The usability evaluation on mobile learning apps with gamification for middle-aged women

Syahida Mohtar1 · Nazeen Jomhari2 · Nor Azyra Omar2 · Mumtaz Begum Peer Mustafa2 · Zulkifli Mohd Yusoff3

Received: 10 March 2022 / Accepted: 10 July 2022 / Published online: 22 July 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Mobile learning and gamification are becoming increasingly popular, especially in education. Numerous researches suggest that mobile learning apps with gamification can improve student learning. However, there are fewer studies that analyze the gamification effects and usability of mobile apps for middle-aged women. The aim of this study is to propose a new mobile learning app with gamification aspects for the use of middle-aged women in memorizing Arabic words. In the preliminary study, pre and post-tests were conducted to determine the effect of the use of the gamification application on 61 middle-aged women using Kahoot! game. The results from the Kahoot! game shows that participants enjoy learning and can memorize Arabic words effectively, indicating that gamification can motivate learning among middle-aged women. Then the usability of a new mobile app prototype containing learning notes and quizzes for user self-learning purposes was evaluated using three (3) criteria namely effectiveness, efficiency, and satisfaction. The considered indicators to measure the application’s effectiveness are (a) tapping behaviour (number of taps used to perform a specific task), (b) number of attempts (how many times users tried to complete the task successfully), and (c) quiz score (score obtained by user after self-learning using the mobile app). Two variables were used to measure the efficiency of the application: the duration of task completion and the number of requests for help. In order to evaluate the satisfaction, a questionnaire was used to determine how respondents rated the mobile app’s usability. The usability evaluation results show that the proposed mobile app was effective as users could utilize all the modules provided and successfully answer the quiz questions.

Keywords Mobile learning apps · Usability · Gamification · Middle-aged

* Syahida Mohtar
msyahida@utem.edu.my

Extended author information available on the last page of the article
1 Introduction

Nowadays, mobile devices are not only used for communication and entertainment, but also in the field of education, including the study of the Quranic Arabic. Traditionally, adult Muslims preferred to visit a mosque or surau to learn more about Islam through religious lectures or attend a Quranic class. Unfortunately, the circumstances of the Covid-19 pandemic outbreak have disrupted teaching and learning practices through physical and face-to-face contact. As a result, many have taken the initiative to watch religious lectures or learn the Quran through video conferencing on the Internet, YouTube, or other similar platforms. However, learning enthusiasm and encouragement starts to decrease over time as they become bored and have less fun in the learning process.

Muslim adults who are busy working usually find it hard to find the time to attend Quranic Arabic classes. In order to rectify this problem, some mobile app developers have created apps with the objective of teaching Muslims Quranic Arabic vocabulary items so that they could begin to learn and understand Quranic Arabic (Mustaffa et al., 2019). For instance, a study by Abdul Ghani et al. (2022) investigated the impact of mobile digital games on Arabic language acquisition among university students. The findings revealed that mobile digital games significantly improved the Arabic vocabulary skills of the students. Although there are existing studies focusing on the impact of mobile games, however, its effectiveness has not yet been fully proven for middle-aged women. Furthermore, according to (Salam et al., 2019), the development of Quranic Arabic vocabulary apps is still in its infancy by looking at the small number of apps available on the Play Store compared to English vocabulary apps. Hence, the positive effects and outcomes of mobile learning games in terms of learning for middle-aged women have yet to be investigated.

The Research questions for this study are: (1) To what extent does a game-based application such as Kahoot perceive learning competency among middle-aged adults? (2) What is the relationship between mobile learning and usability? (3) What is the relationship between usability and student achievement in the learning process?

2 Literature review

2.1 Gamification

Gamification, is an incorporation of game design elements in a non-game to solve task at hand effectively (Khaleel et al., 2016). The popularity of gamification in the field of education has been rising for years now. It is a special approach to increase learners’ engagement by incorporating game-based elements into an educational environment (Smiderle et al., 2020). Numerous past studies have shown that gamification does impact learning outcomes and motivation levels.
By designing and implementing mobile based learning application integrate with gamification concept can improve students’ knowledge, understanding (Tundjung-sari, 2018) and teaching approaches especially in language learning (Md. Yunus & Azmanuddin bin Azman, 2019).

The percentage of smartphone users in Malaysia among middle-aged adults continues to increase from 75.9% in 2018 to 78.0% in 2019 (MCMC, 2019). This shows that middle-aged adults are proficient in using mobile applications, including the ability to browse the Internet using a mobile phone. This enables middle-aged adults to use game-base application such as Kahoot! as it is easily accessible using smartphone (Wichadee & Pattanapichet, 2018).

Kahoot! game was released in 2013. Nowadays, it is among the most popular game-based learning platforms (Wang & Tahir, 2020). It can help educators to turn the learning session into a gameshow without building it from the scratch. It is a game-based learning tool that are designed through videos. Other than it is free, what educators need to do is simply prepare the questions and answers into the site to create an instantly playable game. The result can be retrieved right away for further evaluation if needed. Besides that, for learners, they can get immediate individual feedback in terms of points from the game regarding their level of knowledge or understanding.

Middle-aged women need a special tool that can influence them to stay focus, enjoy and motivated due to cognitive impairment. The findings from a research carried by Md. Yunus and Azmanuddin bin Azman (2019), shows that visual cues in Kahoot! game can give a positive influence to the learners especially those who needs special education. Learners’ attention will be created especially through true and false questions. The embedded visual and audio elements in Kahoot! game will be very helpful and its good features help in promoting engagement, motivation, and learning among middle-aged (Kalleny, 2020). Besides that, Kahoot! game will give a positive impact towards learners’ engagement (Ciaramella, 2017), and can drive their motivation as well as to increase their desire to win (Zarzycka-Piskorz, 2016).

Most importantly, Kahoot! game can be freely and easily use in the current situation of COVID-19 pandemic. Learners can use, learn, and play with it whenever and wherever they want without any limitation. Therefore, older adults including middle-aged can benefit from Kahoot! which it can be used as an alternative to formative assessment in learning new language (Zakaria & Hashim, 2020), such as learning Quranic Arabic for the purpose of understanding the Quran.

