SN Computer Science (2022) 3:238
https://doi.org/10.1007/s42979-022-01138-5

ORIGINAL RESEARCH

Perceived Level of Usability as an Evaluation Metric in Adaptive E-learning

A Case Study with Dyslexic Children

Weam Gaoud Alghabban[1,2] · **Robert Hendley**[2]

Received: 14 September 2021 / Accepted: 6 April 2022 / Published online: 23 April 2022

Abstract
The global spread of COVID-19 has shifted the learning process towards e-learning. In this context, a critical challenge for researchers is to understand and evaluate the effectiveness of e-learning, especially when the learning is adapted to the needs of individual users. In this work we argue that the learner's perception of the level of usability of a system is a valuable metric that gives an insight into the learners' engagement and motivation to learn. Little attention has been paid to this metric. In this paper we explore why this is important and valuable. We present a case study which uses the System Usability Scale (SUS) questionnaire to measure the user's perception of usability as an indirect (proxy) measure of engagement. A between-subject experiment was conducted with 41 learners with dyslexia. The intervention group used the adaptive version of the e-learning system that matched the material to the needs of the learner. The control group used a standard version. At the end, learning gain and SUS scores were assessed. The correlation between learning performance and the perceived level of usability was positive and moderate (0.517, $p < 0.05$) among participants in the intervention group. However, learning performance and perceived level of usability were unrelated in the control group ($-$ 0.364, $p > 0.05$). From this, and other work, it appears that using a learner's assessment of the usability of a system is an effective way to measure their attitude to their learning. It reflects their perception of its suitability to their needs and this, in turn, is likely to affect their engagement and motivation. As such, this provides an effective instrument to judge whether adaptation based on learner needs has been successful.

Keywords E-learning evaluation · Adaptation · Perceived usability · System Usability Scale · Learning gain · Dyslexia

Introduction

Learning has changed dramatically recently, especially with the novel coronavirus (COVID-19) pandemic, and electronic learning (e-learning) has seen a substantial rise. The current challenge of e-learning is not only to make resources available to everyone, at any time and place, and in different formats, but also to offer learning properly—where the learner's motivation, needs, characteristics and background are taken into consideration [42]. These factors highlight a new paradigm as a solution to the limitations of traditional e-learning [66], which does not provide personalization of learning content, interactivity and real-time execution. This new paradigm is called adaptive e-learning. Adaptive e-learning systems can solve these issues by altering the structure, content and presentation of the material to suit the different needs of the different learners [63].

Evaluating the effectiveness of adaptive e-learning systems is a topic that has received attention from many researchers, and a number of studies have identified success factors to maximize the benefits of these systems. Most of the research has focussed on the technical success of these e-learning systems and how learners' characteristics affect the teaching [23]. At the same time, the technology has

"This article is part of the topical collection "Computer Supported Education" guest edited by James Uhomoibhi and Beno Csapó"

✉ Weam Gaoud Alghabban
walghabban@ut.edu.sa

[1] Computer Science Department, Alwajh University College, University of Tabuk, Tabuk 71491, Kingdom of Saudi Arabia

[2] School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

become more accessible and reliable [2]. More recent studies have targeted how learners' attitudes and interactions are significant factors in success. An evaluation of these systems' ability to meet learners' needs and enabling their continuing education is therefore needed.

The learning experiences and achievements of learners is one challenge in the e-learning domain [55]. Learners' satisfaction and learning outcomes are good indicators of the effectiveness of these systems [58]. Another essential indicator of e-learning quality is learners' engagement. Learner engagement refers to the learner's ongoing effort to support their psychological commitment to be motivated, acquire knowledge and achieve learning goals [26]. An increase in the engagement level of learners may serve as a strong predictor of improved performance and achievement [19].

Many approaches are used to measure engagement, such as self-report, interviews, observation and teacher ratings [30]. However, these direct methods can be challenging to implement without affecting the results and they require substantial time and resources. There is therefore a need for indirect measurements to assess engagement efficiently.

Another significant indicator of e-learning quality is usability [37]. Usability refers to the extent to which a product can be used by specified users to achieve specified goals with effectiveness and efficiency. If a learner perceives a system as easy to use, they will be more satisfied and engaged in using the system [12, 68], which in turn helps them achieve learning goals and improve their learning outcome. A highly usable, adaptive e-learning system improves learning performance by enabling learners to concentrate on learning tasks instead of focussing on the system's functionality [51].

Although the level of usability has an impact on a learner's motivation and engagement with their learning [12, 68], the feasibility of using perceived level of usability as a metric to evaluate the effectiveness of adaptation has received little attention [52] (especially in the context of learners with dyslexia—who are the target group in this case study). Without engagement with the learning content, it is difficult to enhance the learning performance of learners with dyslexia [36]. Therefore, assessing these learners' engagement is necessary, in its own right. Also, there are theoretical models and reliable diagnostic tools to identify different characteristics and reading problems of individuals with dyslexia [11, 31]. Therefore, it is appropriate to address these different needs of dyslexia as a basis for adaptation to enhance their learning. In addition, usability research on e-learning largely ignores the learner's perspectives [57], even though the learners are key users of e-learning environments and their characteristics and needs must be the main focus of usability research [12]. So, beyond assessing actual usability, we argue that the learner's assessment (perception) of usability is affected by their overall experience. Therefore, for example, two systems with identical interfaces may be assessed differently, if the content of one is more successfully adapted to the needs of the learner. That is, we can use the learner's usability assessment as a metric to assess the success of the adaptation. This, effectively gives us an indirect way to evaluate their satisfaction and engagement or the value that they attach to the learning.

This research aims to: (i) demonstrate that perceived level of usability is a significant indicator of the effectiveness of adaptation in e-learning systems as an indirect method to measure engagement. We do this by conducting an experimental study of learners with difficulties to understand whether perceived usability is related to learning performance when learning material is adapted to their needs; and (ii) examine a possible correlation between perceived level of usability and grades when the learning material is adapted to learners' needs.

