Student Perception of Usability: A Metric for Evaluating the Benefit When Adapting e-Learning to the Needs of Students with Dyslexia

Weam Gaoud Alghabban[1,2][a] and Robert Hendley[1][b]
[1]School of Computer Science, University of Birmingham, Birmingham, U.K.
[2]Computer Science Department, Alwajh University College, University of Tabuk, Tabuk, Saudi Arabia

Keywords: e-Learning, Student Needs, Engagement, Perceived Usability, System Usability Scale, Dyslexia, Arabic, Skill Level.

Abstract: Usability is now widely recognised as a critical factor to the success of e-learning systems. A highly usable e-learning system increases students' satisfaction and engagement, thereby enhancing learning performance. However, one challenge in e-learning is poor engagement arising from a "one-size-fits-all" approach that presents learning content and activities in the same way to all students. Each student has different characteristics and, therefore, the content should be sensitive to these differences. This study evaluated the students' perceived level of usability of an e-learning system that matches content to reading skill levels of students with dyslexia. 41 students rated their perceived usability of an e-learning system using the System Usability Scale (SUS). Results indicated that when the e-learning system matches content to students' skill level, students perceive greater usability than when the learning is not matched. There was also a moderate, positive correlation between perceived usability and learning gain when e-learning was matched to their skill level. Thus, students assessment of the usability of a system is affected by the degree to which it is suited to their needs. This may be reflected in increased engagement and is associated with higher learning gain.

1 INTRODUCTION

The context of learning is changing radically. The process of teaching and learning is no longer restricted to a traditional classroom environment, due to the advent of electronic learning (e-learning) technologies (Pal and Vanijja, 2020). E-learning, which refers to the use of electronic technologies in educational settings, includes the delivery of educational content via Internet, audio, video and other media (Ozkan and Koseler, 2009). E-learning has become a significant development in the educational field and has been shown to improve the quality of learning (Hamidi and Chavoshi, 2018). Although e-learning is not new, it has not yet been widely used as the primary mode of instruction (Hidalgo et al., 2020). However, due to institutional closures precipitated by the Coronavirus (COVID-19) pandemic, educational institutions worldwide have dramatically increased their reliance on e-learning technologies. Existing literature suggests students' perceptions toward e-learning is mostly positive (Alqurashi, 2019; Rodrigues et al., 2019; Valencia-Arias et al., 2019). However, these studies were conducted when e-learning platforms merely supplemented traditional classroom teaching. With the onset of COVID-19, most educational services have been delivered electronically. Because educators are increasingly reliant upon e-learning technologies, assessing the effectiveness and perceived usability of e-learning environments is a critical task for researchers (Höök, 2000). If an e-learning system is not sufficiently usable, students will become demotivated and retain less material because they have to focus on system functionality rather than content (Ardito et al., 2006).

Engagement with course material has a significant impact on students' learning (Kori et al., 2016). One challenge in e-learning is to keep students motivated and engaged (Moubayed et al., 2020). A "one-size-fits-all" approach that does not consider individual differences in learning and teaching produces poor student engagement (Maravanyika et al., 2017). Since different people learn in different ways, one method to keep each student engaged is to adapt the content to their particular preferences and needs in order to

[a] https://orcid.org/0000-0003-0857-1548
[b] https://orcid.org/0000-0001-7079-3358

Alghabban, W. and Hendley, R.
Student Perception of Usability: A Metric for Evaluating the Benefit When Adapting e-Learning to the Needs of Students with Dyslexia.
DOI: 10.5220/0010452802070219
In *Proceedings of the 13th International Conference on Computer Supported Education (CSEDU 2021) - Volume 1*, pages 207-219
ISBN: 978-989-758-502-9

maximise learning, a process known as adaptive e-learning (Moubayed et al., 2020). As a result, adaptive e-learning systems are becoming increasingly popular (Hariyanto and Köhler, 2020). Making e-learning systems available, effective, and engaging requires an understanding of the target users (Liaw et al., 2007). Thus, students' preferences, cognitive abilities and cultural background should be thoroughly considered when designing e-learning systems (Ardito et al., 2006). In addition, students' interactions with these systems should be as intuitive as possible (Ardito et al., 2006), providing flexibility and adapting to students' unique traits (Brusilovsky and Millán, 2007). In summary, providing positive and appropriate user experiences should be one of the main goals of an e-learning system (Ardito et al., 2006).

One critical component of user experience is usability (Diefenbach et al., 2014) and, especially, a student's *perception* of the usability. Usability refers to how easily and efficiently end users can accomplish their goals within a system (Nielsen, 1993). Usability is therefore a key attribute of software quality and is as important as security, robustness and performance (Harrati et al., 2016). In the context of e-learning, a good user experience leads to better engagement and satisfaction, which in turn increases the likelihood it will help students achieve learning goals. All students, regardless of their orientation, experience or background, should be able to readily utilise e-learning tools to achieve their goals (Harrati et al., 2016).

E-learning systems have been evaluated from various perspectives such as evaluating students' satisfaction (Alghabban and Hendley, 2020b), engagement and interaction (Moubayed et al., 2020), and learning quality and effectiveness (Alghabban and Hendley, 2020a; Rodrigues et al., 2019). However; there is a lack of research on the examination of the usability of e-learning (Pal and Vanijja, 2020). Because students are key users of e-learning environments (Limayem and Cheung, 2008), their needs and differences must be the main focus of usability practice in the e-learning environment (Ardito et al., 2006). However, recent research by Rodrigues et al. (Rodrigues et al., 2019) shows that usability studies of e-learning largely ignore the perspectives of students. This is a crucial shortcoming, as poor usability may reduce students' motivation to learn.

Another underexamined issue in e-learning is the sensitivity of content to students' needs and difficulties (Zaharias, 2009). Prior studies have shown students are more satisfied when content is effectively presented, well-organised, flexible and useful (Holsapple and Lee-Post, 2006). However, according to a recent literature review by Gunesekera et al. (Gunesekera et al., 2019), little research has examined student-content interaction in the context of e-learning. This is especially true in the context of special needs education, which is the domain in which this work is focused. Therefore, there is a need to evaluate the perceived level of usability of e-learning from the perspective of students when content supports their needs. Accordingly, the aims of this research are:

1. To evaluate the perceived level of usability of an e-learning system when matching content to students' needs. The evaluation was done using a standard usability assessment tool (the System Usability Scale (SUS) questionnaire).
2. To determine whether there is a relationship between the perceived level of usability and learning gain when matching content to students' needs.