2.2 Measuring the effectiveness of mobile game-based apps

Pretest–posttest designs are commonly used in behavioral research, particularly to compare groups and/or to measure changes resulting from experimental treatments. For example, there have been studies on the problem of misspelling Thai words on social media, which is growing rapidly among youth. First-person 3D mobile games have been developed as a solution. Pretests and posttests were used to evaluate the performance of 37 university students and collected data were statistically analyzed using a paired-samples t-test. The results of the study show that the game-based learning
approach significantly improves players’ learning performance of misspelled written Thai words with average, posttest scores were 15.027 points higher than pretest scores (95% CI [17.615, 12.439]) (Saksrisathaporn, 2020).

Another study in (Md. Yunus & Azmanuddin bin Azman, 2019), investigated the effect of Kahoot! on students’ memorization of irregular verbs. This study was conducted to prove that the use of Kahoot! helps the Year 3 students to retain English irregular verbs. A quasi-experimental study was conducted to collect data from 35 students using the pretest and posttest scores. The results, which were analyzed descriptively using sample t-test, showed that students’ performance on the posttest had improved over the use of Kahoot! in the classroom. A similar study done by Hashim et al. (2019) added two more game-based learning methods, namely Socrative and the PowerPoint Challenge Game. Thirty (30) secondary school students participated in the study. The data were analyzed using percentages to compare the results after the three game-based learning interventions. The main results showed that students’ scores in the grammar post-test increased significantly after the intervention compared to the pre-test. Thus, the results show that game-based learning is effective in teaching grammar to English Second Language (ESL) learners.

In addition to the studies in which Kahoot was used, there are other studies in which educational games in the form of Snakes and Ladder were examined as learning media. The use of educational games as learning media should also encourage users to discover and build their own knowledge about learning. Using pretest and posttest methods, data analysis was conducted to determine if there were changes in student learning outcomes. As expected, the results show that student learning outcomes are improved through the use of educational games (Patmanthara et al., 2019).

The study by (Wardaszko & Podgórski, 2017) aimed to evaluate the effectiveness of learning with Mobile Learning Game (MLG) compared to learning with textbook. The design of the study intended to show both short- and long-term learning outcomes in knowledge acquisition. The comparative study was conducted with a group of 160 freshman students majoring in management and finance. The participants were divided into subgroups and were taught using either textbook-based learning or a mobile learning game. Three tests were administered to measure their performance: a pre-test at the beginning of the experiment and two post-tests, the first of which was administered immediately after the learning process and the second were administered 2–3 weeks after the initial learning. The study shows that overall MLG was a very effective learning tool for all students in the experimental group and provided knowledge gains that were measured immediately after the experiment. However, the study was unable to identify long-term effects and many dependencies appear to be quite weak in terms of statistical significance. Table 1 summarizes the studies that employed pretest and posttest approaches in mobile learning and gamification research.

### 2.3 Mobile learning application

The use of mobile applications by middle-aged people is steadily increasing, not only to access their social networks, but also to support, enrich and extend their self-directed learning (Horbatiuk et al., 2019; Morrison & Koole, 2018).
| Author & Year | Research Title                                                                 | Research Area                  | Technology                        | Method                                      |
|---------------|--------------------------------------------------------------------------------|--------------------------------|-----------------------------------|---------------------------------------------|
| (Saksrisathaporn, 2020) | A game-based learning approach to improve students’ spelling in thai           | Spelling thai words            | Mobile game-based learning         | Comparison between pre-test and post-test performance |
| (Md. Yunus & Azmanuddin bin Azman, 2019) | Memory stay Or stray?: Irregular verbs learning using Kahoot!                   | Irregular english verb          | Mobile game-based learning         | Comparison between pre-test and post-test performance |
| (Hashim et al., 2019) | Improving ESL learners’ grammar with gamified-learning                         | English grammar                 | Mobile game-based learning         | Comparison between pre-test and post-test performance |
| (Patmanthara et al., 2019) | The use of ladder snake games to improve learning outcomes in computer networking | Computer networking             | Mobile game-based learning         | Comparison between pre-test and post-test performance |
| (Wardaszko & Podgórska, 2017) | Mobile learning game effectiveness in cognitive learning by adults: A comparative study | Cognitive learning effects     | Mobile game-based learning         | Comparison between pre-test and post-test performance |
Many previous studies have shown that mobile applications can be used as tools to facilitate memorization such as in learning foreign languages. The results of several previous studies have shown that learning a foreign language is beneficial for older people as it improves their cognitive performance, has a positive impact on their spirit, reduces anxiety, boosts their self-confidence, builds new social bonds and provides a sense of self-realization (Klimova, 2020). Since learning a foreign language requires constant effort and motivation, gamification in apps is one of the ways to transform the user experience into a sense of fun while learning. Gamification can make the learning process more innovative, fun and productive. Most importantly, it can help retain learners’ knowledge, which is very important for memorization. Table 2 shows the list of studies that involved middle-aged. In the listed studies, the age range of the middle-aged adult is between 40 to 60 years old. However, most researchers tend to focus on the benefits, readiness, and adoption of the mobile learning app rather than the effectiveness of its use with gamification elements.

2.4 Usability criteria for mobile application

The design and development of the mobile learning application for the elderly population, including middle-aged people, must meet their needs. Older people usually face difficulties when the mobile application has a complicated interface and functionality (Anam & Abid, 2020). Due to cognitive and physical decline with increasing in age, it becomes crucial for the ageing people to use mobile application with high usability. The most important factors to consider when designing mobile applications are the organization of information on the screen, the display of data and color coding, hierarchical menus and screen navigation, and the design of icons, labels, and symbols (Mehrban & Asif, 2010). Therefore, usability needs to be analyzed by the developer when the app is in prototype phase. A usability evaluation is the best way to get a product in the hands of actual targeted users to see if and how they use it prior to the product’s release.