The findings of this research contribute to existing knowledge through providing insight into the viability of using 'perceived level of usability' in evaluating the quality of adaptation in e-learning systems by assessing learners' engagement with the learning content indirectly. This adds novelty and originality to this research in the sense of the metric and methodology used to evaluate this dimension. The approach is implemented in an e-learning system to train learners with dyslexia (the target domain) and to evaluate the benefits of adaptation. It is observed through this research that learning performance is correlated to perceived level of usability when the material is adapted to learners' needs. This implies that when a learner perceives the system as easy to use, they will be motivated and engaged, which in turn enhances learning performance.

This paper is structured in the following way: "Related Work" overviews the background on dyslexia (the domain in which this research is focused), adaptive e-learning systems and different techniques to evaluate the effectiveness of these systems. In "Case Study", the proposed methodology to evaluate this research's aim is described. Next, "Results" presents the results, followed by a detailed discussion of this research in "Discussion". Finally, the conclusion and future research directions are presented in "Conclusions and Future Work".

Related Work

Evaluating the effectiveness of adaptive e-learning systems is critical work, as the use of these technologies has skyrocketed during the COVID-19 pandemic. The literature review in this section covers three aspects: the definition of dyslexia and the difficulties it creates for learning, adaptive e-learning, and techniques used for the evaluation of e-learning systems.

Dyslexia

Dyslexia, a language-based learning disability, is the most common of all specific developmental disorders, affecting around 5–17% of the population [32]. The formal definition of dyslexia, according to the International Dyslexia Association (IDA) in 1994, is: "a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language processing that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge" (p. 2) [45]. The components of this definition are illustrated in Fig. 1.

According to this definition, the major academic skill deficit characterizing children with dyslexia is phonological processing difficulty, including problems with spelling, accuracy and/or fluency of single-word decoding and writing [29]. Reading comprehension can be affected as well [45]. However, the extent to which these deficits occur varies depending on the language's orthography [1]. For example, dyslexia in Arabic is different from dyslexia in English [27]. Two types of languages exist: transparent and non-transparent orthographic languages. Readers in transparent orthographic languages, such as Italian and Spanish, face fewer reading difficulties than readers in non-transparent languages such as French and English [62]. Some languages, such as Arabic and Hebrew, feature two orthographies. Arabic is a widely spoken language with a considerable rate of dyslexia [15]; however, little research has targeted dyslexia in this language, and most of these studies focus on the complexity of Arabic rather than on targeting the needs of learners with dyslexia in the e-learning environment to enhance their experience. This research targets dyslexia in Arabic.

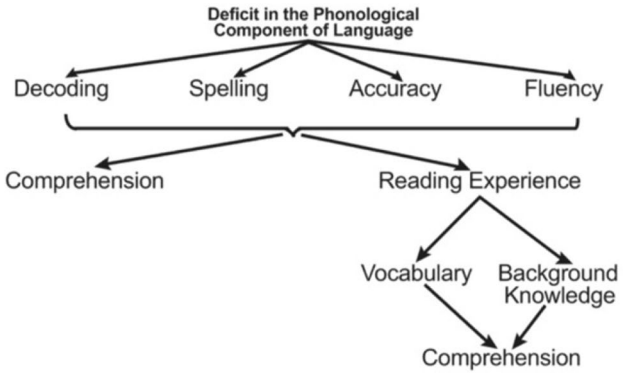

Fig. 1 Components of dyslexia definition adopted by the IDA [29]

Adaptive E-learning

E-learning has a significant impact on learners and has gained acceptance in all academic institutions. E-learning can be defined as "an innovative web-based system based on digital technologies and other forms of educational materials whose primary goal is to provide students with a personalised, learner-centered, open, enjoyable and interactive learning environment supporting and enhancing the learning processes" (p. 95) [57].

In contrast to traditional learning models, the e-learning environment supports the educational process for heterogeneous learners anytime and anywhere [22]. However, traditional, standalone e-learning systems nevertheless have drawbacks when compared to real-life teaching in classrooms, such as the lack of flexible tutorial support, presentation and feedback and the lack of adaptive support [66]. Thus, the e-learning field has shifted to adaptive e-learning, which can adapt learning material according to different learners' requirements [63].

In the e-learning context, adaptation can be defined as a procedure for tailoring the learning environment to accommodate differences between learners [17] to improve learning performance [48]. The objective of adaptation differs from one system to another. Some adaptive e-learning systems aim to enhance the outcome of learners and increase their progress [4], while other systems have different objectives, such as increasing learners' engagement [35], satisfaction and usability [9, 10].

Adaptive e-learning systems can provide suitable learning content according to different learners' characteristics, such as previous knowledge level [21], learning styles [61], and personality [35]. Learners with dyslexia, like normal learners, differ in their needs and characteristics. These differences should be considered individually, instead of treating everyone in the same manner [3]. Among these characteristics of dyslexia are dyslexia type [4] and learning styles [14]. This research targets reading skill level of dyslexia as one characteristic in an adaptive e-learning system [5].

Techniques to Evaluate the Effectiveness of Adaptive E-learning

The prevalence of e-learning highlights the need to evaluate these systems' effectiveness. The evaluation of these systems can be approached from a technical perspective [38] and from a user-centred perspective, as users are the main target of these systems—"the key elements in an e-learning intervention". Thus, learners' needs and abilities must be the main focus of adaptive e-learning when evaluating its effectiveness.

Different evaluation metrics exist. Among these metric are learning gain, learner satisfaction and perceived level of

usability. The following paragraphs explain these metrics in detail.

Learning gain Learning effectiveness is a fundamental factor that should be measured when evaluating adaptive e-learning systems [53]. Learning gain, measured by pre- and post-tests, is a commonly used term to describe any improvement in the learner's knowledge after a specific intervention [34]. Typically, to assess learning gain, the pre-test is administered to a group of learners based on their level of knowledge. Two balanced groups of learners are formed: one group that interacts with the non-adaptive version of the system and the other group that interacts with the adaptive version. If the second group scores higher (in terms of learning gain) than the first group in the post-test, the adaptation is considered a success [34].

