Children and Youth Services Review 119 (2020) 105535

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth

Perceived usability evaluation of Microsoft Teams as an online learning platform during COVID-19 using system usability scale and technology acceptance model in India

Debajyoti Pal*, Vajirasak Vanijja

School of IT, King Mongkut's University of Technology Thonburi, Bangkok, Thailand

ARTICLE INFO

Keywords:
COVID-19
Online learning
Perceived usability
System Usability Scale (SUS)
Technology Acceptance Model (TAM)

ABSTRACT

The COVID-19 pandemic has resulted in a physical shutdown of all types of educational institutes worldwide due to which the education delivery has now shifted to an "*online only*" exclusivity model. In this perspective, perceived usability of the online learning platforms that are currently being used is an important aspect, especially due to the absence of any physical classes. In this work Microsoft Teams is used as the reference platform for which the perceived usability is evaluated. For the evaluation purpose a dual strategy is followed by using the System Usability Scale (SUS), which is a Human Computer Interaction (HCI) based approach, and the Technology Acceptance Model (TAM), which is an Information Systems (IS) based approach. Although both these instruments are popular in their respective domains, yet they have not been considered simultaneously in one work for the purpose of usability evaluation. By doing so, this work attempts to streamline and unify the process of usability evaluation. Results that are obtained from a large-scale survey of university students show the similarity and equivalency between the two methodologies, with the Perceived Ease of Use (PEOU) construct of TAM having greater similarity with SUS. Moreover, this work also considers the digital-divide aspect (mobile vs. web environment) that is prevalent particularly in developing countries like India, and whether it has any effect on the perceived usability. Results show that the consumption platform does not have any effect on the usability aspect.

1. Introduction

Globally, the teaching-learning process is evolving rapidly from a traditional classroom environment to a mixture of traditional plus online learning. Cheap and widespread availability of devices such as smartphones and laptops, together with various applications like YouTube, Facebook, WhatsApp, etc. is changing the way people live, communicate with each other, and even their educational habits (Tiyar & Khoshsima, 2015). Although the concept of online learning is not new, especially after the popularity and widespread success of the MOOC's (Massive Open Online Courses) like Coursera, EdX, and Udemy among others, yet they have never been used as the primary mode of teaching (Hidalgo, Abril, & Parra, 2020; Rahmi, Aldraiweesh, Yahaya, Cumin, & Zeki, 2019). The onset of the COVID-19 pandemic has seriously affected all types of learning institutes globally starting from kindergartens, schools, colleges to the centers of higher education. To curb the spread of the virus by flattening the "*growth curve*" strict measures of social distancing are in place globally, which has in turn resulted in the physical closure of all types of learning institutes. This has forced the educational institutes worldwide to resort to an "*online only*" exclusivity model for the purpose of teaching and learning.

The concept of online learning or e-learning is not new. Availability of cheap and high-speed Internet access, together with the advances made in cloud technologies have helped to promote the flexibility of the learning procedure and supplement it well to the conventional learning methods (Wang, Lew, Lau, & Leow, 2019). Results from extant literatures show that the primary objective of online learning is not only to improve the reach and access of education to the general mass, but also improve the quality of learning along with reducing the cost of education delivery (Hamidi & Chavoshi, 2018; Panigrahi, Srivastava, & Sharma, 2018). Overall, online learning can be beneficial for the students, as they can learn from anywhere and anytime and at their own pace. The perception of the students towards online learning is mostly positive, as evident from the existing works (Alqurashi, 2019; Arias, Naffah, & Hernandez, 2019; Rodrigues, Almeida, Figueiredo, & Lopes, 2019). However, before the onset of COVID-19, the use of various

* Corresponding author at: School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand.
E-mail address: debajyoti.pal@mail.kmutt.ac.th (D. Pal).

https://doi.org/10.1016/j.childyouth.2020.105535
Received 24 July 2020; Received in revised form 25 September 2020; Accepted 27 September 2020
Available online 01 October 2020
0190-7409/

online learning platforms and resources were mainly supplemental in nature, over and above the regular classroom teaching imparted at the schools and universities. Therefore, the COVID-19 scenario has brought forward an unprecedented situation, where there has been a radical change in the mode of education delivery to be strictly "*online only*". Teachers have been compelled to deliver lectures online using some form of an online delivery platform (Abidah, Hidaayatullah, Simamora, Fehabutar, & Mutakinati, 2020). This digital disruption in the education sector has been sudden, untimely and somewhat unorthodox as there has been absolutely no strategic planning or preparation with regards to its implementation worldwide. Therefore, in this chaotic situation with the closure of the educational institutes, one of the very basic necessities is to ensure a decent perceived usability of the online delivery platforms that are being used for the purpose of education delivery, particularly from the students' perspective. Perceived usability is one of the fundamental components of user experience (UX) (Diefenbach, Kolb, & Hassenzahl, 2014). It is reasonable to anticipate that a good UX will lead to greater satisfaction levels, which in turn can guarantee success of the online education delivery platforms. In this work, Microsoft Teams is taken as the reference platform for online education delivery.

The students of today are not equally equipped with digital tools and equipments, and as such, there exists a digital divide (Iivari, Sharma, & Olkkonen, 2020). A number of extant researches have focused and examined on the digital divide that exists particularly related to the access of newer devices and technologies (Song, Wang, & Bergmann, 2020; Srivastava & Shainesh, 2015), however for the young student population the understanding is very limited (Iivari, Kinnula, Juustila, & Kuure, 2018; Mariën & Prodnik, 2014). In an Asian context, especially developing countries like India where the present study has been carried out, the digital divide can be very polarizing, especially in terms of the ownership of smartphones, tablet computers or laptops. The pandemic has further complicated matters, as now the students need to depend solely on the digital devices that they own i.e. smartphones, tablets or laptops for the purpose of their study. For example, a very recent research by authors in Kapasia et al. (2020) during the COVID-19 lockdown period in India and its effect on students of higher education shows that around 86% of the students use their smartphones (mostly Android) for attending the online classes, while a mere 14% use their laptops. Since online learning primarily involves watching multimedia contents either during the live tutoring sessions, or when preloaded video contents are uploaded by the teachers to the learning platform, the type of device used for this purpose may affect the usability and consequently the student's perception of the delivery platform. In fact, in the context of delivery of multimedia content there is enough evidence that the end user experience (often described in terms of the QoE – Quality of Experience) depends on the screen size (Maniar, Bennett, & Gal, 2007; Pal & Vanijja, 2017), and hence can vary between smartphones and laptops (or any computer in general). Moreover, now it has become a norm for the major application developers like Google, Microsoft or Apple to produce both mobile and web-versions of a particular application. The same is true for the online education delivery platforms like Microsoft Teams or Google Meet that have both the mobile and web versions for Android, iOS as well as the Windows environment. While portability and flexibility are some of the advantages of any mobile application, yet in the context of online education since on-screen interaction is an important aspect, hence there is no definitive conclusion as to whether the usability of the learning platform will vary depending on the mobile or the web-version of the applications. Thus, there exists a gap in extant literature specifically with respect to the usability evaluation of the online learning platforms in a mobile and web-based environment.

Another issue is with respect to the best way of measuring the usability of the mobile and web-based applications, so that the results obtained are conclusive and relevant. Current usability evaluation methodologies can be broadly classified into two parallel segments: a) the first approach is more common among the human–computer interaction/usability (HCI) researchers, and b) the second approach is used by the information systems (IS) researchers while studying the adoption of various systems. For the purpose of usability evaluation of various applications, the first approach is more common that uses standardized usability questionnaires intended for the purpose of usability testing by asking the users of an application to assess its usability after its thorough and detailed usage. This technique has been used for assessing the usability of various consumer products (Kortum & Bangor, 2013; Lewis, 2018), computer programming software (Kortum & Johnson, 2013), and even in an educational context (Abuhlfaia & Quincey, 2019; Harrati, Bouchrika, Tari, & Ladjailia, 2016). The System Usability Scale (SUS) (Brooke, 1996), Usability Metric for User Experience (UMUX) (Finstad, 2010), Post-Study System Usability Questionnaire (PSSUQ) (Lewis, 1995) and the After Scenario Questionnaire (ASQ) (Lewis, 1991) are some of the commonly used instruments belonging to the first category for measuring the usability of the applications. The second approach for usability evaluation is common among the information systems and market researchers focusing on the adoption of information systems using various questionnaires and theoretical frameworks. Among them, the Technology Acceptance Model (TAM) originally proposed by (Davis, 1989), has been the most influential one that introduces the concepts of perceived usefulness (PU) and perceived ease of use (PEOU) for measuring the end-user's intention of using a technology. Extant research has shown the appropriateness of TAM in explaining the end-user system usage (Wu, Chen, & Lin, 2007). Yet, there is often criticism about the lack of a statistical evidence with regards to the correlation between perceived usability (as measured by HCI researchers) and the primary components of TAM i.e. PU and PEOU (Tractinsky, 2018) for evaluating any system or application. Therefore, despite having similar objectives of measuring the usability, the two approaches are totally disjoint, and to the best of our knowledge there seems to be no relationship existing between these two approaches. The same is evident for the online learning context also, wherein for measuring the end-user acceptance, either a perceived usability HCI based approach is taken or a TAM (or some other similar) IS based theoretical approach is taken. The important research question is whether these two visually dissimilar approaches are actually related and consistent with each other? Keeping in mind the current COVID-19 scenario that has forced the global education delivery model to be "*online only*", we felt that there is an urgent need not only to evaluate the usability of the online learning platforms focusing on the digital-divide aspect, but to also unify the fragmented efforts undertaken by the HCI and IS research community for usability evaluation in order to create a simpler and more robust approach. Particularly, the following are the goals of the present work:

(a) To determine whether there is a difference in perceived usability while using the online learning applications based on the consumption platform (smartphones vs. laptops) or alternatively the mobile vs. the web version of the applications.
(b) To determine whether there is a relationship between the measures of perceived usability as done by the HCI researchers (SUS) and the IS researchers (TAM). Additionally, which TAM factor (PU or PEOU) has a stronger relation to the measures of perceived usability?