The remainder of this paper is structured as follows. Section 2 introduces the related work. Section 3 describes the research methodology. Results are presented in Section 4, followed by the discussion in Section 5. Finally, the conclusion and further work are presented in Section 6.

2 RELATED WORK

In this section we explore three key aspects: e-learning; assessing student engagement in current research, and finally, dyslexia and the difficulties experienced in learning and the current state of the art in evaluating the perceived usability of e-learning systems of students with dyslexia.

2.1 e-Learning

E-learning has expanded rapidly in all academic institutions; however, there is little consensus on what e-learning means. A systematic literature review by Rodrigues et al. (Rodrigues et al., 2019) defined e-learning 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). Much research emphasizes the benefits of e-learning systems. The expansion of location- and time-independent education and high-speed Internet access have allowed educational institutions to increasingly use e-learning technologies (Ozkan and Koseler, 2009).

However, one challenge in e-learning systems is a "one-size-fits-all" approach that does not consider individual differences in learning and teaching (Allen et al., 2004). Since different people learn in different ways, it is important to match the content to the student's particular preferences and needs in order to maximize and speed up the learning process (Moubayed et al., 2020). As a result, adaptive e-learning systems are becoming increasingly popular, which provides new flexibility and suitability to the peculiarities of the students' specific needs (Hariyanto and Köhler, 2020).

In the context of e-learning systems, adaptation is defined as a procedure for tailoring the educational environment to learners' differences (Brusilovsky, 2012) with the aim of improving learning outcome (Maravanyika et al., 2017). Different student characteristics can be considered. For example, PERSO adapts course material based on students' media preferences and their knowledge (Chorfi and Jemni, 2004). Other studies have attempted to incorporate learning styles as a parameter for the adaptation of the learning process. For example, both (Sihombing et al., 2020) and (Alshammari et al., 2016) adapt based on students' learning style. Another characteristic is students' personality, as reported in (Ghaban and Hendley, 2018).

2.2 Evaluating Student Engagement

The multidisciplinary nature of e-learning systems leads scholars from different fields to evaluate the effectiveness of these systems. Some researchers have focused on evaluating e-learning systems from a technical perspective (Islas et al., 2007). Meanwhile, others have evaluated these systems from the human-factor perspective, taking into account students' satisfaction (Alghabban and Hendley, 2020b), experience (Gilbert et al., 2007), determining e-learning materials' effectiveness (Douglas and Vyver, 2004), and engagement and motivation (Moubayed et al., 2020; Ghaban and Hendley, 2018).

As part of these evaluation metrics, engagement with course material has a significant impact on students' learning (Kori et al., 2016). Student engagement is defined as the student's ongoing effort spent on the learning process to achieve learning goals (Coates, 2006). That is, student engagement is comprised of three categories (Connell et al., 1995): behavioral such as sustained concentration in learning, effort and persistent learning; emotional such as excitement and interest in learning; and psychological such as independence and challenge preferences (Appleton et al., 2006).

With the emergence of e-learning, the focus of earlier research was mainly on how to improve the learning performance and achievement of students. However, lately more studies have focused on students' engagement when interacting with e-learning systems (Ghaban and Hendley, 2018; Lee et al., 2019; Alshammari, 2019). According to Carini et al. (Carini et al., 2006), if the engagement level of students increases, this may serve as a strong predictor of improved performance and achievement. Therefore, engaging more with an e-learning system, can lead the student to use it for longer and also to learn more effectively, which in turn can enhance the students' academic performance. Hence, increasing a student's engagement with their e-learning environment is an important objective.

Many approaches are used by researchers to measure engagement. These include self-report, teacher expert knowledge (Lloyd et al., 2007) and interviews and observation (Fredricks and McColskey, 2012). However, these methods take time and resources and can be difficult to implement without influencing the outcome. Therefore, different methods are often used to measure engagement indirectly. For instance, Ghaban and Hendley (Ghaban and Hendley, 2018) found that students who are more motivated and engaged, will use the e-learning system for longer (They do this by measuring the drop out rate as a proxy for engagement) and that this, in turn, leads to improved learning outcomes. Kangas et al. (Kangas et al., 2017) found that students' satisfaction is an important indicator of their engagement. That is, if a student is more engaged, then his/her interest in and enjoyment of learning is increased too. Also, if a student is more engaged and motivated, then their perception of the system's level of usability will be higher (Ardito et al., 2006; Zaharias and Poylymenakou, 2009).

In this work we measure students assessment of the usability of the e-learning system and then assess: i) whether this varies when the system is adapted to their needs, and ii) whether their learning gain is higher when their perception of the quality of the system is higher. We argue that this is an effective way to assess the benefit of adaptation.

2.3 Dyslexia

The usage of e-learning systems is a growing research area for children with specific learning disabilities. One of the most carefully studied and most common of childhood learning disabilities is dyslexia (Ziegler et al., 2003). It is identified in 80% of the learning disabilities population (Lerner, 1989). Dyslexia, which is called 'reading disorder', has been defined by the main international classification, ICD-10 (WHO, 1992) as ”The main feature of this disorder is a specific and significant impairment in the development of reading skills, which is not solely accounted for by mental age, visual acuity problems, or inadequate schooling” (p. 245). It may affect the following skills: reading word recognition, reading comprehension skill, the performance of tasks requiring reading and oral reading skill (WHO, 1992). The reading performance of children with dyslexia is significantly below the expected level on the basis of age, intelligence, and school placement (Lyon et al., 2003).

Students with dyslexia face other difficulties such as high avoidance, frustration and poor concentration (Oga and Haron, 2012). Additionally, they frequently suffer from poor school attendance, problems with social adjustment and academic failure (WHO, 1992). They also have varied needs, abilities and characteristics, which needs to consider individually. Therefore, it is argued e-learning systems should be responsive to these differences (Brusilovsky and Millán, 2007) rather than treating all students uniformly. One feature that may be considered is *dyslexia type* as reported in (Alghabban and Hendley, 2020a) and learning style (Benmarrakchi et al., 2017b).

Students with dyslexia can become both easily engaged and disengaged with their learning content. They may engage with the games (Vasalou et al., 2017) and collaborative learning environments (Pang and Jen, 2018), and easily disengaged with too much learning content (Baker and Rossi, 2013). Their lack of engagement becomes one of the main causes that affect their academic performance (Sahari and Johari, 2012). Therefore, it is important to measure students' engagement in order to be able to retain their attention and therefore to improve their learning performance.