Different literature studies have different definitions and characteristics of usability (Nayebi et al., 2012). The importance of measuring usability is to investigate the origins of an object’s perceived efficiency or style whether it is able to achieve the highest capacity to provide the best condition for its users. Usability measurement is one of the important steps in mobile application evaluation to address its advantages and disadvantages (Nayebi et al., 2013). In mobile application development, usability studies the simplicity, satisfaction, and elegance of interaction with an application.

Usability criteria for mobile application has evolved over time and has got its relevance in many aspects (Madan & Dubey, 2012). Usability aspects for mobile application built for older adults should focus on its effectiveness, efficiency and users’ satisfaction in terms of its ease of use, ease of learning, willingness to use it in future, and willingness to recommend it to others (Islam et al., 2020). Besides that, effectiveness, efficiency, and satisfaction are also important aspects in usability measurement (Yeni et al., 2020). A research carried out in (Nahar et al., 2019),
| Author & Year       | Research Title                                                                 | Research Area                              | Middle-age range                          |
|--------------------|--------------------------------------------------------------------------------|--------------------------------------------|--------------------------------------------|
| (Klimova, 2020)    | Benefits of the Use of Mobile Applications for Learning a Foreign Language by Elderly Population | Education-Language Learning                | 50—85 years old                           |
| (Cao et al., 2020) | Exploring seniors’ continuance intention to use mobile social network sites in China: a cognitive-affective-conative model | Healthcare-Well Being                      | 50–70 years old                           |
| (Lin & Ho, 2020)   | The development of a mobile user interface ability evaluation system for the elderly | Usability                                  | 50 -59, 60—69, 70—79 and > 80 years old  |
| (Islam et al., 2020)| Chakuri-Bazaar: A Mobile Application for Illiterate and Semi-Literate People for Searching Employment | User-Requirement                           | 18 —55 years old                          |
| (Al-Sabbagh et al., 2019) | Mobile language learning applications for Arabic speaking migrants—a usability perspective | Education-Language Learning                | 15—50 years old                           |
| (Huang et al., 2019) | Middle-aged adults’ attitudes toward health app usage: a comparison with the cognitive-affective-conative model | Healthcare-Well Being                      | 45—65 years old                           |
stated that the most important usability parameters in evaluating a mobile application should include learnability, memorability, efficiency, error, and satisfaction.

Usability has been 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” (ISO, 1998). Therefore, in this study data was analyzed according to usability parameters, which include effectiveness, efficiency, and satisfaction (Marques et al., 2020).

3 Participants

In this study, 61 female participants were recruited to participate in the study. The participants ranged in age from 40 to 60 years old and had different professions and backgrounds. In addition, none of the female participants had experience using Kahoot! to learn Quranic Arabic as well as using similar apps such as Quran Pedia.

4 Methodology

The aim of this study in the initial phase is to investigate the effectiveness of gamification towards middle-aged women using Kahoot! application. To determine the effect of gamification, this study used a quasi-experimental design in which quantitative data were obtained solely from comparing pre-tests and post-tests with the learning intervention. Furthermore, this design was chosen because participants were recruited for this study using the snowball method, without a controlled group.

In the pre-test, participants were asked to answer questions from the Kahoot! application without learning intervention by instructor. The results of the Kahoot! game before and after the learning intervention will indicate and present the performance of the middle-aged women. This is to show the impact of gamification with the Kahoot! game on their Quranic Arabic knowledge.

The instrument used for this research was the Arabic word tests with 35 questions given before and after learning through the Kahoot! game. The Arabic word tests consisted of two quizzes, a red pattern quiz and a green pattern quiz. The red pattern quiz focuses on Arabic words that occur frequently in the Quran (Fig. 1a). The green pattern quiz focuses on Arabic words that have similar pronunciation to Malay words and basic conjugation of the past tense in Arabic verbs (Fig. 1b). After obtaining results on the effects of gamification on the learning of middle-aged women, an evaluation of the usability of a proposed mobile learning app called Quran Pedia was conducted.

In assessing usability of the Quran Pedia, a descriptive study was conducted in which a person’s behavior is observed and described without influencing it in any way. The data is collected quantitatively rather than qualitatively because the goal of this study is to assess the effectiveness of Quran Pedia applications that assist middle-aged women in learning Arabic. Users’ interaction with the application was recorded using mobile screen recordings. Each participant has to install the provided apps in Android mobile devices. They are then given a 5–10-min tutorial on how to
take screen recordings of registration, self-study, and self-assessment. Participants are encouraged to use the app in their free time, and they must record their screen activities while using the apps for data collection. The recorded app activities must then be emailed by the participant to the researcher. Data were collected on tapping behavior, number of attempts, time taken to complete the task, and number of requests for help. At the end of the assessment activity, participants were asked to complete a series of post-test questionnaires.

The post-test questionnaires is one of the important tools in usability research to determine the level of user satisfaction. The purpose of this questionnaires is to determine how the respondents rate the usability objectives. The questionnaires contain 5 questions that capture the level of overall satisfaction, complexity of use (ease of use), learnability (ease of learning), intention to use the application in the future, and willingness to recommend it to others. These questions were rated using a 5-point scale, with response scales ranging from 1 (strongly disagree) to 5 (strongly agree). Cronbach’s alpha tests were used in this study to determine if the multiple-question Likert-scale surveys were reliable (Cronbach, 1951). According to Pallant (2020), to pass the reliability test, the Cronbach’s Alpha values must be greater than 0.7. However, guidelines provided by DeVellis (2017) stated that reliabilities measurement ranges from minimally acceptable (0.6 or greater) to excellent (0.9 or greater). Therefore, it is confirmed that this set of data measures the usability of the mobile app in the acceptable range see Table 3.

5 Data analysis

In determining the effect of gamification, the process began by collecting all the participants’ scores from pre-test and post-test quiz through Kahoot! game. Besides that, participants’ task completion time were collected to investigate is there any relationship between score and task completion time. Once the collecting process was completed, these scores were then converted into percentages. The second step was by transferring all of the scores and completion times for both tests into a table.
form using Microsoft Excel. Then, table was uploaded into R programming as data frame to run the result. Paired sample t-test was run to examine the difference in the mean before and after the learning intervention and Pearson correlation coefficient was used to investigate is there any relationship between test score and test completion time. The result from the data was very significant in proving the effectiveness of Kahoot! game in helping middle-aged women to learn Arabic.