The remaining article is structured as follows. Section 2 introduces the related works. In Section 3 the methodology is discussed, followed by the result analysis in Section 4. Finally, the discussion and conclusion are provided in Section 5.

2. Related works

Evaluation of an online learning platform with regards to its usability as well as the overall experience obtained is an important issue, particularly in the present situation of COVID-19 due to the physical closure of all types of educational institutes. The literature review done in this section focuses of three main aspects: (a) the System Usability Scale (SUS), which is one of the most popular tools used by HCI researchers for the purpose of usability evaluation, (b) the Technology Acceptance Model (TAM) highlighting its similarity to the concept of perceived usability, and (c) the current state of usability evaluation in the online learning scenario involving SUS and TAM in particular.

2.1. The system usability scale (SUS)

SUS is one of the most popular instruments used for assessing the perceived usability, both in usability related studies and surveys (Lewis, 2014, 2018) by the HCI researchers. Extant research has shown that SUS has got a high degree of reliability (normally the Cronbach's alpha coefficient exceeds 0.90), validity, and can be adapted for different contexts (Peres, Pham, & Phillips, 2013). SUS has 10 items in total, with half of the items having a positive tone (the odd number items), and the other half having a negative tone (the even number items). The response is given on a scale of 1 (strongly disagree) to 5 (strongly agree) for each item. The SUS score ranges from 0 to 100 (higher score meaning a better usability) in steps of 2.5 increments. There are a number of reasons for the popularity of SUS, and hence, our decision to include it in this work. First, it is free to use and available in the public domain for a long period of time. Second, it has very good psychometric properties. Third, extensive normative research has been done on SUS, due to which there are several ways of interpreting it (Bangor, Kortum, & Miller, 2009; Sauro & Lewis, 2016). For example, (Bangor et al., 2009) proposed an absolute grading scale with A: > 89; B: 80–89; C: 70–79; D: 60 – 69; and F < 60. In this work, we interpret the SUS scores, as outlined by the authors in (Sauro & Lewis, 2016) by using a Curved Grading Scale (CGS) approach. The CGS grading scheme is considered to be robust as it is based on data obtained from 446 usability studies involving over 5000 completed SUS questionnaires. This scale provides a good way to empirically interpret the meaning of the SUS scores. Table 1 provides a description of this scale.

2.2. The technology acceptance model (TAM)

TAM is one of the most popular and widely used model by IS researchers for predicting the future use of a product or technology. PU and PEOU are the two core constructs of TAM that have 6 items each for the respective measurements. Similar to SUS, the response for the TAM items are also given on a scale of 1 (strongly disagree) to 5 (strongly agree).

In the TAM context, PU refers to the degree to which a person believes that using technology will improve his/her work performance (Davis, 1989). Similarly, PEOU is defined as the degree to which a person believes that using technology will be easy and free from any efforts (Davis, 1989). However, strictly speaking TAM postulates the concepts of PU and PEOU as pre-usage factors, i.e. before using any product or technology. The more a person believes in these factors, the greater will be the usage intention. The original results of TAM in (Davis, 1989) showed that both PU and PEOU are highly correlated and statistically significant with the self-anticipation of the users to likely use a product if it was available to them. The PU and PEOU sub-dimensions had a high reliability of 0.98 and 0.94 respectively, along with appropriate convergent and divergent validities. The 12 items related to PU and PEOU are given in Table 2 as per the original TAM version.

From the above discussion it is evident that TAM focuses on the usability aspect through the lens of PEOU and PU. In fact, (Tractinsky, 2018) concluded in his research about the close intuitive relation between the meaning of usability and "PEOU and PU", the two factors that are instrumental in explaining the behavioral intention. Although he proposed and introduced the conceptual similarity between the two different constructs of "perceived usability" and "PEOU and PU" that is commonly used among the HCI and IS research communities respectively, yet he did not provide any empirical evidence for the same. Moreover, for explaining the usage of any system, product or technology, researchers have used either the concepts of perceived usability (by using SUS, UMUX, etc.) or PEOU and PU (by using TAM or other theoretical models) in a mutually exclusive manner, making the efforts to usability evaluation rather fragmented. Therefore, there is a need to empirically validate the measures of perceived usability and that of PEOU and PU, and check which of them correlate well to each other. One of the objectives of this work is to fill up this research void.

However, since the core essence of TAM is to measure the likelihood of usage (and not the actual system usage), it does not align well to the current objectives of this study, which is to measure the PEOU and PU after using the system i.e. the online learning platform. Therefore, we have slightly re-worded the original TAM version (the details given in the methodology section) so that now it can measure the actual user experience (after using the product), rather than the likelihood of usage. As already explained previously, since the COVID-19 scenario has forced to shift the focus to an "*online only*" mode of education delivery, therefore, perceived usability is an important aspect and the students must be satisfied after using the online learning platform, for the teaching learning process to be effective. Therefore, we take two additional measures (ratings) into account (a) likelihood of recommendation, and (b) overall experience using the learning platform, and check whether these two measures can be predicted by either perceived usability or the TAM constructs.

2.3. Likelihood to recommend and overall experience

After using any system or product it is reasonable to focus on the overall experience obtained and consequently the likelihood of recommendation. Likelihood to recommend (LTR) is a popular way of measuring user satisfaction and loyalty and was originally proposed by Reichheld (2003). He called it the Net Promoter Score (NPS), and since then this model has been used in many circumstances for measuring the user satisfaction (Lee, 2018; Owen, 2018). One of the reasons behind the popularity of this methodology is it is very simple (contains only one question) that gives the organizations timely data (Reichheld, 2003). It addresses the complicated theme of UX and loyalty simply by one question that enables to address this issue easily, economically, and with good results. The most common form of this item is "*Considering everything, how likely will you recommend this product to a friend or colleague*?". This is a 11-point scale ranging from 0 (not likely at all) to 10 (extremely likely). The use of LTR is recommended specifically when users have a choice of which products to use i.e., they can choose among multiple options (Lewis, 2018; Sauro & Lewis, 2016). This makes it suitable for using in the present scenario due to the presence of several competing online platforms like Microsoft Teams, Google Classroom, Zoom, etc. that are being used for the purpose of delivering lectures. Moreover, extant research has shown that there is a high

Table 1
The curved grading scale used for interpreting SUS (Sauro & Lewis, 2016).

Range of SUS Score	Grading	Percentile Range
84.1–100	A+	96–100
80.8–84.0	A	90–95
78.9–80.7	A−	85–89
77.2–78.8	B+	80–84
74.1–77.1	B	70–79
72.6–74.0	B−	65–69
71.1–72.5	C+	60–64
65.0–71.0	C	41–59
62.7–64.9	C−	35–40
51.7–62.6	D	15–34
0.0–51.6	F	0–14

Table 2
The 12 Items corresponding to TAM (Davis, 1989).

Perceived Usefulness	Perceived Ease of Use
Using "XXXX" in my job would enable me to accomplish tasks more quickly	Learning to operate "XXXX" would be easy for me
Using "XXXX" would improve my job performance	I would find it easy to get "XXXX" to do what I want it to do
Using "XXXX" in my job would increase my productivity	My interaction with "XXXX" would be clear and understandable
Using "XXXX" would enhance my effectiveness on the job	I would find "XXXX" to be flexible to interact with
Using "XXXX" would make it easier to do my job	It would be easy for me to become skillful at using "XXXX"
I would find "XXXX" useful in my job	I would find "XXXX" easy to use

**XXXX = Substitute with "this product/technology" (depending on context).

correlation between SUS and LTR data (correlation coefficient = 0.62) collected from a large sample of 2201 users and over 80 products (Sauro & Lewis, 2016).

After using a product, the users form an opinion about their overall experience (OExp) with the product. Having a good experience is important for the long-term viability of the product. One of the most common and widely used ways of collecting such a perception is by using an item that is closely related and modelled on the LTR item, "*Considering everything, how would you rate your overall experience with this product*?". Like LTR, this item is also on a 11-point scale ranging from 0 (terrible) to 10 (excellent). Extant research has shown significant high correlations of the OExp item with LTR and other measures of perceived usability (Lewis, 2018).