For learners with dyslexia, it is crucial to assess the effectiveness of adaptive e-learning systems based on their needs by assessing learning gain. Previous research has demonstrated the success of adaptive e-learning systems in improving dyslexia learning gain after intervention, as seen in [4, 6]. In addition, assessing whether content learned can be generalised to a new context is a critical outcome in education [50]. This reflects the ability of learners to apply what they have learned in the intervention to analyse new material that is unfamiliar to them [44]. However, knowledge generalisation is rarely assessed in most research [41]. Therefore, a measure of generalisation to assess whether learners generalise and can apply learned content to new content is another metric for evaluating the effectiveness of adaptive e-learning systems. Generalisation can be assessed by using material in the pre- and post-tests that varies systematically in terms of similarity to the material used in the e-learning system [25].

Learner satisfaction Learner satisfaction is a critical element of the quality of e-learning experiences [43], and is considered a valid predictor of a system's effectiveness [33]. Satisfaction refers to the pleasure that a user feels when they perform an action or receive something needed to perform an action [59]. In the Human-Computer Interaction (HCI) field, user satisfaction is visualised as the expression gained from interaction with a system [59]. Learner satisfaction is affected by several factors such as engagement and motivation to interact with the system. It is also associated with the extent to which they believe the system they interact with matches their needs [59].

The post-test questionnaire is used to assess learner satisfaction by asking learners questions about their satisfaction after using a system [34]. Several tools can be used to measure learner satisfaction, such as those of [40, 46]. Among these tools, the E-Learner Satisfaction (ELS) tool [65] is a commonly used tool to assess learner satisfaction in adaptive e-learning [9]. ELS is a reliable and validated questionnaire, consisting of 17 items related to four different factors: learning content, system interface, learning community and personalisation [65]. These items have 7-point Likert scales ranging from "strongly disagree" to "strongly agree" [65]. ELS measures overall satisfaction in addition to satisfaction related to each of the four factors [65]. ELS can be adapted to fit a specific research need [65], and is therefore widely used for a variety of e-learning systems [59].

For children, presenting the Likert scale from "strongly disagree" to "strongly agree" is not useful because of the difficulty they encounter in understanding and interpreting it [56]. One widely used alternative instrument is the Smileyometer [56], shown in Fig. 2. It uses pictorial representations based on a 1–5 Likert scale, or a 1-7 Likert scale as in [49], that enables children to identify their feelings by choosing one face. It can be adapted for any tool [60]. Moreover, it is easy to use, quick to complete and does not require any writing.

For learners with dyslexia, measuring the degree to which they are satisfied with their assigned version of the system is useful. It allows them to indirectly measure whether they perceive the system as suitable for their needs. If the system meets their needs and characteristics, they will be more satisfied and motivated, which in turn affects learning effectiveness [3].

Perceived level of usability The way end users perceive their interaction with systems is a significant aspect of e-learning that can lead to a better user experience. Usability is defined as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [39].

Perceived level of usability can be assessed by asking end users to assess the system's usability after use—using standardised usability questionnaires, as in [28]. A commonly used instrument for evaluating the perceived usability of systems in HCI is the System Usability Scale (SUS) questionnaire [16, 37]. SUS has a high degree of validity and reliability [13] and is a valid instrument for comparing the usability of two or more systems [54]. Furthermore, it can be adapted for different contexts [54]. When using SUS, reliable results are evident even with small samples [64]. It can be adapted to be used by children [4] in addition to being used as a usability tool for Arabic users with a high degree of reliability [7].

The SUS includes 10 mixed-tone items on a five-point Likert scale ranging from strongly disagree (1) to strongly agree (5). The even numbered items have a negative tone and

Fig. 2 A Smileyometer [56]

the odd numbered items have a positive tone. To calculate the overall SUS score, the rules are as follows: for each even numbered item, the respondent's answer scale is reduced by 5, and for each odd numbered statement, the respondent's answer scale is reduced by 1. The overall SUS value of each respondent can be obtained by multiplying the sum of scores by 2.5.

Good learning performance and achievement is associated with strong engagement with the learning content [19]. Hence, motivation to learn is used as a usability measure to evaluate an e-learning system [67]. In this research we evaluate the effectiveness of adaptation in the e-learning environment by assessing the viability of perceived level of usability as an indirect metric to measure engagement, and whether perceived level of usability is related to learners' learning gain. We argue that this is an effective way to assess the benefit of adaptation. If a user perceives a system as being easy to use, this will reflect that they are more engaged and motivated, which indirectly improves their learning performance.

Case Study

This section details the methodology used in this research. An experimental study was conducted with Arabic learners with dyslexia in primary schools. We targeted this group for several reasons. First, previous research, such as [8, 18], has failed to consider the variations in skills and backgrounds among children with dyslexia in e-learning. As a result, learners report lower satisfaction and engagement [3]. Second, a number of studies utilised instructor evaluations of e-learning systems, ignoring the learners' perspective [8, 18]. Third, little research has targeted dyslexia in Arabic speakers [4]. Finally, accessibility to children with dyslexia was more straightforward than those in intermediate or high schools.

A web-based e-learning system was designed and implemented to evaluate perceived usability with Arabic learners with dyslexia in primary schools. The system provided six training sessions, each with 20 different word reading/recognition activities. The material used in the system was derived from the school curriculum. Three fundamental Arabic reading skills were chosen that serve as basic foundations for advanced skills (according to interviews with experts). These reading skills are described fully in [5]:

- Reading letters with short vowels (S1).
- Reading words with Sakin letter(s) (S2).
- Reading words with short vowels and Sakin letters (S3).

Figure 3 shows the material structure of the skills in the system.