In assessing usability of the Quran Pedia prototype, the data were analyzed descriptively using descriptive statistics where the central tendency of criteria and statistical dispersion were applied. Mean (average value), minimum (lowest value) and maximum (highest value) were calculated to get information about the middle or center of a distribution. In addition, standard deviation was also calculated to quantify the amount of variation or dispersion of dataset values.

### 6 Results & discussion

#### 6.1 The effect of gamification

Referring to Table 4, participants’ scores increased on the post-test. On the red pattern quiz, the grade of A improved by 37 participants compared to the pre-test, which is more than 50% of the total participants. None of the participants received grades D and E on the post-test. The figure shows that the 14 participants who received grades D and E on the pre-test received a better grade on the post-test.

The quiz with the green pattern showed an increase in the number of participants who received grades A and B which is 26 and 3 participants, respectively. Majority obtained grade C and grade E in the pre-test, whereby in the post-test, only 15 and 1 participants respectively. The decrease in this number shows that many participants received a better grade in the post-test than in the pre-test. Only 3 participants received grade D and E on the post-test. This could be perceived as no improvement; however, it is likely that this is due to the length of the learning unit. The study session was too short between the pre-test and the post-test. These participants may need more time to learn a foreign language.

As shown in Table 5, the paired sample statistics show the overall results of the test. In this case, the particular interest was the learning outcome in Arabic to determine whether the pre-test or the post-test had higher average test scores and the

| Item     | Item Description                | Corrected Item-Total Correlation | Cronbach’s Alpha if Item Deleted | Cronbach’s Alpha       |
|----------|---------------------------------|----------------------------------|----------------------------------|------------------------|
| U1       | Overall satisfaction            | 0.221                            | 0.844                            | 0.775                  |
| U2       | easy-to-use                     | 0.500                            | 0.749                            |                        |
| U3       | Easy to learn                   | 0.586                            | 0.720                            |                        |
| U4       | Willing to use in the future    | 0.741                            | 0.663                            |                        |
| U5       | Will recommend to others        | 0.770                            | 0.655                            |                        |
A paired sample t-test was used to compare the pre-test and post-test result on the Arabic red pattern and green pattern quizzes. This is to investigate if there was a significant difference in the average of the two tests in terms of score and completion time.

For learning improvement, the hypothesis tested for this area was there is no statistically significant difference in test scores between pre-test and post-test in learning Arabic using Kahoot! game. The result from Table 6 and Fig. 2 showed greater difference in terms of means for the post-test scores for both quizzes. This statistic proved that the middle-aged women’s performance was improving over the use of Kahoot! game. Statistically, the value of sig (p) paired sample t-test is 0.000 which is less than 0.05. Therefore, it can be concluded that Ho is rejected, which indicated that there was significant difference in test scores between pre-test and post-test in learning Arabic using Kahoot! game.

For the ability to complete the task faster, the hypothesis tested for this domain was that there would be no statistically significant difference in task completion time between pre-test and post-test in learning with the Kahoot! game. The result from Table 7 and Fig. 2 showed small difference in terms of means for the post-test completion times for both quizzes. This statistic proved that the time taken for

| Grades | Range of scores (%) | No. of Participants |
|--------|---------------------|---------------------|
| A      | 80—100              | 16 53               |
| B      | 65—79               | 20 5                |
| C      | 50—64               | 11 3                |
| D      | 40—49               | 8 0                |
| E      | 0—39                | 6 0                |
| Total  |                     | 61 61              |

Table 4: Pre-test and Post-test results

| Grades | Range of scores (%) | No. of Participants |
|--------|---------------------|---------------------|
| A      | 80—100              | 16 53               |
| B      | 65—79               | 20 5                |
| C      | 50—64               | 11 3                |
| D      | 40—49               | 8 0                |
| E      | 0—39                | 6 0                |
| Total  |                     | 61 61              |

Table 5: Pre-test and Post-test results (Mean & Std. Deviation)
middle-aged women to answer all questions in both quizzes was improving over the use of Kahoot! game. Statistically, the value of sig (p) paired sample t-test is less than 0.05. Therefore, it can be concluded that Ho is rejected, which indicated that there was significant difference in completion time between pre-test and post-test in learning Arabic using Kahoot! game. In other words, the middle-aged women improved their knowledge of Arabic words and showed small improvement in terms of the time they took to complete the quiz. This is due to no competitive element among the participants was required in this study. Thus, no sense of urgency is elevated. They were allowed to take the quiz at their own pace and had the flexibility to allocate their time depending on the difficulty of the questions see Fig. 3.

**Table 6** Comparison of score and task completion time in pre-test and post-test

|                  | Mean  | 95% Confidence Interval of the Difference | t   | df  | Sig. (2-tailed) |
|------------------|-------|------------------------------------------|-----|-----|-----------------|
|                  | Lower | Upper |                                  |     |     |                 |
| Score            |       |       |                                  |     |     |                 |
| Red Pattern pre  | 0.241 | 0.196 | 0.287                            | 10.568 | 60 | 0.000           |
| Green Pattern pre| 0.284 | 0.235 | 0.332                            | 11.665 | 60 | 0.000           |
| Post             |       |       |                                  |     |     |                 |
| Red Pattern post | -17.520 | -26.368 | -8.672                          | -3.961 | 60 | 0.000           |
| Green Pattern post| -17.303 | -26.626 | -7.981                          | -1.684 | 60 | 0.000           |