2.4. Online learning evaluation models

Pertaining to the context of online learning, there are a few TAM related studies that focus on the usability aspect. Persico, Manca, and Pozzi (2014)), employed TAM in a Moodle based online learning platform for investigating the university students' perception of adopting an online learning environment. The evaluation is done based on usefulness, ease of use, and effectiveness. Close to the previous work, Rodriguez and Lozano (2012) use TAM to assess the usability of a Moodle platform in terms of the usefulness and ease of use related to the actual use of the system. TAM has also been used to measure the teacher's intention to use various tools related to online learning and found positive effects of ease of use on the teaching intention (Nikou & Economides, 2018). Scherer, Siddiq, and Teo (2015) did a large-scale analysis on 1190 teachers in Norway with regards to their perception using various information technologies while teaching online, and among other factors found usefulness to be a major one. An empirical examination of the adoption of web-based course tools in the higher education context is undertaken by Ngai, Poon, and Chan (2007) using a TAM based approach. They extended TAM with an additional factor of technical support and found that it is a dominant one along with PU and PEOU affecting the attitude of students towards the web-based tool usage. Boateng, Mbrokoh, Boateng, Senyo, and Ansong (2016) investigated the determinants of e-learning adoption among students of developing countries. Results reveal that both PU and PEOU affect the attitude and adoption of e-learning systems. A user-centric framework for e-learning technologies based on TAM with additional security, privacy and trust constructs is proposed by Baby and Kannammal (2020). They use a novel network path analysis algorithm for evaluating the adoption scenario and find PU and PEOU to measure the actual system usage. Yalcin and Kutlu (2019) examine the student's acceptance of and intention to use learning management systems for university education based on TAM. PU, PEOU, social norms and computer usage efficacy are found to affect the adoption scenario. Overall, there are several other studies all of which establish the relevance of TAM (particularly PEOU and PU) in the online learning context (Estriegana, Merodio, & Barchino, 2019; Gahtani, 2016; Millat, Lopez, Jover, Abad, & Alegret, 2018; Revythi & Tselios, 2019; Salloum, Alhamad, Al-Emran, Monem, & Shaalan, 2019). Apart from online learning, TAM has also been used to predict the adoption intention in various other contexts like smart-homes (Pal, Triyason, Funilkul, & Chutimaskul, 2018; Park, Kim, Kim, & Kwon, 2018), smart products (Dutot, Bhatiasevi, & Bellallahom, 2019; Pal, Arpnikanondt, Funilkul, & Chutimaskul, 2020), smart healthcare (Bettiga, Lamberti, & Lettieri, 2020; Kamal, Shafiq, & Kakria, 2020), music streaming services (Fernandes & Guerra, 2019; Pal & Triyason, 2018), and many more.

Apart from the TAM related studies as mentioned above, SUS has also been used by some researchers for evaluating the quality of the online learning systems. However, when compared to TAM, the number of such studies are limited. Harrati et al. (2016) conducted a study to apprehend how effective and usable e-learning systems are. They mainly focused on how lecturers interact with the e-learning system based on some pre-defined tasks. They concluded that SUS is not a sufficient measure to express the true acceptance and satisfaction level of lecturers for using the system. A combination of TAM and SUS is used by Revythi and Tselios (2019) for evaluating the acceptance of a learning management system. Instead of using the PEOU component of TAM, they replaced it with SUS for measuring the model. However, no details are provided as to how good SUS is a replacement for the PEOU component of TAM, and whether they measure the same thing. Pereira and Anamaria (2012) have used three different techniques for evaluating a Moodle based learning platform (including SUS) and reported that SUS is an effective tool for measuring the usability. However, they did not report any SUS usability score. Ayad and Rigas (2010) evaluated a gamified vs. non-gamified virtual classroom using the SUS instrument and concluded that the game-based platform had a better performance and perceived usability. Similar work has been done by Orfanou, Tselios, and Katsanos (2015), wherein they evaluated an online learning platform using SUS and validated the SUS questionnaire for their learning management system platform. A Moodle based learning management system is evaluated by Sami, Rutter, and Smith (2016), and an association between the SUS scores and variables like age, IT skills, prior experience with learning management systems and the usage frequency are taken into account. Results show a positive correlation between the SUS scores and the usage frequency of the learning management system. Researchers in Arain, Hussain, Rizvi, and Vighio (2016) developed a mobile learning application called DARSGAH and tested it with 100 university students using SUS, and found the average SUS score to be 84, indicating a good measure of perceived usability for their designed application. Hasibuan, Santoso, Yunita, and Rahmah (2019) used an instrument called the e-learning usability scale (EUS) that is modelled on SUS for evaluating an Indonesian e-learning system. Since, EUS was originally made in English, the authors developed an Indonesian version of the same. A high reliability score (Cronbach's alpha = 0.92) is reported that makes it suitable to be used by usability researchers. A learning management system is tested by Abuhlfaia and Quincey (2019) in a higher education setting specifically with respect to the user interface design, usability and the overall learning experience. SUS is used for measuring the usability that is followed by a thematic analysis from a random subset of data to delve into the details of the usability issues. A SUS score of 62.52 is obtained that is below the accepted standards indicating that usability evaluation in an e-learning context is a complicated scenario. A new personalized and adaptive e-learning system is evaluated by Shi, Awan, and Cristea (2013) using

SUS. An average SUS score of 75.75 is obtained along with a decent reliability (Cronbach' value of 0.85). A simulation-based learning system for international trade is assessed by Luo, Liu, Kuo, and Yuan (2014) using SUS by collecting opinion from the students and teachers over a course of two semesters. While the average SUS scores for the students are 62.01, those from the teachers are 74.45, indicating that the perceived usability is better for the teachers. Simoes and Moraes (2012) designed a Moodle based distance learning platform and tested it using two different methodologies: SUS and heuristic evaluation. Both the methods unveiled that their designed system had serious usability problems.

From the literature review the existing research gaps are evident. First, the notion of perceived usability in an educational context is mostly related to the evaluation of the Moodle based platforms that are customized by the researchers depending on their research objectives. However, it should be kept in mind that these types of Moodle based platforms can supplement the primary classroom teaching methods, but they do not have the capacity to replace them. For example, Zhang, Zhao, Zhou, and Nunamaker (2004) compared e-learning with classroom learning for finding out if e-learning can replace the more traditional form of teaching. They concluded that although e-learning is promising and beneficial in case of lifelong learning and training, it can complement classroom teaching and not replace it. Similar observations are made by Condie and Livingston (2007) where they advocate the use of blended learning tools such that online learning complements traditional classroom learning rather than replacing it. In a more recent study in 2020 (Jong et al., 2020) the authors give 12 tips regarding how to integrate MOOC videos as a part of regular classroom lectures to create a positive blended learning environment for improving the student learning experiences. Therefore, the challenges brought forward by the pandemic are unique and strict due to the physical closure of the educational institutes. In the absence of any form of physical classes it requires the online learning platform to be more dynamic and interactive than before that can ideally substitute the traditional classroom-based learning. Measuring the perceived usability of this type of a learning environment and scenario has not been attempted before and is therefore a challenge. Second, extant research does not stress on the digital divide aspect that might affect the quality of online learning. This is expected, since till date the use of online learning strategies by universities have been secondary for improving the student performance (Abrami, Bernard, Bures, Borokhovski, & Tamim, 2011). However, under the changed circumstances, wherein online learning has become the sole and primary means of education delivery, there is a need to investigate the effects of digital divide that might affect the quality of the learning process. Learning online will increase the screen viewing time and extant research has shown that user experience varies with not only the screen size of the device being used (Pal & Vanijja, 2017) but also on the quality of the used applications (Kortum & Sorber, 2015). Using SUS Kortum and Sorber (2015) evaluated the usability of certain applications for smartphones and tablets (smaller vs. bigger screen) and found out that smartphone applications have a greater usability score. Therefore, screen size as well as the nature of the applications might affect the overall learning experience. As such, in this work we have attempted to measure and compare the perceived usability based upon the consumption platform (smartphones vs. laptops) and check out for any differences. Third, the existing approaches towards measuring the perceived usability are absolutely distinct i.e. they either take a HCI based approach, or an IS based approach. Though the methodologies involved for these two distinctly separate domains are different, yet from many of the works that are discussed here it becomes evident that there is a relationship between these two otherwise dissimilar approaches, particularly if TAM is considered. For the online learning context, we were able to find only one work that uses TAM and SUS simultaneously (Scherer et al., 2015). In this work the authors tried to measure the intention to use a learning management system for which they used TAM; however, they dropped the PEOU construct, and replaced it with SUS instead. However, the authors did not investigate the aspect that how closely the factor structure of SUS is related to the PEOU construct of TAM. Moreover, they did not consider any effect or relation that PU might have with SUS. Therefore, in this work we make an attempt to unify these two parallel research segments by examining the relationship between perceived usability as measured by SUS and its correlation with PEOU and PU, which are the major components of TAM.

3. Methodology

3.1. Online learning platform

For the purpose of data collection Microsoft Teams is chosen to be the reference online learning platform. The decision behind selecting Microsoft Teams is done due to the following reasons. First, the universities had an enterprise version for this application for all its personnel (faculty and staff) and students through the "Office 365 Education" plan. Therefore, in case of any type of service-related issues instant support is available from the IT infrastructure team of the universities. Second, Microsoft Teams provide a good integrated teaching–learning space, offering a lot of features that are comparable, and in some cases even better than any Moodle based online learning platform. Although each of the 5 universities have their own learning management systems, yet they are not effective in handling the requirements of the pandemic. The learning management systems work in an asynchronous mode only wherein the course instructors can upload the course videos and other course contents that they desire to do. However, a real time teacher-student interaction is not possible using such systems, due to which they lack interactivity and personalization. Moreover, with none of the systems it is possible to conduct an online student examination of any format (either objective or subjective type). Until the pandemic classroom teaching had been the primary form of education delivery and the role of the learning management systems were secondary/optional in such a sense (depending on the course instructors). However, with the onset of COVID-19 due to the physical closure of universities an "*online only*" education delivery model had to be followed that the current learning management systems lacked mainly due to their asynchronous nature. Microsoft Teams provides an elegant solution in this aspect as not only it can be used as a learning management system by the instructors, but it supports both synchronous and asynchronous learning. For example, just like a physical classroom has got a specific schedule, using this application it is possible to take live online classes for multiple students at a pre-scheduled time. Moreover, if the course instructors want, they can even record their videos and upload those to the application for the students to view. Therefore, in essence this application works in a synchronous as well as an asynchronous delivery mode, depending upon the choice and preference of the students and the instructors. Third, just like any conventional e-learning system, Microsoft Teams has the option of file sharing. This feature can be used by the course instructors to upload any type of files (power-points, word or pdf documents, and video lectures) that they want the students to refer to. Similarly, quizzes and tests can be designed (both objective and subjective types) for the students to take, and the application can automatically grade the students based on some pre-defined rubric. Moreover, social networking features are also built into the Teams software by allowing personalized as well as group chat facilities. Therefore, this application not only provides a multi-party video-conferencing facility, but also a real classroom like virtual learning environment having a variety of features and available both in the mobile and web formats. Hence, Microsoft Teams closely supports the mission of "*online only*" learning. The closest competitor to Microsoft Teams is Google Classroom. However, the problem with Google Classroom is that it does not support the video-conferencing facility, for which the users must install the Google Meet application separately. Therefore, two different applications are required for simulating the whole learning environment if using the solution from Google, which can be inconvenient for the students. Microsoft Teams provides an elegant solution for this by incorporating all the features into a single application, and hence, used as the reference medium.