Many e-learning systems have been designed for students with dyslexia. Al-Ghurair and Alnaqi (Al-Ghurair and Alnaqi, 2019) aimed to enhance short-term memory of students with dyslexia by developing a story theme in a game-based application. The game's usability was evaluated by experts who observed children using the application and children's opinions were taken after using that system. Similar work has been done by Aljojo et al. (Aljojo et al., 2018) where they assist children with dyslexia in pronouncing Arabic letters using a puzzle game-based system. The children evaluated the system's usability in terms of learnability, efficiency, memorability, errors and satisfaction. Their findings were not clear nor properly discussed. Further, Vasalou et al. (Vasalou et al., 2017) aimed to enhance spelling, word decoding, and the fluency of students with dyslexia (in English) by developing a game-based learning application.

Burac and Cruz (Burac and Cruz, 2020) developed a mobile application which implements a text-to-speech feature to facilitate reading among children with dyslexia in English. The overall usability of the application was assessed by teachers by using a usability questionnaire that measured effectiveness, efficiency, and quality of support, ease of learning and satisfaction. Results showed that the usability of the application was excellent. Similarly, Aldabaybah and Jusoh (Aldabaybah and Jusoh, 2018) developed an assistive technology for dyslexia by empirically identifying a set of accessibility features related to the use of menus, colours, navigation and feedback.

Despite all these works, some gaps are evident. First, many systems fail to account for the variations in abilities and skills among individuals with dyslexia (Al-Ghurair and Alnaqi, 2019; Aljojo et al., 2018; Burac and Cruz, 2020; Vasalou et al., 2017). As a result, students with dyslexia using these systems may exhibit lower engagement and satisfaction (Alghabban and Hendley, 2020b). Second, much research has utilised teacher evaluations of e-learning systems, ignoring the perspective of the students (Burac and Cruz, 2020; Aldabaybah and Jusoh, 2018; Aljojo et al., 2018). Thus, whether matching content to reading skill level improves perceived usability of e-learning for students with dyslexia remains unknown. The urgency of this question is exacerbated by the COVID-19 pandemic, which has obstructed access to traditional special needs provision.

To address these gaps, this research aims to evaluate the *perception* of the usability of an e-learning system by students with dyslexia and to compare this across two conditions: i) when it is matched to their needs (specifically their reading skill level) and ii) when it is not.

3 METHODOLOGY

An empirical study was conducted to evaluate the perceived level of usability when matching e-learning content to students' characteristics, specifically reading skill level among Arabic children with dyslexia. Details about the study's hypotheses, the proposed e-learning system, data collection and experimental procedure are presented in following sections.

3.1 Study Hypotheses and Variables

For this study, two hypotheses were formulated:

1. Matching learning content to reading skill level of students with dyslexia achieves significantly bet-

ter perceived level of usability compared to non-matched content.

2. There is a positive correlation between the perceived level of usability and learning gain when matching learning content to the reading skill level of students with dyslexia.

Students with dyslexia in this study were divided into two groups. Students in the experimental group interacted with a matched version of the e-learning system that matches content to each student's reading skill level. Students in the control group interacted with a standard version of the system with identical layout and interface to the matched version but without matching the content to their reading skill level. The perceived level of usability and learning gain were the main dependent variables measured in this study. Additionally, this study investigated whether there was a correlation between the perceived usability and learning gain. The effect of other adaptations and the effect on learning gain are reported elsewhere.

3.2 e-Learning System

The Reading Enhancing System for Dyslexia (RESD) was designed in order to support the evaluation of the perceived usability. The RESD system is a dynamic, web-based e-learning system for Arabic children with dyslexia in elementary schools. It provides different word recognition exercises, as word recognition is a strong predictor of reading fluency and improves students' ability to comprehend novel passages (Lo et al., 2011). There are two versions of the RESD system to support the experimental conditions. The matched version matches training content of exercises to the reading skill level of each student, while the standard version fixes the content to suit all reading skill levels. Both versions of the system are identical in layout; the only difference is the provided content in the exercises.

The system trained students by providing a number of different word reading recognition activities. In total, there were six training sessions, each with 20 activities. The content of training sessions was derived from the study school curriculum and targeted three main reading skills. These skills are the fundamental literacy skills that serve as building blocks for more advanced skills and are presented in Table 1. The design of matching content in the RESD system is presented in Figure 1. The system relies on a reliable, offline reading skill level diagnostic tool, as described in section 3.3. Students accessed training content matching their reading skill level, according to the algorithm presented in Figure 2. For example, if a student's reading skill required them to master reading letters with short vowels, then the exercises assessed and reinforced this skill.

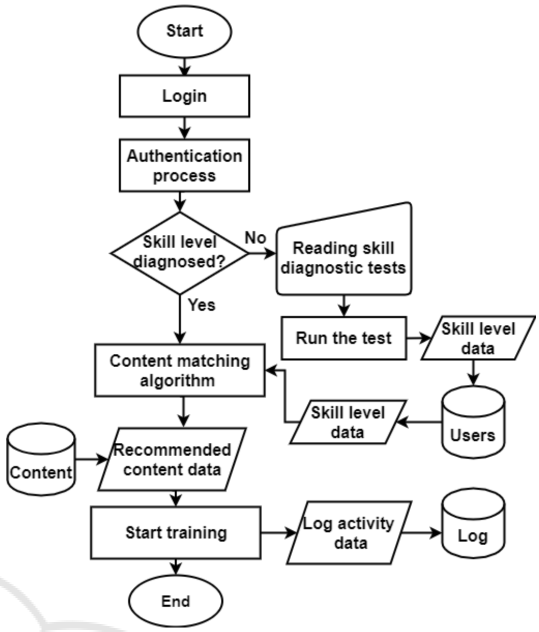

Figure 1: Design of matching content based upon dyslexia reading skill level in the RESD system.

A screenshot of a training activity of the system is presented in Figure 3. As shown in Figure 3, the question was presented at the top of the screen in written form and accompanied by audio. The target word was missing and denoted as dots inside a rectangular box under the query, and its audio had a question mark icon beside it. The student listened to the query and selected the correct target word among three choices, of which only one was correct. If the student chooses the correct response, the interface will provide spoken and written praise. If the student chooses an incorrect response, negative written and spoken feedback was provided. Feedback on training progress is provided as a progress bar at the top of the screen as shown in Figure 3. Because motivation increases confidence and self-esteem of students with dyslexia (Benmarrakchi et al., 2017a), motivational messages were presented after completing every seven activities, as shown in Figure 4.