**Fig. 2** Box plot showing participants mean scores for pre-test and post-test

**Table 7** Correlation coefficient of the learning improvement (score) and task completion time

|                  | N  | Pearson correlation between score & task completion time | Sig  |
|------------------|----|---------------------------------------------------------|------|
|                  |    | Score & task completion time                            |      |
| Red Pattern      |    |                                                         |      |
| pre              | 61 | -0.25                                                   | 0.05 |
| post             | 61 | -0.13                                                   | 0.33 |
| Green Pattern    |    |                                                         |      |
| pre              | 61 | -0.07                                                   | 0.58 |
| post             | 61 | -0.32                                                   | 0.01 |
Figure 4 shows the Pearson correlation was calculated to check the relationship between score and task completion time. As shown in Table 7, statistically, no significant correlation was noticed between score and task completion time for both quizzes. The results showed that the three weeks of learning experience did not lead to higher valence to complete the task faster, which provides no positive relationship between scores obtained by participants and time taken for them to complete the task. However, learning experience and Kahoot! game is effective and relevance to the middle-aged women in helping them to learn Arabic. In addition, older students tend to take more time to complete a test compared to younger students, but age was not significantly related to their test performance (Landrum et al., 2009). Based on observations of the participants’ behavior throughout the process, Kahoot! game has a capability to maximize their engagement (Ciaramella, 2017), and their motivation (Zarzycka-Piskorz, 2016). On top of that, this study also proved that learning grammar using traditional method will be terrified to middle-aged women (Md. Yunus & Azmanuddin bin Azman, 2019).
6.2 Usability evaluation

Quran Pedia usability was evaluated using parameters defined by ISO (1998), which includes effectiveness, efficiency, and satisfaction.

6.2.1 Effectiveness

Effectiveness is the accuracy and completeness with which users try to achieve specified goals. Mobile application effectiveness can be measured based on how easy the application can be used by observing user behaviours (Nayabi et al., 2012). The results of effectiveness evaluation are shown in Table 8. The considered indicators to measure the application’s effectiveness are (a) tapping behaviour (number of taps used to perform a specific task), (b) number of attempts (how many times users tried to complete the task successfully), and (c) quiz score (score obtained by user after self-learning using mobile app).

**Number of taps** The optimal (minimum) number of taps required to perform the registration, login, payment and reading tasks 1, 2, 3 and 4 were 1. The results showed that participants took more taps than the optimal number of taps, on average, which is an indication of effective use of the application by the participants. To perform registration, make payment and login, results showed that some participants faced some difficulties. For learning task 1 and learning task 2, the participants tapped on average 22.1 and 10.9 times to learn the Arabic words, respectively. It deemed reasonable that some participants would require higher number of taps to learn because of lack of knowledge and wanted to be more confident before entering the quiz session. To login to the quiz after self-learning, it showed that some participants faced some difficulties, maximum 35 taps. For quiz 1 and quiz 2, on average, all participants managed to answer all questions. Based on minimum taps, it showed that some participants faced some technical issues and lack of guidance while answering the quiz, hence, did not managed to complete all provided questions.

**Number of attempts** On average, the participants only took 1.0 to 1.9 attempts to complete each task, which in turn indicated that participants completed their tasks effectively using the application.

**Quiz score** Based on the result, on average, participants only managed to score half of full marks from both quizzes. This is because some participants accidentally logged out from the apps because of the technical issue. Other than that, some participants only learnt one time about green and red pattern word during learning task 1 and task 2. Acquired minimum knowledge about Arabic words led them to score low marks.

6.2.2 Efficiency

Efficiency can be assessed by examining the amount of effort a user spends to complete a task. Efficiency use to indicate high productivity level of users’ performance
| Evaluation Criteria | Data Type            | Task                        | Mean and SD | Min | Max |
|---------------------|----------------------|-----------------------------|-------------|-----|-----|
| Effectiveness       | Tapping Behavior     | Apps Registration           | 11.3 ± 14.7 | 1   | 55  |
|                     |                      | Apps Login                 | 7.0 ± 8.1   | 1   | 35  |
|                     |                      | Apps Payment               | 11.9 ± 10.8 | 1   | 56  |
|                     |                      | Reading Task 1: Apps Manual| 2.9 ± 1.7   | 1   | 9   |
|                     |                      | Reading Task 2: How to Understand Qur’an | 3.0 ± 1.9 | 1   | 10  |
|                     |                      | Reading Task 3: Why Easy to Understand Qur’an | 2.7 ± 1.2 | 1   | 7   |
|                     |                      | Reading Task 4: Arabic Word Pattern | 3.0 ± 1.5 | 1   | 10  |
|                     |                      | Learning Task 1: Red Pattern Word | 22.1 ± 20.5 | 1 | 119 |
|                     |                      | Learning Task 2: Green Pattern Word | 10.9 ± 9.2 | 1 | 45  |
|                     |                      | Quiz Login                 | 7.8 ± 7.6   | 1   | 35  |
|                     |                      | Quiz 1                     | 19.7 ± 1.8  | 8   | 20  |
|                     |                      | Quiz 2                     | 20.0 ± 0    | 20  | 20  |
| Number of Attempts  |                      | Apps Registration           | 1.3 ± 0.5   | 1   | 3   |
|                     |                      | Apps Login                 | 1.2 ± 0.4   | 1   | 2   |
|                     |                      | Apps Payment               | 1.1 ± 0.2   | 1   | 2   |
|                     |                      | Reading Task 1: Apps Manual| 1.1 ± 0.4   | 1   | 3   |
|                     |                      | Reading Task 2: How to Understand Qur’an | 1.2 ± 0.4 | 1 | 3   |
|                     |                      | Reading Task 3: Why Easy to Understand Qur’an | 1.1 ± 0.5 | 1 | 5   |
|                     |                      | Reading Task 4: Arabic Word Pattern | 1.0 ± 0.2 | 1 | 2   |
|                     |                      | Learning Task 1: Red Pattern Word | 1.9 ± 1.5 | 1 | 11  |
|                     |                      | Learning Task 2: Green Pattern Word | 1.2 ± 0.5 | 1 | 4   |
|                     |                      | Quiz Login                 | 1.1 ± 0.4   | 1   | 3   |
|                     |                      | Quiz 1                     | 1.1 ± 0.5   | 1   | 4   |
|                     |                      | Quiz 2                     | 1.0 ± 0.2   | 1   | 2   |

*Quiz Score (out of 20 points)*

| Quiz Score (out of 20 points) | Quiz 1 | Quiz 2 |
|-------------------------------|--------|--------|
| Quiz 1                        | 10.1 ± 4.6 | 0   | 19   |
| Quiz 2                        | 9.0 ± 3.4  | 0   | 16   |
within a specific time after they have learned the system. Efficiency also refers to the speed of the users’ performance to tasks [29]. Therefore, this study focused on how fast users can complete particular tasks and the number of helps needed to complete the tasks. The results of efficiency evaluation are shown in Table 9. Two variables were used to measure the application’s efficiency: task completion time and number of times help asked for (number of times help asked from the researcher).