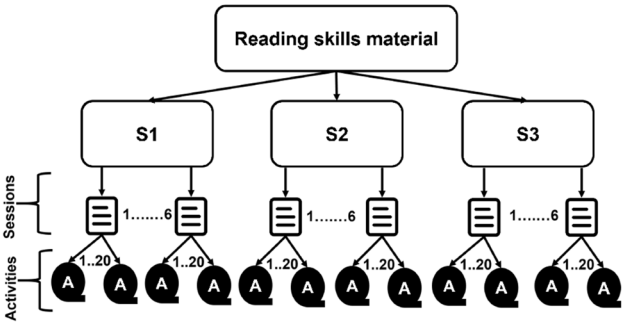

Fig. 3 The training material structure in the e-learning system

Fig. 4 A screenshot example of a training session of S3

In each piece of reading skill material, the difficulty increases gradually. In S1, the activities begin with letters with a simple short vowel (fat-ha (/a/)), followed by (kasra (/i/)) to the advanced (dammah (/u/)). In S2, the activities begin simply (one Sakin letter at the end of two-letter words) to advanced (two Sakin letters in the middle or at the end of three- and four-letters words). A Sakin letter is a letter with a small circle on top of it indicating no vowel. In S3, the activities begin simply (three-letter words with fat-ha (/a/)) and become more advanced (three- and four-letters words with different short vowels and Sakin letters).

An example of a training activity from S3 is shown in Fig. 4. At the top of the screen, the activity question is presented. The target word is denoted as dots inside a rectangular box under the query with the corresponding audio beside it (a question mark icon). The learner can listen to the target word audio and then choose the correct word from among three choices, of which only one is correct. If the learner selects the correct answer, written and spoken praise are provided. Otherwise, negative written and spoken feedback is provided. Training progress is provided as a progress bar at the top of the screen. After every seven completed activities, a motivational message is presented to increase the learner's confidence [15].

Two versions of the system are developed to support the experimental conditions. The standard version fixes the material to cover all three reading skills, while the matched version matches the activities to the reading skill of each learner. If a learner already masters a skill, then the system will not cover the activities of this skill. Rather, it will provide the activities that need to be mastered by the learner. Both versions of the system are identical in layout; the only difference is the provided material in the activities to investigate whether adapting learning material affects learners' perceived level of usability.

Hypotheses

One hypothesis was formulated: *Perceived usability positively correlates to learning gain when learning material is adapted to learners' needs.*

The independent variables were the standard and matched versions of the e-learning system, while the dependent variables were learning gain and the perceived level of usability.

Measurements and Data Collection Tools

The reading skill level of learners was determined using three reliable reading skill level diagnostic tests [5]. These tests meet the requirements of the standardised tests for students with learning difficulties in the Kingdom of Saudi Arabia (KSA). The tests are: (1) reading letters with short vowels, (2) reading words with Sakin letters, and (3) reading words with short vowels and Sakin letters. Each subject was asked to read words in these tests aloud to determine their ability to read the words correctly.

Two measurements were used in this experiment: learning gain and perceived level of usability. Learning gain was assessed directly after finishing the experiment and a learning gain measure was derived (post-test score − pre-test score). These tests included different words from the curriculum and were validated by special education experts. Two types of learning gain were measured in this study. The first type was learning gain of seen words (words included in the training sessions) to evaluate the effectiveness of the adaptation. The second type was learning gain of unseen words (words that were not included in training sessions) to investigate the ability to generalise. The results of learning gain are reported in [6]. In this research, we used the results of learning gain of seen words to investigate the correlation with the perceived level of usability.

Perceived level of usability was measured using the SUS tool, replacing the 1–5 Likert-scale with the Smileyometer. The results of perceived level of usability are reported in [5].

Procedure

The process of this research was conducted in accordance with institutional ethical policy. Signed consent forms from subjects' parents/guardians and primary schools were obtained.

The experiment procedure is illustrated in Fig. 5. In one session, subjects were welcomed and introduced to the experiment's process. Next, their demographic information (grade, age) was collected and diagnostic tests, including the pre-test, were administered. Then, all subjects were randomly assigned to the balanced control or intervention group. The control group used the standard version of the learning system, while the intervention group used the matched version.

The training task was completed in a quiet room inside each subject's school for three weeks during two sessions per week (a total of 6 sessions). Each session lasted approximately 30 min. Each subject worked individually with the system without knowing which experimental condition they had been allocated.

At the end of all the training sessions, subjects were immediately directed to complete the post-test followed by the SUS tool.

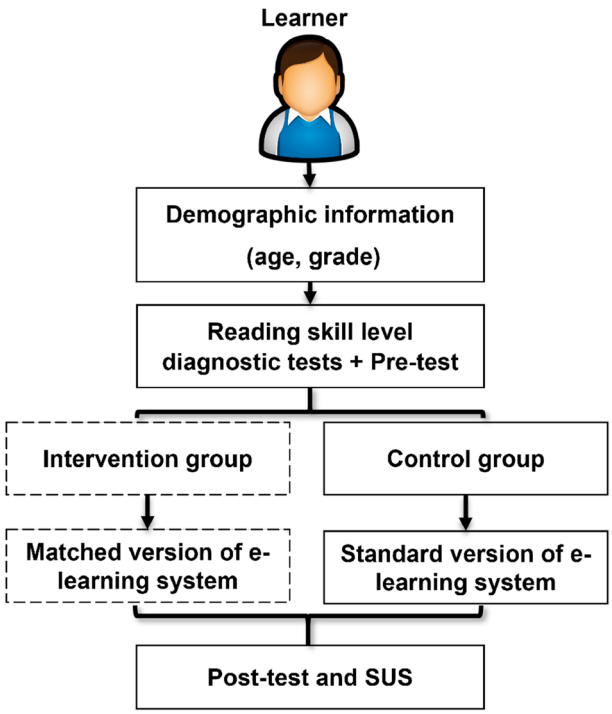

Fig. 5 Experiment procedure

Subjects

Participants who were native speakers of Arabic, officially diagnosed with dyslexia from different primary schools in Jeddah city, KSA participated in this experiment. The sample ($n = 41$) was split into two subgroups, matched for age, grade, reading skill level and prior reading accuracy (from the pre-test) and randomly assigned either to the intervention or the control condition (see Table 1). All subjects had previous experience using electronic devices. Due to the separation of boys and girls in KSA educational institutions, researchers did not have access to male learners. This had the advantage of reducing variance between subjects.