The training sessions' layout was consistent for all reading skills, keeping the question, target word, audio icon, and answer choices in the same place. The user interface followed best design practices in terms of font, type, colour, and background colour for e-learning systems targeting Arabic dyslexia (AlRowais et al., 2013) and used encouraging spoken feedback.

Table 1: The reading skills and content represented in the RESD system.

No.	Reading skill	Content
1	Reading letters with short vowels	Reading letters with fat-ha /a/
		Reading letters with kasra /i/
		Reading letters with damma /u/
2	Reading words with consonant sections	Reading one consonant section in 2-letters words (consonant section in the end of a word)
		Reading one consonant section in 3-letters words (consonant section could be in the middle or the end of a word)
		Reading two consonant sections in 4-letters words (consonant sections are in the middle and end of a word)
3	Reading words with short vowels and consonant sections	Reading 3-letters words with fat-ha /a/
		Reading 3-letters words with fat-ha /a/ and kasra /i/
		Reading 3-letters words with fat-ha /a/ and damma /u/
		Reading 3-letters words with fat-ha /a/, kasra /i/ and damma /u/
		Reading 4-letters words with all short vowels and consonant sections

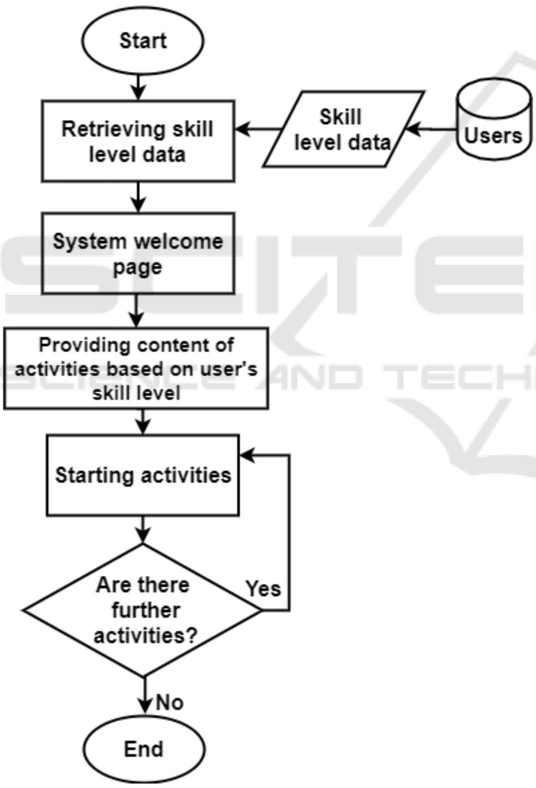

Figure 2: The proposed content matching algorithm based upon reading skill level.

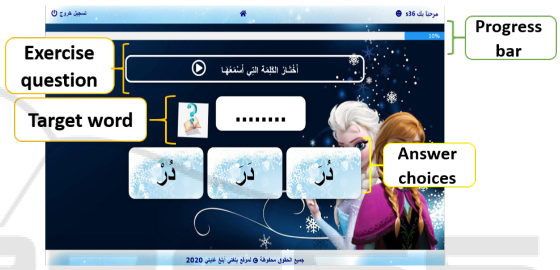

Figure 3: A screenshot of the RESD system.

Figure 4: A motivational message.

3.3 Data Collection

Data was gathered by using three tools: reading skill level diagnostic tests, pre- and post-tests, and the SUS questionnaire.

Reading skill level diagnostic tests consisted of three subtests: (1) letter reading with short vowels, (2) word reading with consonant sections and (3) word reading with short vowels and consonant sections. These tests met the requirements of the reliable and standardised tests approved by the Ministry of Education in Saudi Arabia (SA) for students with learning difficulties (Bukhari et al., 2016). The skill level test for letter reading tested the student's ability to accurately recognise letters (spelling) in three-letter words with short vowels. The skill level test for words reading with consonant sections, included a list of ten vowelised words with consonant sections within two and three letters-words. The skill level test for word reading with short vowels and consonant sections included a list of ten vowelised words used to assess the students' level of this skill. The words differed in the number of letters and using the combination of short vowels and consonants. For each reading skill level test, each participant was asked to read these words aloud to detect his or her ability to spell letters correctly (as in the first test) and to read the words correctly (as in the second and third tests).

Pre- and post-tests are commonly used to derive learning gain, or changes in learning outcomes that have been achieved after a specific intervention (Pickering, 2017). The pre-test contained 15 different vowelised words from the curriculum and was validated by special educational experts. The pre-test met the standardised test requirements for students with learning difficulties (Bukhari et al., 2016). The pre-test was used at the beginning of the study to assess initial reading performance of students and participants were divided into two groups, balanced by the pre-test scores. The post-test, which is the same as the pre-test, is used upon completion of the intervention to measure learning gain. Learning gain is calculated by subtracting the score of the pre-test from the score of the post-test. Using the same tests in this study was the best choice to allow a precise comparison of reading abilities (Bonacina et al., 2015).

SUS was used in this study to assess children' perception of usability for the following reasons. The SUS tool is one of the most widely used instruments for Human-Computer Interaction (HCI) researchers to evaluate the perceived usability of systems (Harrati et al., 2016). It has a high degree of reliability and validity (Bangor et al., 2008), and can be adapted for various contexts (Peres et al., 2013). It is a valid instrument to compare the usability of two or more systems (Peres et al., 2013). With SUS, reliable results are evident even with small samples (Tullis and Stetson, 2004). It has also been adapted to evaluate the usability by children (Alghabban and Hendley, 2020a; Putnam et al., 2020) and also used as a standard usability tool for Arabic users with a high degree of reliability (AlGhannam et al., 2018).

The SUS includes 10 mixed-tone items. The odd numbered items have a positive tone, and the even number a negative tone. All items are on a five-point Likert scale from strongly disagree (1) to strongly agree (5). SUS uses alternative positive and negative questions.

Several pictorial elements are used to convey the meaning of the Likert scale for children (Read, 2012; Putnam et al., 2020). A widely-used visual representation in academia is the Smileyometer (Read, 2012), as shown in Figure 5. In this study, the Likert scale used by SUS were converted to Smileyometers, since a visual representation of the Likert scale of agreement matches children's cognitive abilities (Alghabban and Hendley, 2020a; Putnam et al., 2020).