3.2. Data collection and procedure

The duration of the online classes lasted from 1st March 2020 to 30th June 2020 (around 3 months). After this period an online survey was conducted from 5th July 2020 to 12th July 2020 for collecting the information. The questionnaire was distributed using Google Forms. All the participants in the survey were students of science, engineering or management disciplines. Students of the science discipline were enrolled in one of the following three courses: differential equations, zoology I or ICT in education. For the engineering discipline students were enrolled in at least one of the following four courses: computational thinking, statistics for scientists, mathematics I, and database management systems. The management students were enrolled either in the business analytics or the operations management course. The list of the enrolled students in each of the courses were obtained from two sources: either from the instructors teaching the course or from the student registration department of the university. Obtaining this initial pool of potential student list was done online by emailing to the respective course instructors or the administrative staff due to the physical closure of the institutes. Initially those instructors or administrative staffs were contacted whom the authors knew personally, and they were further requested to contact relevant people they knew. Thus, there was a certain degree of snowballing effect for the selected sampling strategy. Following this procedure, a list of 1764 enrolled students were obtained belonging to 5 different colleges in India. Participation in the survey was purely voluntary and did not have any link with the final course evaluation. Before taking the actual survey, the participants had to first complete an Institutional Review Board (IRB) approved online consent form. The survey link to the Google Forms was either emailed to the students or shared through relevant channels of Microsoft Teams and WhatsApp. At the beginning of the survey the following instruction was given:

> *"Thank you for agreeing to take part in this evaluation of Microsoft Teams. It will take no longer than 15 to 20 min to complete the entire survey. In this survey, you will be given some questions to rate your experience using Microsoft Teams as a part of your course. The primary goal is to use this questionnaire and your learning experience with this application to obtain a general perception of its usability. Please note that this is not a test of you – you are helping us to understand your experience with this software. Please read and mark each item carefully as they differ in whether a high number or a low number indicates a good or poor user experience. Your first impression is just fine. Let's get started".*

For each of the participants the order of the entire questionnaire was random. This was done to minimize the chances of response bias. The participants accessed Microsoft Teams either using their smartphones (mobile-version) or laptops (web-version) and rated their experience accordingly. The questionnaire details for SUS and the revised TAM version are provided in Table 3. For SUS, the participants gave their ratings on a 5-point Likert scale (1 – *strongly disagree* to 5 – *strongly agree*). In case the participants did not have any opinion related to any specific item, they were instructed to select the middle point (score of 3), instead of leaving it blank. For SUS the participants were also made aware of the situation that half of the statements were positive and the other half negative, and therefore be careful of this scenario and provide their ratings accordingly. In case of TAM, a 7-point Likert scale rating was used (1 – *extremely disagree* to 7 – *extremely agree*). For the TAM questionnaire all the items had the same (positive) tone, and therefore, scores to the right of the scale would indicate a better user experience. In this case also, instead of leaving any question blank, the participants were asked to fill up the mid-value (score of 4), in case of any doubt to keep consistency with SUS.

After finishing the usability questionnaire there was a general section wherein the participants were asked to give their overall assessment of the learning platform. The LTR and OExp items that are previously discussed were presented in this part.

4. Result analysis

All the data analysis is done using SPSS version 17.0.

4.1. Participant demographics

To begin with, the participant demographics are provided in Table 4. Initially, the survey had been sent out to 1764 participants. However, some of the participants either did not respond back, or did not complete the full survey. For the purpose of final data analysis 1595 samples were retained, giving a response rate of close to 90%. The proportion of male and female students were roughly equal in our sample. Almost 64% of the students' age was 21 or below, with a median age of 21 years. Majority of them resided in an urban area (69.53%) and had previous experience using some form of online learning (84.6%). Around 61% of the participants accessed Microsoft Teams from their smartphones, whereas the remaining used the web version on their laptops.

4.2. Initial data preprocessing

In case of SUS the participants gave their ratings on a 5-point scale, whereas for TAM ratings were obtained on a 7-point scale. Since different scales are used therefore to begin with both the SUS and TAM ratings are converted to a uniform score ranging from 0 to 100. In case of SUS to begin with each item's score contribution is calculated, which ranges from 0 to 4. For the positively-worded items (1, 3, 5, 7 and 9), the score contribution is the scale position minus 1, whereas for the negatively-worded items (2, 4, 6, 8 and 10) the contribution is 5 minus the scale position. To get the overall SUS score the sum of the item score contributions is multiplied by 2.5. Therefore, SUS scores range from 0 to 100 in steps of 2.5 points. However, in case of TAM the score calculation is a little bit different as it involves all positively worded items. First, the mean of the item scores is calculated, then 1 is subtracted from the mean (for normalizing in the 0 to 6 range), and finally multiplied by 16.67. This generates a final value in the same range of 0 to 100 as SUS. For obtaining the overall TAM score, the mean of PU and PEOU is computed. This computation process normalizes the SUS and TAM scores in the same range that is used for the purpose of further analysis.

4.3. Reliability of the questionnaires

For measuring the reliability of each set of questionnaire the Cronbach's alpha value is calculated. Reliability is the extent to which an instrument will give the same results if the measurement is to be taken again under the same conditions. Extant research shows that for the questionnaire to be reliable the Cronbach's alpha values should be at least 0.70 (Hair, Anderson, Tatham, & Black, 1998). In the present case the Cronbach's alpha value obtained for SUS is ($\alpha = 0.90$) and that for the TAM version is ($\alpha = 0.88$). Specifically, the alpha values for PEOU and PU sub-items are 0.91 and 0.87 respectively. Therefore, both the questionnaires are reliable and have sufficient internal consistency.

4.4. Concurrent validity

Concurrent validity is another measure of checking the efficacy of the designed survey and can be used as an evidence for justifying the use of a particular questionnaire for predicting the outcomes. It is a type

Table 3
Questionnaire details used in the experiment.

Instrument	Items	Questionnaire Details
SUS [Adapted and Modified from 25]	SUS_{01}	I think that I would like to use this application frequently
	SUS_{02}	I found the application unnecessarily complex
	SUS_{03}	I thought the application was easy to use
	SUS_{04}	I think that I need the support of a technical person to be able to use this application
	SUS_{05}	I found the various functions in the application were well integrated
	SUS_{06}	I thought there was too much inconsistency in this application
	SUS_{07}	I would imagine that most people would learn to use this application very quickly
	SUS_{08}	I found the application very awkward to use
	SUS_{09}	I felt very confident using the application
	SUS_{10}	I needed to learn a lot of things before I could get going with this application
TAM [Adapted and Modified from 29]	TAM_{01}	Using this application in my studies enables me to accomplish tasks more quickly than other applications in its class
	TAM_{02}	Using this application improves my study performance
	TAM_{03}	Using this application in my study increases my productivity
	TAM_{04}	Using this application enhances the effectiveness of my study
	TAM_{05}	Using this application makes it easier to do my studies
	TAM_{06}	I have found this application useful in my study
	TAM_{07}	Learning to use this application was easy for me
	TAM_{08}	I found it easy to get this application to do what I wanted it to do
	TAM_{09}	My interaction with this application was clear and understandable
	TAM_{10}	I found this application to be flexible to interact with
	TAM_{11}	It was easy for me to become skillful at using this application
	TAM_{12}	I found this application easy to use

Table 4
Demographics of the participants (N = 1595).

Characteristics	Value	Frequency	Percentage (%)
Age	18 to 21 years	1021	64.01
	Greater than 21 years	574	35.99
Sex	Male	819	51.35
	Female	776	48.65
Monthly family income (in INR)	Below 20,000	933	58.49
	20,000–30,000	326	20.44
	30,000–40,000	101	6.33
	Greater than 40,000	235	14.74
Residential area	Urban	1109	69.53
	Rural	486	30.47
Area of study	Science	562	35.23
	Engineering	924	57.93
	Management	109	6.84
Level of study	Graduate	971	60.88
	Postgraduate	624	39.12
Consumption platform	Smartphones	985	61.76
	Laptops	610	38.24

Table 5
Correlation matrix between all the measures of perceived usability.