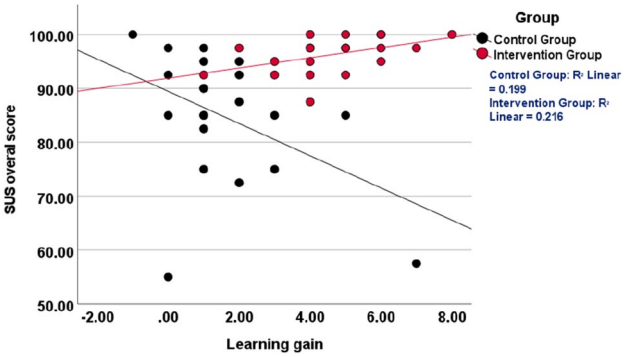

Fig. 6 Scatter-plot of learning gain by SUS overall score by groups

Results

The results were analysed using IBM SPSS Statistical package (version 27).

To assess the relationship between learning gain and perceived level of usability for both groups, a Spearman's correlation was run. As shown in Fig. 6, there was a moderate, positive correlation between learning gain and perceived usability among subjects in the intervention group ($r_s = 0.517, p = 0.02$). In contrast, perceived usability and learning gain were unrelated among subjects in the control group ($r_s = -0.364, p = 0.105$). Therefore, perceived usability was associated with greater learning gain among learners using the matched version of e-learning.

Further investigation was conducted to determine which factors are related to SUS score in each group (i.e., learners' grades), as shown in Fig. 7. In general, among subjects in the intervention group, specifically in Grade 2, there was a moderate but statistically significant positive correlation between perceived usability and learning gain, $r_s = 0.623, p = 0.041$. However, there was no statistically significant correlation among subjects in Grade 3 ($r_s = 0.5, p = 0.5$) and Grade 4 ($r_s = 0.289, p = 0.637$). In contrast, usability and learning gain were unrelated among subjects in the control group in all grades, Grade 2 ($r_s = -0.543, p = 0.105$), Grade 3 ($r_s = -0.687, p = 0.870$) and Grade 4 ($r_s = -0.803, p = 0.102$).

Discussion

The COVID-19 pandemic has forced educational institutions to switch to e-learning [47], making it necessary to evaluate the quality of e-learning systems. A learner's degree of engagement is affected by the learning system's usability, which impacts the learner's motivation to engage in the material [20].

Table 1 Subjects' distribution in terms of grade, age, reading skill level and pre-test

Group	Control group ($N = 21$) Mean (SD)	Intervention group ($N = 20$) Mean (SD)	Groups comparison
Female	21	20	
Second grade	10	11	
Third grade	6	4	
Fourth grade	5	5	
7 Years-old	4	5	
8 Years-old	7	6	
9 Years-old	6	5	
10 Years-old	4	4	
Reading skill level (S1)	3	3	
Reading skill level (S2)	5	6	
Reading skill level (S3)	13	11	
Reading accuracy (pre-test)	4.19 (1.99)	3.9 (1.74)	$U = 196, Z = 0.371, p = 0.711$
Reading accuracy (learning gain)	1.57 (1.80)	4.6 (1.67)	$U = 46.5, Z = 4.307, p = 0.000017$

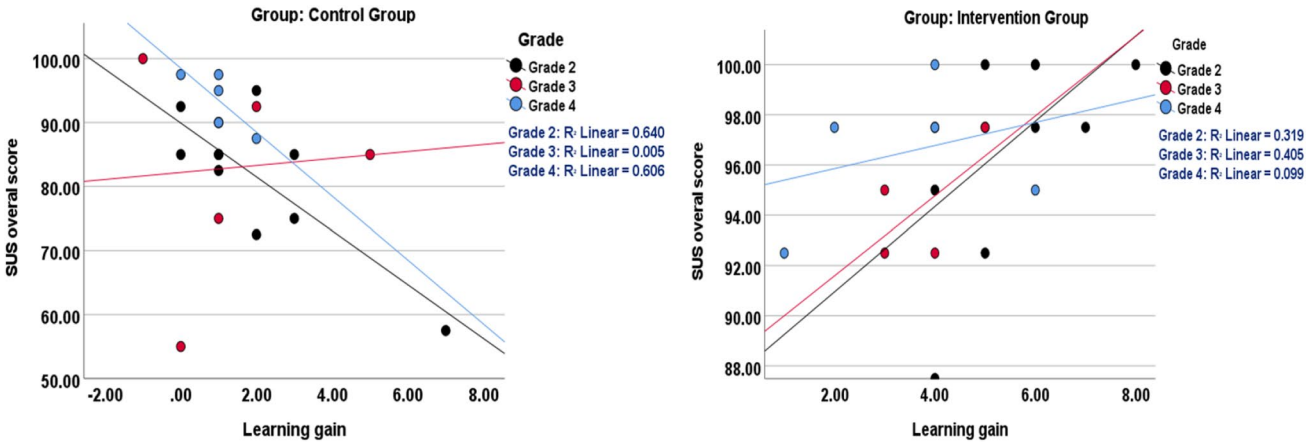

Fig. 7 Scatter-plot of learning gain by SUS score by grade for the control and intervention groups

This research aims to fill a gap in the literature by assessing the viability of using the perceived level of usability when evaluating an adaptive e-learning system from the learners' perspective. This research contributes to the field of adaptive e-learning systems evaluation by revealing the significance of the learner's perception of the usability. Our argument is that the learner will (perhaps not consciously) recognise that their learning is matched to their needs. They will, therefore attach more value to the learning and this will be reflected in the usability scores that they award. This is an effective way to assess whether the adaptation has been effective when two, otherwise identical versions of a system are compared.

These research results indicate that perceived level of usability is correlated to learners' learning gain when the learning material is adapted to their needs. This means that if learners perceive an e-learning system as more usable, their engagement is enhanced, which in turn affects their learning performance. This finding is in line with prior research [5, 10].

Perceived level of usability can also, therefore, be used as an indirect metric for measuring learner engagement. This is parallel to a previous study [35], that measured dropout rate as a proxy for engagement indirectly. This means that learners who are more engaged will use the e-learning system for longer and that this, in turn, leads to improved learning outcomes.