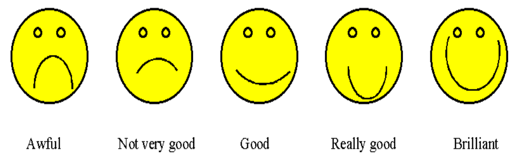

Figure 5: The Smileyometer (Read, 2012).

3.4 Procedure

The study was subject to ethical approval by University of Birmingham, and for each child explicit consent was obtained from parents/guardians and schools, prior to study participation. All children were selected from elementary schools in SA.

The procedure timetable is presented in Table 2. In one session, children were welcomed and introduced to the study's objectives. Then, their demographic information (age, grade) was collected and reading skill level diagnostic tests, including the reading pre-test, were administered. Subsequently, children were divided into two groups: the experimental group and the control group. The two groups had similar distribution of age, grade and prior reading performance assessed by the pre-test. The experimental group used the matched version of the RESD, and the control group used the standard version. The experiment took place in person, in a quiet room, within each child's school. Each child individually interacted with the system. Six training sessions were conducted, two sessions each week, each one lasted for 30 minutes.

At the end of the total training sessions, the post-test and SUS were administered.

4 RESULTS

4.1 Participants

The participants were all female students, aged 7-10 years at different elementary schools in SA. They were all officially diagnosed with dyslexia. Forty-one students with a mean age of 8.4 years (Grade 2 to 4) were included in this study. All participants had previous experience with electronic devices. Due to the customary separation of males and females in SA educational organizations, researchers did not have access to male students. This also has an advantage by

Table 2: Procedure timetable.

Week #	Session	Activity
1	Pre-starting	Overview of the study, collecting demographic information, conducting reading skill level tests, pre-test
2	Training sessions	Training session 1 and 2
3		Training session 3 and 4
4		Training session 5 and 6
4	Learning gain, SUS	Post-test, SUS

reducing variances between participants.

All participants were native Arabic speakers and the study was conducted in Arabic. The sample characteristics are presented in Figure 6. All of the participants completed the experiment.

4.2 The Effect of Skill Level Content Matching on Perceived Level of Usability

The experimental sample includes 20 students in the experimental group and 21 in the control group. The groups were balanced for prior level of reading performance (p-value for pre-test = $0.711 > 0.05$) and age ($p = 0.808 > 0.05$). Analyses were performed in IBM SPSS.

First, the perceived level of usability was calculated for both matched and standard variants of the RESD system. The usability scores for the matched version of the system based on reading skill level (M = 96.25, SD = 3.39) and the standard version (M = 84.76, SD = 12.09) were acceptable since their average score is greater than 70 (Bangor et al., 2008). Thus, both versions of the system were assessed as usable and valuable in the learning process, and the students with dyslexia were generally satisfied and found them easy to use.

The two versions of the system (matched and standard) were also compared to get a deep insight into their usability and to investigate whether the provision of matching content has any impact on usability. Because data were not normally distributed, an independent sample Mann-Whitney U test was conducted to compare the two conditions. Distributions of the usability scores for both versions were similar, as assessed by visual inspection. The results indicated that there was a statistically significant difference between the overall usability score of the two versions, $U = 358$, $p = 0.0001$. Therefore, the first hypothesis is confirmed, and it can be concluded that the matched version of the e-learning system based on reading skill level of students with dyslexia yields significantly higher levels of perceived usability than a standard version.

When investigating differences among individual items on the SUS tool, we found, for instance, that children using the matched version provided higher ratings to item 8 ("I found the website was very convenient to use") than those using the standard version. Thus, e-learning systems that detect and respond to students' characteristics elicit higher user satisfaction, offer greater academic support, and enhance engagement and satisfaction.

4.3 Perceived Level of Usability and Learning Gain Correlation

A Spearman's rank-order correlation was run to assess the relationship between learning gain and perceived level of usability of the experimental and control groups. The relationship between these two variables is monotonic, as assessed by visual inspection of a scatterplot.

Among participants in the experimental group, there was a moderate positive correlation between perceived usability and learning gain, $r = 0.517$, $p = 0.02 < 0.05$. By contrast, usability and learning gain were unrelated among participants in the control group, $r = -0.364$, $p = 0.105 > 0.05$. Therefore, usability was associated with greater learning gain among students using the matched e-learning mechanism.

5 DISCUSSION

This work aimed to fill a research gap in evaluating the perceived level of usability of an e-learning system that matches the needs of students with dyslexia. This work contrasts with other research on e-learning that does not consider individual differences in learning among students with dyslexia (Al-Ghurair and Alnaqi, 2019; Aljojo et al., 2018; Burac and Cruz, 2020; Vasalou et al., 2017) or relied on teacher reports of usability rather than student perception (Burac and Cruz, 2020; Aljojo et al., 2018; Aldabaybah and Jusoh, 2018). Students have the most influential interactions with the content (Gunesekera et al., 2019),

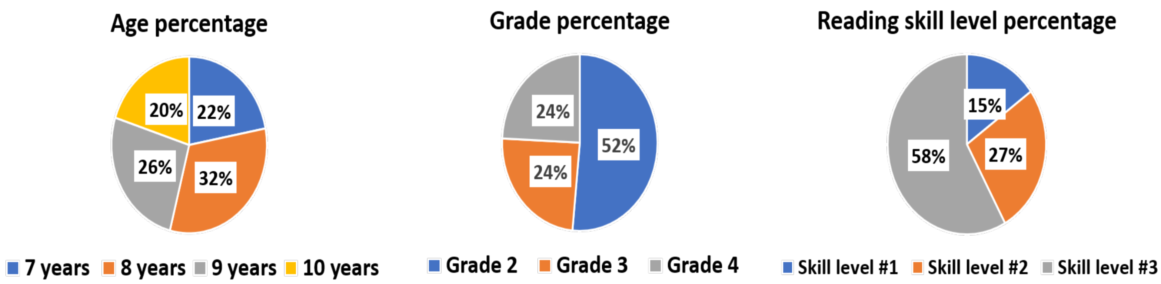

Figure 6: The experimental sample characteristics.

and so evaluating their perception of the usability and their engagement, from their perspective, is an important aspect to ensure the effectiveness and usefulness of these tools (Plata and Alado, 2014).