Construct	SUS	TAM	PU	PEOU	LTR	OExp
SUS	**1**					
TAM	0.845	**1**				
PU	0.661	0.914	**1**			
PEOU	0.872	0.906	0.664	**1**		
LTR	0.778	0.822	0.721	0.728	**1**	
OExp	0.803	0.856	0.705	0.767	0.883	**1**

of criterion validity and is often used when two different instruments are used for measuring the same phenomenon. Since for the present case two different instruments (SUS and TAM) are used for measuring the same concept of perceived usability, hence checking the concurrent validity is justified. According to extant research, concurrent validity is satisfied if the correlation between the different metrics exceeds 0.30 (Hair et al., 1998). It is obvious that if two metrics measure the same thing, then one would expect a higher correlation between them. In Table 5 we present the correlation matrix for all the measures of perceived usability (SUS and TAM) that has been used in this study. Results show that all the values exceed the threshold of 0.30, and are statistically significant ($p < 0.05$), thereby signifying that concurrent validity is satisfied.

4.5. Construct validity

The construct validity is checked by conducting a parallel analysis. Parallel analysis is a powerful technique by which it can be determined that how many components or factors should be retained after doing the factor analysis. Originally, this technique was proposed by authors in O'Connor (2000) and has since then been used widely for measuring the construct validity (Lewis, 2018, 1995, 2014, 2018). This technique is based on extracting the eigenvalues from random datasets that parallels the actual raw dataset with regards to the number of cases and variables. For SUS parallel analysis indicated a one-factor solution, but in case of TAM it indicated a two-factor solution. For aligning the items with the factors an unweighted least square with varimax rotation procedure is used. The first six items load on the same factor of PU having different weights ranging from 0.712 to 0.764, whereas the last six items load on the other factor of PEOU having weights ranging from 0.787 to 0.853. These results are expected as the items load onto their corresponding factors.

4.6. Discriminant validity

The discriminant validity tests whether the measurements that are not supposed to be related are actually unrelated. For calculating discriminant validity, the average variance extracted (AVE) value is calculated for each construct and compared with the shared variances. The Fornell Larcker criterion of discriminant validity states that the variance extracted for each construct should be greater than any squared correlation among the constructs, implying that the constructs are empirically distinct. Alternatively, the square-root of AVE of each construct must be more than the correlation coefficient between the other constructs. Table 6 shows the test of discriminant validity.

4.7. Comparing the means from SUS and modified TAM

Table 7 provides the results with regards to the minimum, maximum and the mean scores obtained from the two different methodologies i.e. the SUS approach and the TAM approach. Since the adjective rating scales like the Curved Grading Scale (CGS) provide an easy way

Table 6
Test of discriminant validity.

Construct	SUS	TAM	LTR	OExp
SUS	0.922			
TAM	0.845	0.885		
LTR	0.778	0.822	0.914	
OExp	0.803	0.856	0.883	0.897

*Note: The diagonal elements represent the square-root of AVE.

of representation and checking the product usability, therefore in Table 7 the CGS grades are also reported. Based on current studies, a mean SUS score of 68 is considered to be the reference level (Lewis, 2018, 1991, 1995, 2018; Sauro & Lewis, 2016). Any score above 68 is thus above average, and that below 68 is considered to be below average. Moreover, in the present case as evident from the results, the mean grade obtained by following either approaches are similar although the standard deviations are different. Further, no significant difference is observed between the mean SUS and TAM scores by carrying out an independent sample *t-test* ($t = 0.47, df = 1593, p = 0.63$) for the entire sample. Therefore, the HCI based approach and the IS based approach seems to be equivalent and consistent with each other by producing similar results.

4.8. Regression analysis

A regression analysis is conducted taking likelihood to recommend and overall experience as the dependent variables under the following three conditions: (a) predicting LTR and OExp with PU and PEOU (b) predicting LTR and OExp with PU and SUS (c) predicting LTR and OExp with PEOU and SUS. Substituting SUS alternatively, with PEOU and then with PU while predicting LTR and OExp enables to gain an insight that it is more closely related to which of the factors among the two. Table 8 presents the results of the regression analysis. All the values presented in Table 8 are statistically significant i.e. $p < 0.05$.

The results show that PU and PEOU account for around 66.7% of the variance in LTR with a 95% confidence interval, and the β coefficients are significant for the predictors. When SUS is used instead of PEOU the variance explained increases to 68.6%, the results still being significant. Moreover, when both PEOU and SUS are used for predicting LTR, the variance explained is maximum at 72.3% and the weightage of the SUS component is slightly greater than the PEOU component at 0.532 and 0.513 respectively. The trend is similar for the second case of predicting the overall experience OExp. PU and PEOU accounts for 68.2% of the variability in OExp, the β coefficients being significant for both the predictors. Again, substituting SUS with PEOU there is a slight increase in variance to 71.8%. Finally, the variance explained is maximum when PEOU and SUS are taken together in the equation for predicting OExp.

4.9. Effect of consumption platform and gender on perceived usability

In order to analyze the effect of the consumption platform (smartphones vs. laptops) along with the gender of the participants on the perceived usability of Microsoft Teams a 2 (consumption platform) × 2 (gender) factorial analysis of variance (two-way ANOVA) is conducted on the mean SUS and TAM scores. The summary of the test is provided in Table 9 for both the cases. The advantage of carrying out the two-way analysis instead of the one-way ANOVA is that in addition to the main effects of each factor it is also possible to obtain the interaction effect between the factors. The (consumption platform × gender) interaction effect is shown in Fig. 1. The results from two-way ANOVA show that there is no main effect of the consumption platform either on the mean SUS scores, $F\ (1, 1591) = 0.153, p = 0.696$ or the mean scores obtained from TAM, $F\ (1, 1591) = 1.777, p = 0.183$ and for the combined sample also, $F\ (1, 2816) = 0.490, p = 0.484$. The main effect of gender is also found to be non-significant under all the three conditions: $F\ (1, 1591) = 0.471, p = 0.492$ (for SUS), $F\ (1, 1591) = 0.353$, $p = 0.552$ (modified TAM) and $F\ (1, 2816) = 0.023, p = 0.878$ (overall

Table 8
Results of regression analysis.

Predicting	Predictors	Variance Explained	Regression Weights	
			β_1	β_2
LTR	PU and PEOU	0.667	0.432	0.441
LTR	PU and SUS	0.686	0.428	0.495
LTR	PEOU and SUS	0.723	0.513	0.532
OExp	PU and PEOU	0.682	0.336	0.587
OExp	PU and SUS	0.718	0.378	0.595
OExp	PEOU and SUS	0.746	0.546	0.601

Note: The β weights are for the predictors in order. For e.g. in the second row 0.432 and 0.441 represent the weights for PU and PEOU respectively while predicting LTR.

Table 9
Effect of consumption platform and gender on perceived usability.

Method	Variable	F Statistic	Significance (p value)
SUS	Gender	0.471	0.492
	Platform	0.153	0.696
	Gender × Platform	0.178	0.674
TAM	Gender	0.353	0.552
	Platform	1.777	0.183
	Gender × Platform	0.038	0.845
Combined Sample	Gender	0.023	0.878
	Platform	0.490	0.484
	Gender × Platform	0.033	0.856

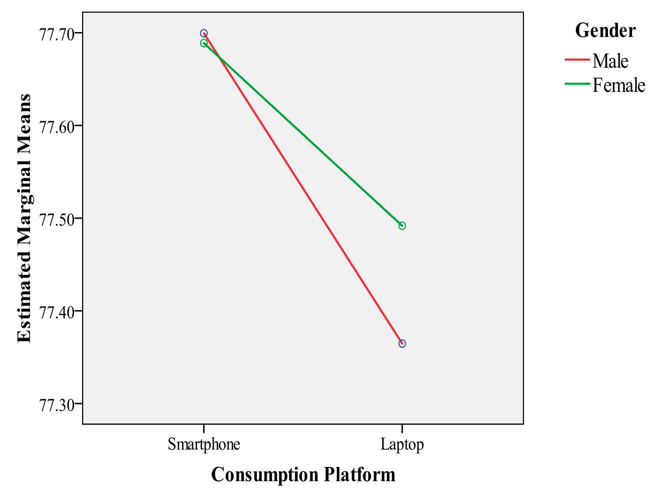
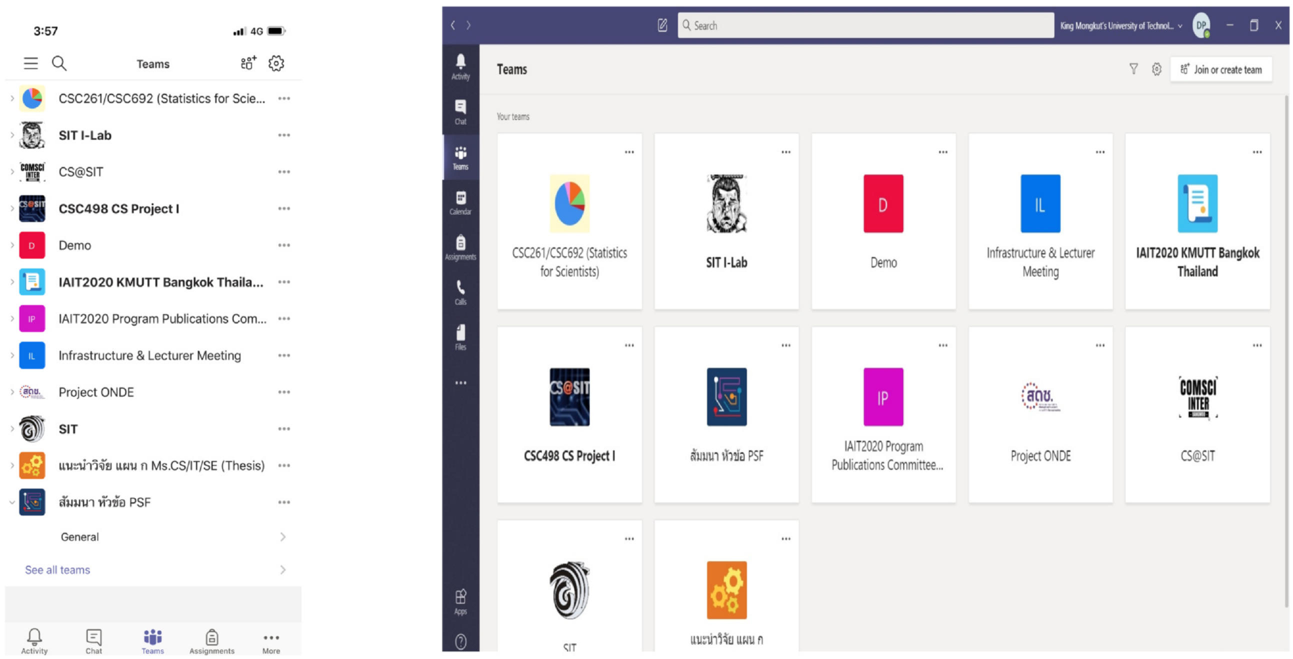


Fig. 1. Interaction effect between consumption platform and gender.