These research results also confirm previous study findings that showed that perceived level of usability significantly correlates with learners' engagement [24]. Learners who perceive the system as easy to use, will be more engaged and motivated.

Furthermore, the results show that perceived level of usability is a key element in evaluating the effectiveness of adaptive e-learning. This confirms previous research findings that showed that a highly usable e-learning system allows learners to focus on the learning process rather than on the functionality of the system [51].

Conclusions and Future Work

This research focussed on evaluating an adaptive e-learning system using perceived level of usability to predict learners' engagement. This work adds to existing knowledge through new insights into perceived level of usability as a significant metric for evaluating these systems' quality by indirectly assessing their engagement and how it can relate to their learning performance.

In future research, it would be useful to conduct a study with other subjects to evaluate the correlation between perceived usability and learning gain, for instance high school and undergraduate learners. Furthermore, it would be useful to investigate whether there is any differentiation in the perceived level of usability in other learning subjects (i.e., math and science). In addition, it would be useful to examine possible relationships between SUS score and age, and male learners.

Acknowledgements Many thanks to the Ministry of Education in Jeddah, the Kingdom of Saudi Arabia, for granting permission to undertake this study in primary schools. We gratefully acknowledge Mrs. Kholoud Alharbi, an assistant in the Department of Special Education in Jeddah, and Mrs. Fatemah Bahasan, a special education teacher in Jeddah, for their wonderful collaboration in facilitating this work. Also, we would like to thank all the participating learners and experts involved in this work.

Funding This work was supported by University of Tabuk in the Kingdom of Saudi Arabia.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Aldabaybah B, Jusoh S. Usability features for arabic assistive technology for dyslexia. In: 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC); 2018. p. 223–228. https://doi.org/10.1109/ICSGRC.2018.8657536.
2. Al-Fraihat D, Joy M, Masa'deh R, Sinclair J. Evaluating e-learning systems success: an empirical study. Comput Hum Behav. 2020;102:67–86. https://doi.org/10.1016/j.chb.2019.08.004.
3. Alghabban WG, Hendley R. The impact of adaptation based on students' dyslexia type: an empirical evaluation of students' satisfaction. In: Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization, UMAP '20 Adjunct. Association for Computing Machinery, New York; 2020. p. 41–46. https://doi.org/10.1145/3386392.3397596.
4. Alghabban WG, Hendley R. Adapting e-learning to dyslexia type: an experimental study to evaluate learning gain and perceived usability. In: Stephanidis C, Harris D, Li WC, Schmorrow DD, Fidopiastis CM, Zaphiris P, Ioannou A, Fang X, Sottilare RA, Schwarz J, editors. HCI international 2020—late breaking papers: cognition, learning and games. Cham: Springer; 2020. p. 519–37. https://doi.org/10.1007/978-3-030-60128-7_39.
5. Alghabban W, Hendley R. Student perception of usability: A metric for evaluating the benefit when adapting e-learning to the needs of students with dyslexia. In: Proceedings of the 13th International Conference on Computer Supported Education—Volume 1: CSEDU. INSTICC, SciTePress; 2021. p. 207–219. https://doi.org/10.5220/0010452802070219.
6. Alghabban WG, Al-Dawsari HM, Hendley R. Understanding the impact on learners' reading performance and behaviour of matching e-learning material to dyslexia type and reading skill level. In: HCI in games: serious and immersive games. Cham: Springer; 2021. p. 135–54. https://doi.org/10.1007/978-3-030-77414-1_11.
7. AlGhannam BA, Albustan SA, Al-Hassan AA, Albustan LA. Towards a standard arabic system usability scale: Psychometric evaluation using communication disorder app. International Journal of Human-Computer Interaction. 2018;34(9):799–804. https://doi.org/10.1080/10447318.2017.1388099.
8. Aljojo N, Munshi A, Almukadi W, Hossain A, Omar N, Aqel B, Almhuemli S, Asirri F, Alshamasi A. Arabic alphabetic puzzle game using eye tracking and chatbot for dyslexia. Int J Interact Mob Technol (iJIM). 2018;12(5):58–80. https://doi.org/10.3991/ijim.v12i5.8957.
9. Alshammari M, Anane R, Hendley RJ. Students' satisfaction in learning style-based adaptation. In: 2015 IEEE 15th international conference on advanced learning technologies; 2015. p. 55–57. https://doi.org/10.1109/ICALT.2015.56.
10. Alshammari M, Anane R, Hendley RJ. Usability and effectiveness evaluation of adaptivity in e-learning systems. In: Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems, CHI EA '16. Association for Computing Machinery, New York; 2016. p. 2984–2991. https://doi.org/10.1145/2851581.2892395.
11. Alsobhi AY, Khan N, Rahanu H. Toward linking dyslexia types and symptoms to the available assistive technologies. In: 2014 IEEE 14th international conference on advanced learning technologies. IEEE; 2014. p. 597–598. https://doi.org/10.1109/ICALT.2014.174.
12. Ardito C, Costabile MF, Marsico MD, Lanzilotti R, Levialdi S, Roselli T, Rossano V. An approach to usability evaluation of e-learning applications. Univ Access Inf Soc. 2006;4(3):270–83. https://doi.org/10.1007/s10209-005-0008-6.
13. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact. 2008;24(6):574–94. https://doi.org/10.1080/10447310802205776.
14. Benmarrakchi F, El Kafi J, Elhore A. Supporting dyslexic's learning style preferences in adaptive virtual learning environment. In: 2016 International conference on engineering MIS (ICEMIS); 2016. p. 1–6. https://doi.org/10.1109/ICEMIS.2016.7745294.
15. Benmarrakchi FE, Kafi JE, Elhore A. Communication technology for users with specific learning disability. Procedia Comput Sci. 2017;110:258–65. https://doi.org/10.1016/j.procs.2017.06.093.
16. Brooke J. Sus—a quick and dirty usability scale. In: Usability evaluation in industry. Oxford: Taylor & Francis; 1996. p. 189–94.
17. Brusilovsky P. Adaptive hypermedia for education and training. In: Durlach PJ, Lesgold AM, editors. Adaptive technologies for training and education, vol. 46. Cambridge University Press Cambridge; 2012. p. 46–68. https://doi.org/10.1007/978-3-540-72079-9_1.
18. Burac MAP, Cruz JD. Development and usability evaluation on individualized reading enhancing application for dyslexia (IREAD): A mobile assistive application. IOP Conf Ser Mater Sci Eng. 2020;803:1–7. https://doi.org/10.1088/1757-899X/803/1/012015.
19. Carini RM, Kuh GD, Klein SP. Student engagement and student learning: testing the linkages*. Res High Educ. 2006;47:1–32. https://doi.org/10.1007/s11162-005-8150-9.
20. Chin SPP, Tsui E, Lee CS. Enhancing learning effectiveness by adopting a knowledge-based usability guidelines. VINE J Inf Knowl Manag Syst. 2016. https://doi.org/10.1108/VJIKMS-02-2014-0015.
21. Chorfi H, Jemni M. Perso: towards an adaptive e-learning system. J Interact Learn Res. 2004;15(4):433–447. https://www.learntechlib.org/p/18900.
22. Chrysafiadi K, Virvou M. Student modeling approaches: a literature review for the last decade. Expert Syst Appl. 2013;40(11):4715–29. https://doi.org/10.1016/j.eswa.2013.02.007.
23. Cidral WA, Oliveira T, Di Felice M, Aparicio M. E-learning success determinants: Brazilian empirical study. Comput Educ. 2018;122:273–90. https://doi.org/10.1016/j.compedu.2017.12.001.
24. Dahleez KA, El-Saleh AA, Alawi AMA, Abdelfattah FA. Higher education student engagement in times of pandemic: the role of e-learning system usability and teacher behavior. Int J Educ Manag. 2021. https://doi.org/10.1108/IJEM-04-2021-0120.
25. Daly EJ III, Martens BK, Kilmer A, Massie DR. The effects of instructional match and content overlap on generalized reading performance. J Appl Behav Anal. 1996;29(4):507–18. https://doi.org/10.1901/jaba.1996.29-507.
26. Dixson MD. Measuring student engagement in the online course: the online student engagement scale (ose). Online Learn. 2015. https://doi.org/10.24059/olj.v19i4.561.
27. Elbeheri G, Everatt J. Literacy ability and phonological processing skills amongst dyslexic and non-dyslexic speakers of Arabic. Read Writ. 2007;20:273–94. https://doi.org/10.1007/s11145-006-9031-0.
28. Finstad K. The usability metric for user experience. Interact Comput. 2010;22(5):323–7. https://doi.org/10.1016/j.intcom.2010.04.004.
29. Fletcher JM, Lyon GR, Fuchs LS, Barnes MA. Learning disabilities: from identification to intervention. 2nd ed. New York: Guilford Publications; 2018.