Although both versions of the system had the same user interface, this work suggests both factors can be enhanced by programs which match students' abilities and needs. The findings of this work are consistent with previous research (Alghabban and Hendley, 2020a; Alshammari et al., 2016) that shows students engage with and enjoy products more when they target their specific skills, needs, and attributes and that those products are perceived as more usable. Such satisfaction precipitates improved learning performance by allowing students to focus on their learning rather than the functionality of the system (Orfanou et al., 2015).

Furthermore, this work's findings yielded a positive correlation between perceived level of usability and learning gain when the e-learning system matches the students' needs. These results parallel those from previous studies (Ardito et al., 2006; Zaharias and Poylymenakou, 2009), which found that student engagement and satisfaction mediated the link between usability and academic progress. Students recognise the content that is useful (Holsapple and Lee-Post, 2006) and suits their needs (Alghabban and Hendley, 2020b; Alghabban and Hendley, 2020a; Sihombing et al., 2020), resulting in higher satisfaction (Alghabban and Hendley, 2020b; Holsapple and Lee-Post, 2006), perceived usability (Alghabban and Hendley, 2020a; Sihombing et al., 2020) and learning performance (Alghabban and Hendley, 2020a; Alshammari et al., 2016).

Because the system's interfaces and layout were otherwise identical, these group differences cannot be attributed to the customization of the user interface. Thus, students with dyslexia are more satisfied with e-learning platforms when their content and learning activities are tailored to their reading skill level. In addition, the results suggest that the skill level of students is one significant characteristic in education (Essalmi et al., 2010) that should be incorporated in e-learning systems in order to enhance engagement and students' experience. This is in line with classroom practice where, once the student's reading level has been determined, teachers select the most appropriate materials for each student (Dolgin, 1975). The findings suggest that students notice when content is tailored to their needs, and that this tailored content increases perceived course quality. This is consistent with previous research (Zaharias, 2009) that shows that focusing on users' needs makes them more active in an e-learning course. Whilst they may not be able to explicitly assess this match, they are, at least subconsciously, aware of it, and this will be reflected in their assessment of aspects of the system that do not change between conditions (Alghabban and Hendley, 2020b).

6 CONCLUSION AND FUTURE WORK

This research evaluated student perceptions of the usability of an e-learning system and how this changes when matching content to the needs of students with dyslexia. We did this for two reasons. First, evaluating the perceived level of usability of e-learning systems is a critical task for researchers (Höök, 2000), and recently, it has become of great importance due to increased use of e-learning technologies spurred by the COVID-19 pandemic (Pal and Vanijja, 2020). That is, a usable e-learning system leads to better engagement and satisfaction, which in turn increases the likelihood that it will help students achieve their learning goals (Alghabban and Hendley, 2020a; Alshammari et al., 2016). Second, we believe that this is an effective metric for assessing the effectiveness of adaptation based on students' needs. By comparing two conditions (adapted and non-adapted) we can measure the change in a student's attitude to the adaptation. This can augment existing metrics (such as learning gain) to give additional insight into whether the adaptation is beneficial.

An experimental evaluation of the e-learning system's usability was conducted with 41 elementary school children with dyslexia, and it yielded significant results. Findings indicated that matching content to children's reading skill level results in a higher level of perceived usability than non-matched content. Thus, when e-learning meets students' needs, they report that the tools are more usable and more engaging and produce greater learning gains.

This research has, however, some limitations. The research targeted dyslexia in the Arabic language. The structure and orthography of Arabic are different from other languages such as English. Therefore, the manifestation of dyslexia in Arabic is different from dyslexia in English (Elbeheri and Everatt, 2007). Thus, further investigation is required to check whether these results can be generalised to other languages, other age groups and male students. Moreover, further investigation may include other student characteristics, such as learning style and personality.

Although this research's findings are promising, further exploration of outcomes associated with the use of e-learning tools is needed. Education through computer-based platforms is complex, involving student self-efficacy, the quality of the learning content, ease of use (Li et al., 2012), the extent to which the tools suit students' different characteristics (Sihombing et al., 2020), the user interface design, interactivity and engagement (Plata and Alado, 2014). The perceived level of usability is one element of the overall user experience that this research evaluated, while further research may focus on other aspect of user experience.

ACKNOWLEDGEMENTS

This work was supported by University of Tabuk, Tabuk, Saudi Arabia. Special thanks to the Ministry of Education in Jeddah, Saudi Arabia for giving the permission to undertake this study in different schools. In addition, many thanks to all participating students and teachers.

REFERENCES

Al-Ghurair, N. and Alnaqi, G. (2019). Adaptive arabic application for enhancing short-term memory of dyslexic children. *Journal of Engineering Research*, 7(1):1–11.

Aldabaybah, B. and Jusoh, S. (2018). Usability features for arabic assistive technology for dyslexia. In *2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC)*, pages 223–228.

Alghabban, W. G. and Hendley, R. (2020a). Adapting e-learning to dyslexia type: An experimental study to evaluate learning gain and perceived usability. In Stephanidis, C., Harris, D., Li, W.-C., Schmorrow, D. D., Fidopiastis, C. M., Zaphiris, P., Ioannou, A., Fang, X., Sottilare, R. A., and Schwarz, J., editors, *HCI International 2020 – Late Breaking Papers: Cognition, Learning and Games*, pages 519–537, Cham. Springer International Publishing.

Alghabban, W. G. and Hendley, R. (2020b). 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, page 41–46, New York, NY, USA. Association for Computing Machinery.

AlGhannam, B. A., Albustan, S. A., Al-Hassan, A. A., and Albustan, L. A. (2018). Towards a standard arabic system usability scale: Psychometric evaluation using communication disorder app. *International Journal of Human–Computer Interaction*, 34(9):799–804.

Aljojo, N., Munshi, A., Almukadi, W., Hossain, A., Omar, N., Aqel, B., Almhuemli, S., Asirri, F., and Alshamasi, A. (2018). Arabic alphabetic puzzle game using eye tracking and chatbot for dyslexia. *International Journal of Interactive Mobile Technologies (iJIM)*, 12(5):58–80.

Allen, M., Mabry, E., Mattrey, M., Bourhis, J., Titsworth, S., and Burrell, N. (2004). Evaluating the effectiveness of distance learning: A comparison using meta-analysis. *Journal of Communication*, 54(3):402–420.

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

AlRowais, F., Wald, M., and Wills, G. (2013). An arabic framework for dyslexia training tools. In *1st International Conference on Technology for Helping People with Special Needs (ICTHP-2013) (19/02/13 - 20/02/13)*, pages 63–68.