Table 7
Mean scores from SUS and TAM related to the CGS grades.

Methodology	Min Score	Max Score	Mean Score	Std Dev	Min Grade	Max Grade	Mean Grade
SUS	48.10	99.15	77.20	8.34	F	A+	B+
TAM	43.39	97.23	78.04	11.59	F	A+	B+



(a) (b)

Fig. 2. An example of the home screen of Microsoft Teams for (a) the mobile version (b) the web version.

sample). Finally, the interaction effect between gender and platform is also found to be non-significant for all the three cases: F (1, 1591) = 0.178, p = 0.674 (SUS), F (1, 1591) = 0.038, p = 0.845 (modified TAM) and F (1, 2816) = 0.033, p = 0.856 (overall sample). As evident from Fig. 1, numerically the females give a higher usability score than their male counterparts. Moreover, although marginal, but the perceived usability score of Microsoft Teams for the mobile version is higher than the web version. However, statistically the results are not significant.

5. Discussion and conclusion

In the current study the perceived usability of an online learning platform is evaluated taking Microsoft Teams as the reference. Considering the present situation of COVID-19 pandemic prevailing across the globe, it is important to conduct a usability evaluation of the popularly used tools for the purpose of online education delivery, more so because the education delivery mode has shifted to "*online only*". The usability of the learning tools, therefore become an important aspect for ensuring the online learning to be effective and useful for the students. In addition to this there is the problem of digital divide, specifically in the developing countries, where the consumption platform can vary. Although, most of the application developers these days focus on developing applications both for the mobile and the web environment, yet, related to the present context of online education delivery little is known about the perceived usability of these applications across the two different platforms. This is an aspect the present study investigates into. Another aspect into the usability evaluation is the plethora of different approaches and techniques used, primarily that can be classified under the HCI or the IS domains, and whether these two different approaches are equivalent and consistent with each other. Therefore, another aim of this work is to investigate the relationships between the measures of perceived usability as evaluated by the HCI researchers by taking SUS as the reference instrument due to its widespread popularity and validity with the measures of perceived usefulness (PU) and perceived ease of use (PEOU) that are the core components of the hugely popular technology acceptance model (TAM) used by the IS researchers.

5.1. Perceived usability and the effect of consumption platform

Both in case of SUS and TAM neither the consumption platform nor the gender of the participants is found to have any significant effects (direct or interaction) on the perceived usability (scores obtained from SUS and TAM). Extant literatures suggest that the end-user experience depends to a great extent on the screen size of the devices used (Maniar et al., 2007; Pal & Vanijja, 2017). Similarly, the overall usability of the applications also increases with an increase in the screen real estate (Robertson et al., 2005). Considering an average screen size of 6.5 in. for smartphones and 13 in. for laptops, there is almost a doubling of the screen real estate, yet it does not translate to a better perceived usability. The results obtained in this study are a bit unexpected as conventionally one would expect the usability of the web-based platform to be better than the mobile platform. There can be several reasons behind the current findings.

First, in case of the smartphones the screen real estate is limited. This forces the application developers to reduce the functionalities, features and contents for the mobile platform and keep them to the bare essential ones. For example, in case of Microsoft Teams when using the mobile version of the application it is not possible to initiate meetings, however the users can join anytime to any scheduled meeting(s) that they have permission to. Not supporting some features and providing only the most essential ones enable the developers to choose carefully what they want to display on the smartphone screens, featuring only those that are mostly used by the users. This helps them to offset the drawbacks of having a smaller screen size, by keeping the user interface simple, yet functional. Therefore, for the students perceived usability for both the platforms are similar. For example, in Fig. 2 the home screen of Microsoft Teams is shown for both the mobile and the web versions. Although, the content shown is more for the web version, yet for the mobile version the layout is designed in a simple easy to navigate vertical manner having all the essential and frequently used features. Second, nearly most of the students use a smartphone as of today. Extant research has shown that if the end-users are familiar in using some device for a long period of time, then in that case such a familiarity translates to a perception of greater usability (Kortum & Johnson, 2013; McLellan, Muddimer, & Peres, 2012). Moreover, using the same

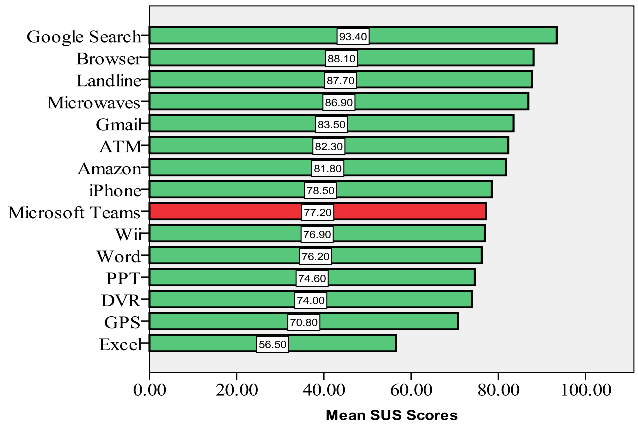


Fig. 3. Comparison of SUS scores of commonly used products as reported in (Kortum & Bangor, 2013) with Microsoft Teams.

device for multiple purposes makes the users experts, which helps in boosting their usability perceptions (Kortum & Johnson, 2013; McLellan et al., 2012). Since the smartphones are heavily used by the students, therefore it makes them more familiar and comfortable with their usage, rather than the laptops which are less portable than the smartphones. Therefore, the smartphones despite having a smaller screen real estate when compared to laptops have the same level of perceived usability as evident from the current scenario. Third, due to the popularity and the widespread use of the smartphones, the mobile platform has matured a lot. Thus, the application developers now give equal efforts in developing their applications for both the platforms. Thus, the usability does not vary depending on the consumption platform.

Since in this work we used SUS as one of the methodologies for evaluating the perceived usability of the online learning platform, additionally we were interested in checking out how Microsoft Teams performed with respect to the other consumer grade products in terms of the SUS scores of perceived usability. For this the work of Kortum and Bangor (2013) is referred to wherein they carried out a large-scale evaluation on 14 different consumer grade products. Fig. 3 shows the comparison of the SUS scores for the various scenarios. As evident from the figure, the usability of Microsoft Teams lies somewhere in the middle region indicating a decent usability of the application.

5.2. SUS and modified TAM: The equivalence

Both the questionnaires used in the survey (SUS and TAM) have acceptable high levels of reliability. Consistent with previous studies, parallel analysis indicated the presence of a one-factor solution for SUS (Kortum & Bangor, 2013; Lewis, 2018). For TAM a two-factor solution is obtained with items 1 to 6 aligning on one factor (PU) and those from 7 to 12 on the other factor (PEOU). The correlation matrix shows that the scores obtained from SUS and TAM are not only highly correlated with each other and significant, but also PEOU correlates significantly more than PU with the SUS scores. When the mean scores obtained from SUS and TAM are compared the differences are not found to be statistically significant, indicating that both the approaches are essentially equivalent and measure the same thing i.e. perceived usability. The results from the different regression models (predicting LTR and OExp) show that both the sub-components of TAM i.e. PU and PEOU are statistically significant. Moreover, all the regression models are roughly similar with regards to the variance explained (the *adjusted* R^2 values) and the beta coefficients, particularly when SUS is replaced with PEOU. Although marginal, the explanatory power of the model is maximum when both SUS and PEOU are accounted for. Concluding from all the above observations it appears that both SUS as well as TAM are good instruments for measuring the perceived usability, and they can be used interchangeably by researchers based on the research context and their preferences. Moreover, focusing exclusively on TAM the PEOU component appears to be more closely related to the concept of perceived usability than the PU component.

5.3. Conclusion and future research direction

In this work we presented the perceived usability of Microsoft Teams taking it as a reference for the online learning platform in times of the COVID-19 pandemic. For measuring the perceived usability both HCI and IS based approaches are used by taking the SUS and TAM instruments as the baseline measurements respectively. Both the approaches give identical results, and we can conclude that SUS as well as the TAM scores are indicative of the likelihood to recommend and the overall experience obtained after using Microsoft Teams, which is an interesting finding for researchers working on user experience. The use of SUS and PEOU in the regression models interchangeably without any significant changes in the weight of the β coefficients or the variance explained indicates that SUS and PEOU are metrics that are developed independent from each other, yet both of them measure the same concept of perceived usability. Therefore, the two approaches are consistent with each other. Moreover, SUS is correlated in a better manner with PEOU, rather than PU. The effect of the consumption platform is found to be not significant for measuring the perceived usability, despite the fact that screen real estate for laptops is at least double when compared to smartphones. This indicates that the application developers put in equal efforts when developing their applications for both the platforms, such that the negative effects of a smaller screen real estate in case of the smartphones do not lead to a bad user experience. The mobile platform has matured enough with the growing popularity of the smartphones that has translated to a greater level of perceived usability.