**Task completion time** Other than quiz 1 and quiz 2, the results showed that participants took an average below 156 s (within 2.5 min) to complete each task. These results indicate that participants managed to complete the task very fast and efficiently using the application. Quiz 1 and quiz 2 consisted of 20 questions. Participants were required to answer each question within 30 s, which in total maximum 600 s for each quiz. Based on the result, on average, participants managed to answer all questions within 135 s and 231 s for quiz 1 and quiz 2, respectively. However, to register, make payment and login, results showed that some participants took a very long time to complete these tasks.

**Number of times asked for help** Number of times asked for help by participants were recorded. This is very important because if the participants asked for help multiple times, it gives an indication of poor user interface (UI) design. Thus, asked for help from the researcher once is considered optimal. The results show that on average, the participants asked for help below than the optimal value to complete all tasks. These findings indicated that the UI was an efficient to help them to use and perform task using the application.

### 6.2.3 Satisfaction

Satisfaction is the users’ comfort with and positive attitudes towards the application. This is to examine whether this mobile application meets user expectation. In this study, five questions were used to measure satisfaction. The results in Table 10 shows that the average score of each satisfaction measure was between 3.5 to 3.9. The participants found this app was easy to learn (mean: 3.8) and easy to use (mean: 3.9). Their willingness to use this app in future (3.8) and recommend it to others (mean 3.9) was also encouraging. Finally, the overall satisfaction score was 3.5, which indicated that participants were satisfied with this application. However, there was a lot of feedback given by the participants who said that this application needs further improvements to make it better for future use, especially those who have given a minimum rating.

### 7 Limitations and future work

This study focuses only on middle-aged women who are between 40 and 60 years old. The results need to be verified with other age groups and genders, as the advantages and disadvantages of using Kahoot! as a game-based learning tool may differ.
The structural skills or the level of motivation and engagement may affect males and females differently. Therefore, it is recommended that further studies should be devoted to male N-NAS learners and female learners of different ages to investigate their learning ability through the Kahoot! game.

The second limitation is related to the sample size. To ensure the reliability and validity of the study, a larger scale could be chosen. This could be achieved by including a larger number of participants instead of including only a group of 61 participants. Finally, there is the issue of learning time. As mentioned in the result, in this study there was no positive relationship between the score and the task completion time. By increasing learning time, probably can encourage learners to become more confident and more adept at using Kahoot! game. More questions should be included in future work to evaluate students’ competence in recalling the translations of Arabic terms to Malay language that have been learned. Longitudinal studies towards adult students in the use of Arabic language learning are also encouraged to be conducted in the future by using more complex question forms.

In terms of usability, this study has some limitations as well that also provide the implications for future work. Firstly, limitations regarding the limited number of the participants. The sample size of the evaluation study was small. Large-scale studies with more participants need to be done in order to understand strengths and weaknesses of Quran Pedia in helping N-NAS to learn Arabic. Moreover, more study needs to be undertaken in order to confirm the strength of the association between Quran Pedia and its effectiveness in teaching Arabic words. Secondly, to evaluate this application, we did not collect the data regarding number of errors. Errors occur when a user fails to achieve the desired goal while using it. As stated, usability evaluation is crucial for application further improvements. Therefore, number of errors, severity of these errors, and easiness to recover from the errors must be considered in the future work (Nahar et al., 2019). Third, in this study, we did not collect the qualitative data for performance and assessment rating. Future study can be enhanced by providing free-text space in questionnaire. Participants can give their feedback or any idea regarding the strength and weakness of this application. Forth, this study only focussed on user-based evaluations. According to Nayebi et al. (2013), to have more accurate result, expert-based evaluations could be combined with user-based evaluations. Lastly, the results of usability metrics were not compared with the evaluation results of any other different kind of mobile language learning applications for middle-aged. Thus, future research could be carried out to compare this mobile application with other applications that are available in the market to provide a more effective validation.

8 Conclusion

This study was conducted to examine the performance of N-NAS in learning Quranic Arabic using the Kahoot! game as well as to evaluate the usability of a prototype mobile app called Quran Pedia.