Alshammari, M., Anane, R., and Hendley, R. J. (2016). 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, page 2984–2991, New York, NY, USA. Association for Computing Machinery.

Alshammari, M. T. (2019). Design and learning effectiveness evaluation of gamification in e-learning systems. *International Journal of Advanced Computer Science and Applications*, 10(9):204–208.

Appleton, J. J., Christenson, S. L., Kim, D., and Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology*, 44(5):427 – 445.

Ardito, C.and Costabile, M. F., Marsico, M. D., Lanzilotti, R., Levialdi, S., Roselli, T., and Rossano, V. (2006). An approach to usability evaluation of e-learning applications. *Universal Access in the Information Society*, 4(3):270–283.

Baker, R. S. and Rossi, L. M. (2013). Assessing the disengaged behaviors of learners. *Design recommendations for intelligent tutoring systems I*, 1:153.

Bangor, A., Kortum, P. T., and Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction*, 24(6):574–594.

Benmarrakchi, F., Kafi, J. E., and Elhore, A. (2017a). Communication technology for users with specific learning disabilities. *Procedia Computer Science*, 110:258 – 265.

Benmarrakchi, F. E., Kafi, J. E., and Elhore, A. (2017b). User modeling approach for dyslexic students in virtual learning environments. *International Journal of Cloud Applications and Computing (IJCAC)*, 7(2):1–9.

Bonacina, S., Cancer, A., Lanzi, P. L., Lorusso, M. L., and Antonietti, A. (2015). Improving reading skills in students with dyslexia: the efficacy of a sublexical training with rhythmic background. *Frontiers in Psychology*, 6:1–8.

Brusilovsky, P. (2012). Adaptive hypermedia for education and training. In Durlach, P. J. and Lesgold, A. M., editors, *Adaptive technologies for training and education*, volume 46, pages 46–68. Cambridge University Press Cambridge.

Brusilovsky, P. and Millán, E. (2007). *User Models for Adaptive Hypermedia and Adaptive Educational Systems*, pages 3–53. Springer Berlin Heidelberg, Berlin, Heidelberg.

Bukhari, Y. A., AlOud, A. S., Abughanem, T. A., AlMayah, S. A., Al-Shabib, M. S., and Al-Jaber, S. A. (2016). *Alaikhtibarat Altashkhisiat Lidhuyi Sueubat Altaalum Fi Madatay Allughat Alarabia Wa Alriyadiat Fi Almarhalat Alebtidaeiia [Diagnostic Tests for People with Learning Difficulties in the Subjects of Arabic Language and Mathematics at the Primary Stage]*. General Administration for Special Education, The General Administration for Evaluation and Quality of Education, Ministry of Education, Saudi Arabia.

Burac, M. A. P. and Cruz, J. D. (2020). Development and usability evaluation on individualized reading enhancing application for dyslexia (IREAD): A mobile assistive application. *IOP Conference Series: Materials Science and Engineering*, 803:1–7.

Carini, R. M., Kuh, G. D., and Klein, S. P. (2006). Student engagement and student learning: Testing the linkages*. *Research in Higher Education*, 47:1–32.

Chorfi, H. and Jemni, M. (2004). Perso: Towards an adaptive e-learning system. *Journal of Interactive Learning Research*, 15(4):433–447.

Coates, H. (2006). *Student engagement in campus-based and online education: University connections*. Routledge.

Connell, J. P., Halpem-Felsher, B. L., Clifford, E., Crichlow, W., and Usinger, P. (1995). Hanging in there: Behavioral, psychological, and contextual factors affecting whether african american adolescents stay in high school. *Journal of Adolescent Research*, 10(1):41–63.

Diefenbach, S., Kolb, N., and Hassenzahl, M. (2014). The 'hedonic' in human-computer interaction: History, contributions, and future research directions. In *Proceedings of the 2014 Conference on Designing Interactive Systems*, DIS '14, page 305–314, New York, NY, USA. Association for Computing Machinery.

Dolgin, A. B. (1975). How to match reading materials to student reading levels. *The Social Studies*, 66(6):249–252.

Douglas, D. E. and Vyver, G. V. D. (2004). Effectiveness of e-learning course materials for learning database management systems: An experimental investigation. *Journal of Computer Information Systems*, 44(4):41–48.

Elbeheri, G. and Everatt, J. (2007). Literacy ability and phonological processing skills amongst dyslexic and non-dyslexic speakers of arabic. *Reading and Writing*, 20:273–294.

Essalmi, F., Ayed, L. J. B., Jemni, M., Kinshuk, and Graf, S. (2010). A fully personalization strategy of e-learning scenarios. *Computers in Human Behavior*, 26(4):581 – 591.

Fredricks, J. A. and McColskey, W. (2012). *The Measurement of Student Engagement: A Comparative Analysis of Various Methods and Student Self-report Instruments*, pages 763–782. Springer US, Boston, MA.

Ghaban, W. and Hendley, R. (2018). Investigating the interaction between personalities and the benefit of gamification. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32*, pages 1–13.

Gilbert, J., Morton, S., and Rowley, J. (2007). e-learning: The student experience. *British Journal of Educational Technology*, 38(4):560–573.

Gunesekera, A. I., Bao, Y., and Kibelloh, M. (2019). The role of usability on e-learning user interactions and satisfaction: a literature review. *Journal of Systems and Information Technology*, 21:368–394.

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

Hariyanto, D. and Köhler, T. (2020). A web-based adaptive e-learning application for engineering students: An expert-based evaluation. *International Journal of Engineering Pedagogy (iJEP)*, 10(2):60–71.

Harrati, N., Bouchrika, I., Tari, A., and 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.

Hidalgo, F. J. P., Abril, C. A. H., and Parra, M. E. G. (2020). Moocs: Origins, concept and didactic applications: A systematic review of the literature (2012–2019). *Technology, Knowledge and Learning*, 25:853–879.

Höök, K. (2000). Steps to take before intelligent user interfaces become real. *Interacting with Computers*, 12(4):409–426.

Holsapple, C. W. and Lee-Post, A. (2006). Defining, assessing, and promoting e-learning success: An information systems perspective*. *Decision Sciences Journal of Innovative Education*, 4(1):67–85.