The first drawback of this work is the use of Microsoft Teams only as the reference platform for measuring the perceived usability of the online learning platforms. Although, we have provided justifications for taking this decision, still for future works it will be better to consider other popular learning tools into account, for e.g. Google Classroom. A comparative analysis can be done between the two platforms to bring out the real state of art of the online learning environment in terms of the perceived usability. The second drawback stems from the methodology that is used to collect data i.e. the survey approach. Generally, in usability research apart from surveys usability studies are also carried out for the purpose of data collection. However, no such usability study is done for the present case due to the lockdown scenario because of COVID-19. We tried to compensate for this by considering a large enough sample, however future studies can focus on the traditional usability testing in a controlled environment. The third drawback can be attributed to the geographical location of the students who participated in the survey. All the students come from one specific country, although they are both from rural and urban areas. However, since the usage behavior and consequently the perception of usability can change with culture, therefore future studies must focus on a cross-country approach so that the current findings can be generalized. Fourth, in this work the effectiveness of the online learning platform in terms of the perceived usability is measured from a student perspective. However, the course instructors i.e. the teachers are on the other side of the same coin and gaining insights to their views about the usability aspect of the online learning platforms is also important and can be done as future work. The current work focuses heavily on the current pandemic situation by evaluating the usability of Microsoft Teams as the learning platform, however it is equally important to focus on how it will be possible to sustain the use of online learning after the pandemic, instead of simply switching back to the traditional face-to-face teaching routine. Therefore, it is equally important to evaluate the sustainability of the learning platforms like Microsoft Teams as it will help in providing insights to the extent to which these new tools can help in achieving their intended and potential benefits. Apart from usability, future works can focus on this sustainability aspect too. Fifth, considering the closeness of PEOU construct of TAM with SUS, future IS related studies that use TAM as the reference framework, can replace the questionnaires related to perceived ease of use with SUS, and check the predictive capability of the model. An interchangeable use of these constructs in different contexts may give newer insights to the usability aspect. Finally, how accurately SUS represents the usability of online learning platforms like Microsoft Teams is too early to comment. The number of works using SUS for evaluating usability of online learning applications are relatively few, therefore more such usability studies need to be carried out for developing strong norms. In general, online learning is a complex scenario having multiple factors like the quality of the course contents, quality of the video lectures, the extent of support provided by the system, the UI design of the learning system, interactivity, and learnability of the system that might affect their usage (Junus, Santoso, Isal, & Utomo, 2015). Usability is just a small part of the overall user experience that the current work tries to capture, while future studies may focus on the broad user experience aspect.

Funding

None.

Availability of data and material

Not applicable.

Code availability

Not applicable.

CRediT authorship contribution statement

Debajyoti Pal: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Vajirasak Vanijja:** Investigation, Methodology, Supervision, Writing - review & editing.

Declaration of Competing Interest

All the authors of this paper declare that there is no conflict of interest.

References

Abidah, A., Hidaayatullah, H. N., Simamora, R. M., Fehabutar, D., & Mutakinati, L. (2020). The impact of COVID-19 to Indonesian Education and its Relation to the philosophy of "Merdeka Belajar". *Studies in Philosophy of Science and Education, 1*(1), 38–49.

Abrami, P. C., Bernard, R. M., Bures, E. M., Borokhovski, E., & Tamim, R. M. (2011). Interaction in distance education and online learning: Using evidence and theory to improve practice. *Journal of Computing in Higher Education, 23*, 82–103.

Abuhlfaia, K., & Quincey, E. D. (2019). Evaluating the Usability of an E-Learning Platform within Higher Education from a Student Perspective. *Proceedings of the 3rd International Conference on Education and E-learning (ICEEL 2019)* (pp. 1–7). New York: ACM.

Abuhlfaia, K., & Quincey, E. (2019). Evaluating the usability of an E-learning platform within higher education from a student perspective. *Proceedings of the 2019 3rd International Conference on Education and E-Learning (ICEEL 2019)* (pp. 1–7). New York, NY, USA: Association for Computing Machinery.

Alqurashi, E. (2019). Predicting student satisfaction and perceived learning within online learning environments. *Distance Education, 40*(1), 133–148.

Arain, A. A., Hussain, Z., Rizvi, W. H., & Vighio, M. S. (2016). Evaluating usability of M-learning application in the context of higher education institute. *Proceedings of the International Conference on Learning and Collaboration Technologies: vol. 9753*, (pp. 259–268). Toronto, Canada: Springer.

Arias, A. V., Naffah, S. C., & Hernandez, J. B. (2019). A proposed model of E-learning tools acceptance among university students in developing countries. *Education and Information Technologies, 24*, 1057–1071.

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

Baby, A., & Kannammal, A. (2020). Network path analysis for developing an enhanced TAM Model: A user-centric E-learning perspective. *Computers in Human Behavior, 107*, 106081.

Bangor, A., Kortum, P. T., & Miller, J. T. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies, 4*(3), 114–123.

Bettiga, D., Lamberti, L., & Lettieri, E. (2020). Individuals' Adoption of smart technologies for preventive health care: A structural equation modeling approach. *Health Care Management Science, 23*, 203–214.

Boateng, R., Mbrokoh, A. S., Boateng, L., Senyo, P. K., & Ansong, E. (2016). Determinants of E-learning adoption among students of developing countries. *The International Journal of Information and Learning Technology, 33*(4), 248–262.

Brooke, J. (1996). SUS: A "Quick and Dirty" Usability Scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.). *Usability Evaluation in Industry* (pp. 189–194). Taylor and Francis.

Condie, R., & Livingston, K. (2007). Blending online learning with traditional approaches: changing practices. *British Journal of Educational Technology, 38*(2), 337–348.

Davis, D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

Diefenbach, S., Kolb, N., & Hassenzahl, M. (2014). The "Hedonic" in Human-computer Interaction: History, Contributions, and Future Research Directions. *Proceedings of the 2014 Conference on Designing Interactive Systems – DIS 14* (pp. 305–314). New York: ACM.

Dutot, V., Bhatiasevi, V., & Bellallahom, N. (2019). Applying the technology acceptance model in a three-countries study of Smartwatch adoption. *The Journal of High Technology Management Research, 30*(1), 1–14.

Estriegana, R., Merodio, J. A. M., & Barchino, R. (2019). Student acceptance of virtual laboratory and practical work: An extension of the technology acceptance model. *Computers & Education, 135*, 1–14.

Fernandes, T., & Guerra, J. (2019). Drivers and deterrents of music streaming services purchase intention. *International Journal of Electronic Business, 15*(1), 21–42.

Finstad, K. (2010). The usability matric for user experience. *Interacting with Computers, 22*(5), 323–327.

Gahtani, S. S. A. (2016). Empirical investigation of E-learning acceptance and assimilation: A structural equation model. *Applied Computing and Informatics, 12*(1), 27–50.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). New Jersey: Prentice Hall International Inc.

Hamidi, H., & Chavoshi, A. (2018). Analysis of the essential factors for the adoption of mobile learning in higher education: A case of study of students of the university of technology. *Telematics and Informatics, 35*(4), 1053–1070. https://doi.org/10.1016/j.tele.2017.09.016.

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

Hasibuan, D. P., Santoso, H. B., Yunita, A., & Rahmah, A. (2019). An Indonesian adaptation of the E-learning usability scale. *Proceedings of 4th International Conference on Computing and Applied Informatics (ICCAI 2019), Medan, Indonesia* (pp. 1–7). .

Hidalgo, E. J. P., Abril, C. A. H., & Parra, M. E. G. (2020). *MOOCs: Origins, Concepts and didactic applications: A systematic review of the literature (2012–2019)*. Technology, Knowledge and Learning1–27.

Iivari, N., Kinnula, M., Juustila, T. M., & Kuure, L. (2018). Exclusions in social inclusion projects: Struggles in involving children in digital technology development. *Information Systems Journal, 28*(6), 1020–1048.

Iivari, N., Sharma, S., & Olkkonen, L. V. (2020). Digital transformation of everyday life – How COVID-19 pandemic transformed the basic education of the young generation and why information management research should Care? *International Journal of Information Management,* 102183. https://doi.org/10.1016/j.ijinfomgt.2020.102183.

Jong, P. G. M., Pickering, J. D., Hendriks, R. A., Swinnerton, B. J., Goshtasbpour, F., & Reinders, M. E. J. (2020). Twelve tips for integrating massive open online course content into classroom teaching. *Medical Teacher, 42*(4), 393–397.

Junus, I. S., Santoso, H. B., Isal, R. Y. K., & Utomo, A. Y. (2015). Usability evaluation of the student centered e-learning environment. *The International Review of the Student-centered E-learning Environment, 16*(4), 62–82.

Kamal, S. A., Shafiq, M., & Kakria, P. (2020). Investigating Acceptance of telemedicine services through an extended technology acceptance model (TAM). *Technology in Society, 60*, 101212. https://doi.org/10.1016/j.techsoc.2019.101212.