The result of paired samples t-test between pre-test and post-test scores shows that the Kahoot! game has a statistically significant effect on participants’ performance. It is
| Evaluation Criteria | Data Type | Task Description | Mean and SD | Min | Max |
|---------------------|-----------|------------------|-------------|-----|-----|
| Efficiency          | Task Completion Time (seconds) | Apps Registration | 156.0 ± 459.0 | 2.3 | 3120 |
|                     |           | Apps Login       | 96.8 ± 305.0 | 0.8 | 1680 |
|                     |           | Apps Payment     | 148.0 ± 250.0 | 1   | 1260 |
|                     |           | Reading Task 1: Apps Manual | 49.3 ± 10.5 | 1 | 840 |
|                     |           | Reading Task 2: How to Understand Qur’an | 30.0 ± 77.4 | 0.6 | 600 |
|                     |           | Reading Task 3: Why Easy to Understand Qur’an | 16.9 ± 18.5 | 0.6 | 120 |
|                     |           | Reading Task 4: Arabic Word Pattern | 68.1 ± 63.6 | 1 | 380 |
|                     |           | Learning Task 1: Red Pattern Word | 80.0 ± 66.3 | 1 | 270 |
|                     |           | Learning Task 2: Green Pattern Word | 42.7 ± 46.3 | 1 | 340 |
|                     |           | Quiz Login       | 40.6 ± 70.4 | 0.6 | 426 |
|                     |           | Quiz 1           | 135.0 ± 72.1 | 60 | 420 |
|                     |           | Quiz 2           | 231.0 ± 87.2 | 60 | 420 |
| Number of Times Asking for Help | Apps Registration | 0.9 ± 1.0 | 0 | 4 |
|                     |           | Apps Login       | 0.4 ± 0.7 | 0 | 2 |
|                     |           | Apps Payment     | 1.0 ± 1.1 | 0 | 5 |
|                     |           | Reading Task 1: Apps Manual | 0.2 ± 0.4 | 0 | 1 |
|                     |           | Reading Task 2: How to Understand Qur’an | 0.2 ± 0.4 | 0 | 1 |
|                     |           | Reading Task 3: Why Easy to Understand Qur’an | 0.2 ± 0.4 | 0 | 1 |
|                     |           | Reading Task 4: Arabic Word Pattern | 0.1 ± 0.5 | 0 | 2 |
|                     |           | Learning Task 1: Red Pattern Word | 0.1 ± 0.3 | 0 | 1 |
|                     |           | Learning Task 2: Green Pattern Word | 0.1 ± 0.3 | 0 | 1 |
|                     |           | Quiz Login       | 0.2 ± 0.5 | 0 | 2 |
|                     |           | Quiz 1           | 0.5 ± 0.6 | 0 | 2 |
|                     |           | Quiz 2           | 0.3 ± 0.6 | 0 | 3 |
proved that Kahoot! game is able to train, evaluate and improve learners’ structural ability in learning Arabic especially among middle-aged women (Tundjungsari, 2018). The results of this study suggest that gamification can be an effective approach in the teaching and learning process (Md. Yunus & Azmanuddin bin Azman, 2019). The results also show that adult learners perceive Kahoot! as a fun and effective tool as they focus on their learning through the immediate feedback they received (Kalleny, 2020). This can boost student engagement and motivation through its special and entertaining features (Zarzycka-Piskorz, 2016). Most importantly, it has the ability to retain learners’ knowledge, which is very important in language learning (Ciaramella, 2017). So, we found that gamification is effective and relevant for middle-aged women and helps them to memorize Arabic words and translations. However, there are some suggestions mentioned in the limitations that could be considered in future research related to the implementation of gamification in education, especially in language learning. Regardless of the shortcomings, this study has demonstrated the effectiveness of Kahoot! game in language learning.

The usability evaluation measured the most important aspects of usability goals, including effectiveness, efficiency, and satisfaction. The results of the evaluation study show that the application was effective and efficient, and users were satisfied in terms of its ease of use, learning effect, willingness to use it in the future, and willingness to recommend it to others. However, this Quran Pedia prototype app requires some improvements to fix its technical issues and to increase its usability in order to provide a better user experience for middle-aged people such as:

a. Upgrading the user interface and improving the process of registration, payment and login modules.

b. Embedding audio elements in self-learning modules to aid memorization with melody (Ciaramella, 2017).

c. Improving the self-evaluation thru quiz and self-monitoring modules by displaying number of questions, answer time for each question, score, and task completion time for purpose of monitoring and make it more effective (Kalleny, 2020).

| Evaluation Criteria | Data Type        | Mean | Std. Deviation | Min | Max |
|---------------------|------------------|------|----------------|-----|-----|
| Satisfaction        | Overall Satisfaction | 3.5  | 1.0            | 2   | 5   |
|                     | Easy to Use      | 3.9  | 0.9            | 1   | 5   |
|                     | Easy to Learn    | 3.8  | 0.9            | 2   | 5   |
|                     | Future Use       | 3.8  | 1.9            | 1   | 5   |
|                     | Recommended to Others | 3.9  | 0.9            | 2   | 5   |

Acknowledgements We would like to thank you to Universiti Teknikal Malaysia Melaka (UTeM) for sponsoring the corresponding author in this research numbered UTeM.02.13.04/500-4/12/16/1/2(91)

Funding The corresponding author is sponsoring by Universiti Teknikal Malaysia Melaka (UTeM) numbered UTeM.02.13.04/500-4/12/16/1/2(91) to pursue her PhD degree in these studies.
Data availability The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that there is no conflict of interest regarding the publication of this paper.

References

Abdul Ghani, M. T., Hamzah, M., Wan Daud, W. A. A., & Muhamad Romli, T. R. (2022). The impact of mobile digital game in learning Arabic language at tertiary Level. Contemporary Educational Technology, 14(1).

Al-Sabbagh, K. W., Bradley, L., & Bartram, L. (2019). Mobile language learning applications for Arabic speaking migrants—a usability perspective. Language Learning in Higher Education, 9(1), 71–95.

Anam, R., & Abid, A. (2020). Usability study of smart phone messaging for elderly and low-literate users. International Journal of Advanced Computer Science and Applications, 11, 108–115.

Cao, Y., Qin, X., Li, J., Long, Q., & Hu, B. (2020). Exploring seniors’ continuance intention to use mobile social network sites in China: a cognitive-affective-conative model. Universal Access in the Information Society, 1–22.

ciramella, K. E. (2017). The effects of Kahoot! on vocabulary acquisition and retention of students with learning disabilities and other health impairments: Rowan University.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297–334.

DeVellis, R. F. (2017). Scale development: Theory and applications thousand oaks, CA: Sage.(4th ed.).

Hashim, H., Rafiq, R. M., & Md Yunus, M. (2019). Improving ESL learners’ grammar with Gamified-learning. Arab World English Journal (AWEJ) Special Issue on CALL(5).

Horbatiuk, L., Kravchenko, N., Aliksieieva, H., Lipych, V., & Rozumna, T. (2019). Use of Mobile Applications for Foreign Language Lexical Competence Formation. Journal of History Culture and Art Research, 8(3), 113–124.

Huang, Y.-M., Lou, S.-J., Huang, T.-C., & Jeng, Y.-L. (2019). Middle-aged adults’ attitudes toward health app usage: A comparison with the cognitive-affective-conative model. Universal Access in the Information Society, 18(4), 927–938.

Islam, M. N., Ahmed, M. A., & Islam, A. N. (2020). Chakuri-bazaar: A mobile application for illiterate and semi-literate people for searching employment. International Journal of Mobile Human Computer Interaction (IJMHCI), 12(2), 22–39.