30. Fredricks JA, McColskey W. The measurement of student engagement: a comparative analysis of various methods and student self-report instruments. US, Boston: Springer; 2012. p. 763–82. https://doi.org/10.1007/978-1-4614-2018-7_37.
31. Friedmann N, Coltheart M. Types of developmental dyslexia. In: Bar-On A, Ravid D, editors. Handbook of communication disorders: theoretical, empirical, and applied linguistics perspectives. Berlin: De Gruyter Mouton; 2016. p. 1–37.
32. Gabrieli JDE. Dyslexia: a new synergy between education and cognitive neuroscience. Science. 2009;325(5938):280–3. https://doi.org/10.1126/science.1171999.
33. Gatian AW. Is user satisfaction a valid measure of system effectiveness? Inf Manag. 1994;26(3):119–31. https://doi.org/10.1016/0378-7206(94)90036-1.
34. Gena C. Methods and techniques for the evaluation of user-adaptive systems. Knowl Eng Rev. 2005;20(1):1–37. https://doi.org/10.1017/S0269888905000299.
35. Ghaban W, Hendley R. Investigating the interaction between personalities and the benefit of gamification. In: Proceedings of the 32nd international bcs human computer interaction conference vol. 32; 2018. p. 1–13. https://doi.org/10.14236/ewic/HCI2018.41.
36. Hamid SSA, Admodisastro N, Manshor N, Ghani AAA, Kamaruddin A. Engagement prediction in the adaptive learning model for students with dyslexia. In: Proceedings of the 4th international conference on human-computer interaction and user experience in Indonesia, CHIuXiD '18. Association for Computing Machinery, New York; 2018. p. 66–73. https://doi.org/10.1145/3205946.3205956.
37. Harrati N, Bouchrika I, Tari A, Ladjailia A. Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis. Comput Hum Behav. 2016;61:463–71. https://doi.org/10.1016/j.chb.2016.03.051.
38. Islas E, Pérez M, Rodriguez G, Paredes I, Ávila I, Mendoza M. E-learning tools evaluation and roadmap development for an electrical utility. J Theor Appl Electron Commer Res. 2007;2(1):63–75. https://doi.org/10.3390/jtaer2010006.
39. ISO: ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs): Part 11: Guidance on usability; 1998.
40. Jung HJ. Ubiquitous learning: determinants impacting learners' satisfaction and performance with smartphones. Lang Learn Technol. 2014;18(3):97–119.
41. Kozlowski SW, Gully SM, Brown KG, Salas E, Smith EM, Nason ER. Effects of training goals and goal orientation traits on multidimensional training outcomes and performance adaptability. Organ Behav Hum Decis Process. 2001;85(1):1–31. https://doi.org/10.1006/obhd.2000.2930.
42. Kulaglić S, Mujačić S, Serdarević IK, Kasapović S. Influence of learning styles on improving efficiency of adaptive educational hypermedia systems. In: 2013 12th International conference on information technology based higher education and training (ITHET); 2013. p. 1–7. https://doi.org/10.1109/ITHET.2013.6671060
43. Kuo YC, Walker AE, Belland BR, Schroder KE. A predictive study of student satisfaction in online education programs. Int Rev Res Open Distrib Learn. 2013;14(1):16–39. https://doi.org/10.19173/irrodl.v14i1.1338.
44. Lo YY, Cooke NL, Starling ALP. Using a repeated reading program to improve generalization of oral reading fluency. Educ Treat Child. 2014;34(1):115–140. http://www.jstor.org/stable/42900103
45. Lyon GR, Shaywitz SE, Shaywitz BA. A definition of dyslexia. Ann Dyslexia. 2003;53(1):1–14. https://doi.org/10.1007/s11881-003-0001-9.
46. Mahdavi I, Fazlollaht H, Heidarzade A, Mahdavi-Am N, Rooshan Y. A heuristic methodology for multi-criteria evaluation of web-based e-learning systems based on user satisfaction. J Appl Sci. 2008;8(24):4603–9. https://doi.org/10.3923/jas.2008.4603.4609.
47. Mallillin LLD, Mendoza LC, Mallillin JB, Felix RC, Lipayon IC. Implementation and readiness of online learning pedagogy: a transition to Covid 19 pandemic. Eur J Open Educ E-learn Stud. 2020. https://doi.org/10.46827/ejoe.v5i2.3321.
48. Maravanyika M, Dlodlo N, Jere N. An adaptive recommender-system based framework for personalised teaching and learning on e-learning platforms. In: 2017 IST-Africa Week Conference (IST-Africa); 2017. p. 1–9. https://doi.org/10.23919/ISTAFRICA.2017.8102297.
49. Murchland S, Kernot J, Parkyn H. Children's satisfaction with assistive technology solutions for schoolwork using the quest 2.1: children's version. Assist Technol. 2011;23(3):162–76. https://doi.org/10.1080/10400435.2011.588990.
50. Nist L, Joseph LM. Effectiveness and efficiency of flashcard drill instructional methods on urban first-graders' word recognition, acquisition, maintenance, and generalization. Sch Psychol Rev. 2008;37(3):294–308. https://doi.org/10.1080/02796015.2008.12087877.
51. Orfanou K, Tselios N, Katsanos C. Perceived usability evaluation of learning management systems: empirical evaluation of the system usability scale. Int Rev Res Open Distrib Learn. 2015;16(2):227–46. https://doi.org/10.19173/irrodl.v16i2.1955.
52. Pal D, Vanijja V. Perceived usability evaluation of microsoft teams as an online learning platform during covid-19 using system usability scale and technology acceptance model in india. Child Youth Serv Rev. 2020;119:1–12. https://doi.org/10.1016/j.childyouth.2020.105535.
53. Paramythis A, Weibelzahl S, Masthoff J. Layered evaluation of interactive adaptive systems: framework and formative methods. User Model User-Adap Inter. 2010;20(5):383–453. https://doi.org/10.1007/s11257-010-9082-4.
54. Peres SC, Pham T, Phillips R. Validation of the system usability scale (sus): Sus in the wild. Proc Hum Fact Ergon Soc Annu Meet. 2013;57(1):192–6. https://doi.org/10.1177/1541931213571043.
55. Rajabalee YB, Santally MI. Learner satisfaction, engagement and performances in an online module: implications for institutional e-learning policy. Educ Inf Technol. 2021;26(3):2623–56. https://doi.org/10.1007/s10639-020-10375-1.
56. Read JC, MacFarlane S. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In: Proceedings of the 2006 conference on interaction design and children, IDC '06. Association for Computing Machinery, New York; 2006. p. 81–88. https://doi.org/10.1145/1139073.1139096
57. Rodrigues H, Almeida F, Figueiredo V, Lopes SL. Tracking e-learning through published papers: a systematic review. Comput Educ. 2019;136:87–98. https://doi.org/10.1016/j.compedu.2019.03.007.
58. Sampson P, Leonard J, Ballenger J, Coleman JC. Student satisfaction of online courses for educational leadership. Online J Distance Learn Administr. 2010. https://eric.ed.gov/?id=EJ914153
59. Shee DY, Wang YS. Multi-criteria evaluation of the web-based e-learning system: a methodology based on learner satisfaction and its applications. Comput Educ. 2008;50(3):894–905. https://doi.org/10.1016/j.compedu.2006.09.005.
60. Shields BJ, Palermo TM, Powers JD, Grewe SD, Smith GA. Predictors of a child's ability to use a visual analogue scale. Child Care Health Dev. 2003;29(4):281–90. https://doi.org/10.1046/j.1365-2214.2003.00343.x.
61. Sihombing JH, Laksitowening KA, Darwiyanto E. Personalized e-learning content based on felder-silverman learning style model. In: 2020 8th International conference on information and

