms. This is also in-line with the emerging Bring Your Own Device (BYOD) policies (Johnson et al., 2016). 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, both systems were considered easy to learn by the students. Students found both systems easy to use (item 3) with well integrated functions (item 5) and felt confident with both systems (item 9) without the need for considerable technical support (item 4). However, their intention to use both systems were not as high as it would expected (item 1). This may be due to the assessment nature of both systems though.

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 when compared to that of the computer-based assessment. Students found the use of mobile devices for the assessment more attractive, enjoyable and pleasing. The findings agree with previous research that provided evidence for the positive student attitudes towards the use of mobile devices in assessments (Authors, 2018). In addition, students found the mobile-based assessment more innovative and creative capturing more their attention. This may be due to the fact that, even students had some previous experience in using digital devices in education they had rather limited experience in using mobile devices 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. Screen size did not seem to be an issue for the specific assessment task. As previous research suggests, not all tasks may benefit from larger screen displays (Raptis, Tselios, Kjeldskov, and Skov, 2013). Students felt in control of the interaction and 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.

The usability and user experiences in technology-supported learning environments are needed in order to know what students think about the systems and therefore are considerable determinants for the technology adoption. Moreover, they are significant important for the improvement of the quality and effectiveness of computer-mediated instruction (Crowther, 2004).

Despite the fact that the capabilities of smartphones technologies have improved dramatically, their use in education in Greece is rather limited. This is due to organizational, privacy and securrity issues (Economides and Grousopoulou, 2010) or even usability issues (Papadakis, Kalogiannakis, Sifaki, and Vidakis, 2018). Based on the study findings, the usability of mobile devices, when used in summative assessments with closed-type questions, was perceived to be equivalent with that of desktop computers. The same holds for the user experience dimensions such as perspicuity, efficiency, dependability and stimulation. Therefore, from the usability and user experience perspectives, the study findings provide some empirical evidence that mobile devices can possibly replace or complement desktop computers for on-line summative assessments in higher education.

However, one of the limitations of the study is that the assessment consisted only of closed-type questions. Different question types and assessment tasks, that may require different interaction types, should be considered in a future research. Moreover, more research is needed to investigate the effect of age and gender on perceived usability and user experience. Another important issue that needs further investigation is the equivalence in exam performance between computer and mobile-based assessments. Previous findings are inconsistent about student scores in computer and mobile-based assessments (Authors, 2013). Further evidence is needed regarding the equivalence of computer and mobile-based assessments modes in terms of student achievement (Authors, 2017).

Despite its aforementioned limitations, the study provided some initial evidence, that at least from the perspectives of usability and user experience, mobile devices could be an alternative to desktop computers in BYOD assessments in higher education.

References

Authors (2018).

Authors (2017).

Authors (2016).

Abrantes, S. and Gouveia, L.B. (2010). "A study on the usage of mobile devices in collaborative environments vs desktops: An approach based on flow experience," *International Conference on e-Business (ICE-B)*, Athens, pp. 1-4.

Adepu, S and Adler, R.F. (2016). "A comparison of performance and preference on mobile devices vs. desktop computers," *IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, NY, pp. 1-7.

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.

Bidaki, M. Z., Sanati, A. R., and Ghannad, F. R. (2013). Producing and introducing mobile books, as a new model of providing learning content in medical sciences. Procedia – Social and Behavioral Sciences, 83, 99–102.

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.

Chae, M., and Kim, J. (2004). “Do size and structure matter to mobile users? An empirical study of the effects of screen size, information structure, and task complexity on user activities with standard web phones”, *Behaviour and Information Technology*, Vol.23, No.3, pp.165-181.

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.

Findlater, L., and McGrenere, J. (2008). “Impact of screen size on performance, awareness, and user satisfaction with adaptive graphical user interfaces”. Proceedings CHI 2008, ACM Press, pp. 1247-1256.

Guler, C., Kilic, E., and Cavus, H. (2014). “A comparison of difficulties in instructional design processes: Mobile vs. desktop”, *Computers in Human Behavior*, Vol. 39, pp. 128–135.

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

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *NMC horizon report: 2016 higher education edition*. Austin, Texas: The New Media Consortium.

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”, *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.

Ong, K., Jarvelin, K., Sanderson, M., and Scholer, F. (2017). “Using information scent to understand mobile and desktop web search behaviour”, *Proceedings SIGIR’17*, Shinjuku, Tokyo, Japan

Papadakis S., Kalogiannakis M., Sifaki E., and Vidakis N. (2018). “Access Moodle Using Smart Mobile Phones. A Case Study in a Greek University”. In: Brooks A., Brooks E., Vidakis N. (eds) Interactivity, Game Creation, Design, Learning, and Innovation. ArtsIT 2017, DLI 2017. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 229. Springer, Cham.

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.

Raptis, D., Tselios, N., Kjeldskov, J. and Skov, M. (2013). "Does size matter? Investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency", Proceedings of the 15th international conference on Human computer interaction with mobile devices and services, pp. 127-136, Munich, Germany.

Romero, C., Ventura, S., and De Bra, P. (2009). “Using Mobile and Web-Based Computerized Tests to Evaluate University Students”, *Computer Applications in Engineering Education*, Vol.17, No.4, 435–447.

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.

Vavoula, G. and Mike Sharples, M. (2009). “Meeting the Challenges in Evaluating Mobile Learning: a 3-level Evaluation Framework”, *International journal of mobile and blended learning*, Vol. 1, No. 2, pp.54-75.