ISO. (1998). 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs). The International Organization for Standardization, 45(9).

Kalleny, N. K. (2020). Advantages of Kahoot! Game-based formative assessments along with methods of its use and application during the COVID-19 Pandemic in various live learning sessions. Journal of Microscopy and Ultrastructure, 8(4), 175.

Khaleel, F. L., Ashaari, N., Meriam, T., Wook, T., & Ismail, A. (2016). The architecture of dynamic gamification elements based learning content. Journal of Convergence Information Technology, 11(3), 164–177.

Klimova, B. (2020). Benefits of the Use of Mobile Applications for Learning a Foreign Language by Elderly Population. Procedia Computer Science, 176, 2184–2191.

Landrum, R. E., Carlson, H., & Manwaring, W. (2009). The relationship between time to complete a test and test performance. Psychology Learning & Teaching, 8(2), 53–56.

Lin, C. J., & Ho, S.-H. (2020). The development of a mobile user interface ability evaluation system for the elderly. Applied Ergonomics, 89, 103215.

Madan, A., & Dubey, S. K. (2012). Usability evaluation methods: A literature review. International Journal of Engineering Science and Technology, 4(2), 590–599.

Marques, A. D. B., Moreira, T. M. M., Jorge, T. V., Rabelo, S. M. S., Carvalho, R. E. F. L. D., & Felipe, G. F. (2020). Usability of a mobile application on diabetic foot self-care. Revista brasileira de enfermagem, 73, e20180862.
Md. Yunus, M., & Azmanuddin bin Azman, M. (2019). Memory Stay Or Stray?: Irregular Verbs Learning Using Kahoot! *Arab World English Journal* (5), 206–219. https://doi.org/10.24093/awej/call5.15

Mehrban, M., & Asif, M. (2010). Challenges and Strategies in Mobile Phones Interface for elder people. In *Morrison, D., & Koole, M. (2018). Learning on-the-go: Older adults’ use of mobile devices to enhance self-directed, informal learning. Journal of Interactive Learning Research, 29*(3), 423–443.

Mustaffa, F. Y., Salam, A. R., Maskun, Z. F., & Abdullah, A. H. (2019). Determining Areas of Improvement in Quranic Arabic Vocabulary Learning Mobile Applications through Analysis of App User Reviews. *International Journal of Engineering and Advanced Technology, 8*(5C), 1188–1191. https://doi.org/10.35940/ijeat.E1168.0585C19

Nahar, L., Jaafar, A., & Sulaiman, R. (2019). Usability evaluation of a mobile phone based braille learning application" mbraille". *Malaysian Journal of Computer Science*, 108–117.

Nayebi, F., Desharnais, J.-M., & Abran, A. (2012). The state of the art of mobile application usability evaluation. Paper presented at the 2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE).

Nayebi, F., Desharnais, J.-M., & Abran, A. (2013). An expert-based framework for evaluating ios application usability. Paper presented at the 2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*: Routledge.

Patmanthara, S., Yuliana, O., Dwiyanto, F., & Wibawa, A. (2019). The use of ladder snake games to improve learning outcomes in computer networking. *International Journal of Emerging Technologies in Learning (iJET), 14*(21), 243–249.

Saksirisathaporn, K. (2020). A Game-Based Learning Approach to Improve Students’ Spelling in Thai. *(IJACSA) International Journal of Advanced Computer Science and Applications, 11.*

Salam, A., Mustaffa, F., Sharif, N. M., & Sirri, A. (2019). Features and Teaching/Learning Activities Used In Educational Android Mobile Applications to Teach Quranic Arabic Vocabulary. *Sign, 7*(A8), A9.

Smiderle, R., Rigo, S. J., Marques, L. B., de Miranda Coelho, J. A. P., & Jaques, P. A. (2020). The impact of gamification on students’ learning, engagement and behavior based on their personality traits. *Smart Learning Environments, 7*(1), 1–11.

Tundjungsari, V. (2018). Mobile Learning Design Using Gamification for Teaching and Learning in Algorithms and Programming Language. Paper presented at the International Conference on Interactive Collaborative Learning.

Wang, A. I., & Tahir, R. (2020). The effect of using Kahoot! for learning–A literature review. *Computers & Education, 149*, 103818.

Wardaszko, M., & Podgórski, B. (2017). Mobile learning game effectiveness in cognitive learning by adults: A comparative study. *Simulation & Gaming, 48*(4), 435–454.

Wichadee, S., & Pattanapichet, F. (2018). Enhancement of performance and motivation through application of digital games in an English language class. *Teaching English with Technology, 18*(1), 77–92.

Yeni, S., Cagiltay, K., & Karasu, N. (2020). Usability investigation of an educational mobile application for individuals with intellectual disabilities. *Universal Access in the Information Society, 19*(3), 619–632.

Zakaria, N. Y. K., & Hashim, H. (2020). Game-Based Assessment in Academic Writing Course for Pre-Service Teachers. *TESOL International Journal, 15*(1), 65–73.

Zarzycka-Piskorz, E. (2016). Kahoot it or not? Can games be motivating in learning grammar? *Teaching English with Technology, 16*(3), 17–36.

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Authors and Affiliations

Syahida Mohtar$^1$ · Nazean Jomhari$^2$ · Nor Azyra Omar$^2$ · Mumtaz Begum Peer Mustafa$^2$ · Zulkifli Mohd Yusoff$^3$

Nazean Jomhari
nazean@um.edu.my

Nor Azyra Omar
1712032@siswa.um.edu.my

Mumtaz Begum Peer Mustafa
mumtaz@um.edu.my

Zulkifli Mohd Yusoff
zulkifliy@um.edu.my

$^1$ Department of Software Engineering, Faculty of Information & Communication Technology, Universiti Teknikal Malaysia Melaka Hang Tuah Jaya, Durian Tunggal, 76100 Melaka, Malaysia

$^2$ Department of Software Engineering, Faculty of Computer Science and Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

$^3$ Department of Al-Quran and Al-Hadith, Academy of Islamic Studies, Universiti Malaya, 50603 Kuala Lumpur, Malaysia