communication technology (ICoICT); 2020. p. 1–6. https://doi.org/10.1109/ICoICT49345.2020.9166452
62. Spencer K. Predicting word-spelling difficulty in 7- to 11-year-olds. J Res Read. 1999;22(3):283–92. https://doi.org/10.1111/1467-9817.00091.
63. Tseng JC, Chu HC, Hwang GJ, Tsai CC. Development of an adaptive learning system with two sources of personalization information. Comput Educ. 2008;51(2):776–86. https://doi.org/10.1016/j.compedu.2007.08.002.
64. Tullis TS, Stetson JN. A comparison of questionnaires for assessing website usability. In: Usability professional association conference, vol. 1. Minneapolis; 2004. p. 1–12.
65. Wang YS. Assessment of learner satisfaction with asynchronous electronic learning systems. Inf Manag. 2003;41(1):75–86. https://doi.org/10.1016/S0378-7206(03)00028-4.
66. Xu D, Wang H, Su K. Intelligent student profiling with fuzzy models. In: Proceedings of the 35th Annual Hawaii international conference on system sciences; 2002. p. 1–8. https://doi.org/10.1109/HICSS.2002.994005.
67. Zaharias P. A usability evaluation method for e-learning: focus on motivation to learn. Association for Computing Machinery, New York; 2006. p. 1571-1576. https://doi.org/10.1145/1125451.1125738
68. Zaharias P, Poylymenakou A. Developing a usability evaluation method for e-learning applications: beyond functional usability. Int J Hum Comput Interact. 2009;25(1):75–98. https://doi.org/10.1080/10447310802546716.