Islas, E., Pérez, M., Rodriguez, G., Paredes, I., Ávila, I., and Mendoza, M. (2007). E-learning tools evaluation and roadmap development for an electrical utility. *Journal of Theoretical and Applied Electronic Commerce Research*, 2(1):63–75.

Kangas, M., Siklander, P., Randolph, J., and Ruokamo, H. (2017). Teachers' engagement and students' satisfaction with a playful learning environment. *Teaching and Teacher Education*, 63:274 – 284.

Kori, K., Pedaste, M., Altin, H., Tõnisson, E., and Palts, T. (2016). Factors that influence students' motivation to start and to continue studying information technology in estonia. *IEEE Transactions on Education*, 59(4):255–262.

Lee, J., Song, H.-D., and Hong, A. (2019). Exploring factors, and indicators for measuring students' sustainable engagement in e-learning. *Sustainability*, 11(4):1–12.

Lerner, J. W. (1989). Educational interventions in learning disabilities. *Journal of the American Academy of Child & Adolescent Psychiatry*, 28(3):326 – 331.

Li, Y., Duan, Y., Fu, Z., and Alford, P. (2012). An empirical study on behavioural intention to reuse e-learning systems in rural china. *British Journal of Educational Technology*, 43(6):933–948.

Liaw, S.-S., Huang, H.-M., and Chen, G.-D. (2007). Surveying instructor and learner attitudes toward e-learning. *Computers & Education*, 49(4):1066 – 1080.

Limayem, M. and Cheung, C. M. (2008). Understanding information systems continuance: The case of internet-based learning technologies. *Information & Management*, 45(4):227 – 232.

Lloyd, N. M., Heffernan, N. T., and Ruiz, C. (2007). Predicting student engagement in intelligent tutoring systems using teacher expert knowledge. In *The Educational Data Mining Workshop held at the 13th Conference on Artificial Intelligence in Education*, pages 40–49.

Lo, Y.-y., Cooke, N. L., and Starling, A. L. P. (2011). Using a repeated reading program to improve generalization of oral reading fluency. *Education and Treatment of Children*, 34(1):115–140.

Lyon, G. R., Shaywitz, S. E., and Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1):1–14.

Maravanyika, M., Dlodlo, N., and Jere, N. (2017). An adaptive recommender-system based framework for personalised teaching and learning on e-learning platforms. In *2017 IST-Africa Week Conference (IST-Africa)*, pages 1–9.

Moubayed, A., Injadat, M., Shami, A., and Lutfiyya, H. (2020). Student engagement level in an e-learning environment: Clustering using k-means. *American Journal of Distance Education*, 34(2):137–156.

Nielsen, J. (1993). *Usability Engineering*. Academic Press.

Oga, C. and Haron, F. (2012). Life experiences of individuals living with dyslexia in malaysia: A phenomenological study. *Procedia - Social and Behavioral Sciences*, 46:1129 – 1133. 4th WORLD CONFERENCE ON EDUCATIONAL SCIENCES (WCES-2012) 02-05 February 2012 Barcelona, Spain.

Orfanou, K., Tselios, N., and 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.

Ozkan, S. and Koseler, R. (2009). Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation. *Computers & Education*, 53(4):1285 – 1296. Learning with ICT: New perspectives on help seeking and information searching.

Pal, D. and Vanijja, V. (2020). Perceived usability evaluation of microsoft teams as an online learning platform during covid-19 using system usability scale and technology acceptance model in india. *Children and Youth Services Review*, 119:1–12.

Pang, L. and Jen, C. C. (2018). Inclusive dyslexia-friendly collaborative online learning environment: Malaysia case study. *Education and Information Technologies*, 23:1023–1042.

Peres, S. C., Pham, T., and Phillips, R. (2013). Validation of the system usability scale (sus): Sus in the wild. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1):192–196.

Pickering, J. D. (2017). Measuring learning gain: Comparing anatomy drawing screencasts and paper-based resources. *Anatomical Sciences Education*, 10(4):307–316.

Plata, I. T. and Alado, D. B. (2014). Evaluating the perceived usability of virtual learning environment in teaching ict courses. *Globalilluminators. Org*, 1:63–76.

Putnam, C., Puthenmadom, M., Cuerdo, M. A., Wang, W., and Paul, N. (2020). Adaptation of the system usability scale for user testing with children. CHI EA '20, page 1–7, New York, NY, USA. Association for Computing Machinery.

Read, J. C. (2012). Evaluating artefacts with children: Age and technology effects in the reporting of expected and experienced fun. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, ICMI '12, page 241–248, New York, NY, USA. Association for Computing Machinery.

Rodrigues, H., Almeida, F., Figueiredo, V., and Lopes, S. L. (2019). Tracking e-learning through published papers: A systematic review. *Computers & Education*, 136:87 – 98.

Sahari, S. H. and Johari, A. (2012). Improvising reading classes and classroom environment for children with reading difficulties and dyslexia symptoms. *Procedia - Social and Behavioral Sciences*, 38:100 – 107. ASIA Pacific International Conference on Environment-Behaviour Studies (AicE-Bs), Grand Margherita Hotel, 7-9 December 2010, Kuching, Sarawak, Malaysia.

Sihombing, J. H., Laksitowening, K. A., and Darwiyanto, E. (2020). Personalized e-learning content based on felder-silverman learning style model. In *2020 8th International Conference on Information and Communication Technology (ICoICT)*, pages 1–6.

Tullis, T. S. and Stetson, J. N. (2004). A comparison of questionnaires for assessing website usability. In *Usability professional association conference*, volume 1, pages 1–12. Minneapolis, USA.

Valencia-Arias, A., Chalela-Naffah, S., and Bermúdez-Hernández, J. (2019). A proposed model of e-learning tools acceptance among university students in developing countries. *Education and Information Technologies*, 24(2):1057–1071.

Vasalou, A., Khaled, R., Holmes, W., and Gooch, D. (2017). Digital games-based learning for children with dyslexia: A social constructivist perspective on engagement and learning during group game-play. *Computers & Education*, 114:175 – 192.

WHO (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. World Health Organization.

Zaharias, P. (2009). Usability in the context of e-learning: A framework augmenting 'traditional' usability constructs with instructional design and motivation to learn. *International Journal of Technology and Human Interaction*, 5(4):37–59.

Zaharias, P. and Poylymenakou, A. (2009). Developing a usability evaluation method for e-learning applications: Beyond functional usability. *International Journal of Human–Computer Interaction*, 25(1):75–98.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., and Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology*, 86(3):169 – 193.