Kapasia, N., Paul, P., Roy, A., Saha, J., Zaveri, A., Mallick, R., ... Chouhan, P. (2020). Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 Pandemic in West Bengal, India. *Children and Youth Services Review, 116*, 105194. https://doi.org/10.1016/j.childyouth.2020.105194.

Kortum, P., & Bangor, A. (2013). Usability ratings for everyday products measured with the system usability scale. *International Journal of Human-Computer Interaction, 29*(2), 67–76.

Kortum, P., & Johnson, M. (2013). The relationship between levels of user experience with a product and perceived system usability. *Proceedings of the Human Factors and Ergonomics Society, 57*(1), 197–201.

Kortum, P., & Sorber, M. (2015). Measuring the usability of mobile applications for phones and tablets. *International Journal of Human-Computer Interaction, 31*(8), 518–529.

Lee, S. (2018). Net promoter score: Using NPS to measure IT customer support satisfaction. *Proceedings of the 2018 ACM SIGUCCS Annual Conference (SIGUCCS '18)* (pp. 63–64). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3235715.3235752.

Lewis, J. R. (1991). An After-scenario questionnaire for usability studies: Psychometric evaluation over three trials. *ACM SIGCHI Bulletin, 23*(4), 79. https://doi.org/10.

1145/126729.1056077.
Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction, 7*(1), 57–78.
Lewis, J. R. (2014). Usability: Lessons learned and yet to be learned. *International Journal of Human-Computer Interaction, 30*(9), 663–684.
Lewis, J. R. (2018). Measuring perceived usability: The CSUQ, SUS, and UMUX. *International Journal of Human-Computer Interaction, 34*(12), 1148–1156.
Lewis, J. R. (2018). The system usability scale: Past, present, and future. *International Journal of Human-Computer Interaction, 34*(7), 577–590.
Luo, G. H., Liu, E. Z. F., Kuo, H. W., & Yuan, S. M. (2014). Design and implementation of a simulation-based learning system for international trade. *The International Review of Research in Open and Distributed Learning, 15*(1), 203–226.
Maniar, N., Bennett, E., & Gal, D. (2007). The effect that screen size has on video-based M-learning. *Proceedings of the Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PerComW'07), White Plains, NY* (pp. 145–148). .
Mariën, I., & Prodnik, J. A. (2014). Digital inclusion and user (dis)empowerment: A critical perspective. *Info, 16*(6), 35–47.
McLellan, S., Muddimer, A., & Peres, S. C. (2012). The effect of experience on system usability scale ratings. *Journal of Usability Studies, 7*(2), 56–67.
Millat, I. E., Lopez, F. J. M., Jover, M. P., Abad, J. C. G., & Alegret, A. (2018). An extension of the technology acceptance model for online learning environments. *Interactive Learning Environments, 26*(7), 895–910.
Ngai, E. W. T., Poon, J. K. L., & Chan, Y. H. C. (2007). Empirical examination of the adoption of WebCT using TAM. *Computers & Education, 48*(2), 250–267.
Nikou, S. A., & Economides, A. A. (2018). Factors that influence behavioral intention to use mobile-based assessment: A STEM teachers' perspective. *British Journal of Educational Technology, 50*(2), 587–600.
O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*(3), 396–402.
Orfanou, K., Tselios, N., & Katsanos, C. (2015). Perceived usability evaluation of learning management systems: Empirical evaluation of the system usability scale. *The International Review of Research in Open and Distributed Learning, 16*(2), 227–246.
Owen, R. (2018). Net promoter score and its successful application. *Marketing Wisdom,* 17–29. https://doi.org/10.1007/978-981-10-7724-1_2.
Pal, D., Arpnikanondt, C., Funilkul, S., & Chutimaskul, W. (2020). The adoption analysis of voice based smart IoT products. *IEEE Internet of Things Journal*. https://doi.org/10.1109/JIOT.2020.2991791.
Pal, D., & Triyason, T. (2018). User intention towards a music streaming service: A Thailand case study. *KnE Social Sciences, 3*(1), 1–16.
Pal, D., Triyason, T., Funilkul, S., & Chutimaskul, W. (2018). Smart homes and quality of life for the elderly: Perspective of competing models. *IEEE Access, 6*, 8109–8122.
Pal, D., & Vanijja, V. (2017). A no-reference modular video quality prediction model for H.265/HEVC and VP9 codecs on a mobile device. *Advances in Multimedia, 2017*, 1–19. https://doi.org/10.1155/2017/8317590.
Panigrahi, R., Srivastava, P. R., & Sharma, D. (2018). Online learning: Adoption, continuance, and learning outcome – A review of literature. *International Journal of Information Management, 43*, 1–14. https://doi.org/10.1016/j.ijinfomgt.2018.05.005.
Park, E., Kim, S., Kim, Y., & Kwon, S. J. (2018). Smart Home services as the next mainstream of the ICT industry: Determinants of the adoption of smart home services. *Universal Access in the Information Society, 17*, 175–190.
Pereira, S. A., & Anamaria, M. (2012). The ergonomic evaluation of a virtual learning environment usability. *Work, 41*(1), 1140–1144. https://doi.org/10.3233/WOR-2012-0293-1140.
Peres, S. C., Pham, T., & Phillips, R. (2013). Validation of the system usability scale (SUS): SUS in the wild. *Proceedings of the Human Factors and Ergonomics Society, 57*(1), 192–196.
Persico, D., Manca, S., & Pozzi, F. (2014). Adapting the technology acceptance model to evaluate the innovative potential of E-learning systems. *Computers in Human Behavior, 30*, 614–622.
Rahmi, W. A., Aldraiweesh, A., Yahaya, N., Cumin, Y. B., & Zeki, A. M. (2019). Massive open online courses (MOOCs): Data on higher education. *Data in Brief, 22*, 118–125. https://doi.org/10.1016/j.dib.2018.11.139.
Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review, 81*, 46–54.
Revythi, A., & Tselios, N. (2019). Extension of technology acceptance model by using system usability scale to assess behavioral intention to use E-Learning. *Education and Information Technologies, 24*, 2341–2355.
Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., & Tan, D. (2005). The large-display user experience. *IEEE Computer Graphics and Applications, 25*(4), 44–51.
Rodrigues, H., Almeida, F., Figueiredo, V., & Lopes, S. L. (2019). Tracking E-learning through published papers: A systematic review. *Computers & Education, 136*, 87–98. https://doi.org/10.1016/j.compedu.2019.03.007.
Rodriguez, T. E., & Lozano, P. M. (2012). The acceptance of Moodle technology by business administration students. *Computers & Education, 58*(4), 1085–1093.
Salloum, S. A., Alhamad, A. Q. M., Al-Emran, M., Monem, A. A., & Shaalan, K. (2019). Exploring students' acceptance of E-learning through the development of a comprehensive technology acceptance model. *IEEE Access, 7*, 128445–128462.
Sami, B., Rutter, M., & Smith, S. (2016). The utilization of system usability scale in learning management systems: A case study of Jeddah community college. *Proceedings of the 9th International Conference of Education, Research, and Innovation (ICERI 2016), Seville, Spain* (pp. 5314–5323). .
Sauro, J. & Lewis, J. R. (2016). Quantifying the user experience: practical statistics for user research. (2nd ed.), Cambridge, MA: Morgan Kaufmann. doi: 10.1016/C2010-0-65192-3.
Scherer, R., Siddiq, F., & Teo, T. (2015). Becoming more specific: Measuring and modelling teachers' perceived usefulness of ICT in the context of teaching and learning. *Computers & Education, 88*, 202–214.
Shi, L., Awan, M. S. K., & Cristea, A. I. (2013). Evaluating system functionality in social personalized adaptive E-learning systems. *Proceedings of European Conference on Technology Enhanced Learning, Paphos, Cyprus* (pp. 633–634). .
Simoes, A. P., & Moraes, A. D. (2012). The ergonomic evaluation of a virtual learning environment usability. *Work, 41*(1), 1140–1144.
Song, Z., Wang, C., & Bergmann, L. (2020). China's prefectural digital divide: Spatial analysis and multivariate determinants of ICT diffusion. *International Journal of Information Management, 52*, 102072. https://doi.org/10.1016/j.ijinfomgt.2020.102072.
Srivastava, S. C., & Shainesh, G. (2015). Bridging the service divide through digitally enabled service innovations: Evidence from Indian healthcare service providers. *MIS Quarterly, 39*(1), 245–267.
Tiyar, F. R., & Khoshsima, H. (2015). Understanding Students' Satisfaction and Continuance Intention of e-learning: Application of Expectation-Confirmation Model. *World Journal on Educational Technology, 7*(3), 157–166.
Tractinsky, N. (2018). The usability construct: A dead end? *Human-Computer Interaction, 33*(2), 131–177.
Wang, L. Y. K., Lew, S. L., Lau, S. H., & Leow, M. C. (2019). Usability factors predicting continuance of intention to use cloud E-learning application. *Heliyon, 5*(6), e01788. https://doi.org/10.1016/j.heliyon.2019.e01788.
Wu, J. H., Chen, Y. C., & Lin, L. M. (2007). Empirical evaluation of the revised end user computing acceptance model. *Computers in Human Behavior, 23*(1), 162–174.
Yalcin, M. E., & Kutlu, B. (2019). Examination of students' acceptance of and intention to use learning management systems using extended TAM. *British Journal of Educational Technology, 50*(5), 2414–2432.
Zhang, D., Zhao, J. L., Zhou, L., & Nunamaker, J. F. (2004). Can e-learning replace classroom learning? *Communications of the ACM, 47*(5), 75–79.