Original Research Paper

User Experiences and Usability Evaluation of COVID-19 Application

Shimaa Abdullah Nagro

College of Computing and Informatics, Saudi Electronic University, Saudi Arabia

Abstract: Tawakkalna is an application developed by the Saudi Arabian government to manage the spread of COVID-19 in the country. Even now that the pandemic is under better control, Tawakkalna continues to offer numerous services to its users. People of all age groups, including children and adults, rely on the application on a daily basis. It is important to evaluate the perceived usability of the Tawakkalna (COVID-19 KSA) application. Accordingly, this study aims to evaluate its perceived usability from the perspective of users belonging to all age groups and to identify whether any particular age group is facing difficulties with using it. The study used two existing and well-known questionnaires: The System Usability Scale (SUS) and the Usability Metric for User Experience (UMUX). The study performed descriptive and inferential statistical analysis to present an in-depth understanding and comparison of the survey results. A total of 395 responses were received. The SUS score is 71.1. The UMUX score is 72.0. The results indicate that, overall, the SUS and UMUX scores support each other and grade Tawakkalna's perceived usability as good. The results also indicate that the mean rank of the UMUX score is higher for the 18-30 age group than for all the other age groups. In addition, children under age 12 and adults greater than age 50 face challenges while using Tawakkalna. Finally, this study presents recommendations to improve the application's usability.

Keywords: HCI, Human-Computer Interaction, System Usability Scale, Usability Metric for user Experience, Survey, COVID-19 KSA

Introduction

The 'Corona Virus' outbreak was first observed in the Chinese city of Wuhan. Shortly after, in March 2020, the World Health Organization declared COVID-19 a pandemic. It has since affected all people worldwide. The Kingdom of Saudi Arabia (KSA) is no exception. According to the statistics reported by the Saudi Ministry of Health (MOH), as of January 2023, a total of 827,658 people had contracted COVID-19 and the death toll reached 9560 (WHO, 2023).

To reduce the risk of contact transmission, governments around the world imposed several preventive measures such as curfews and lockdowns, social and physical distancing, isolation and hygiene, and face mask requirements. Lockdowns and curfews in particular drastically affected daily life, making it difficult for people to pursue routine activities and increasing healthcare needs beyond the pandemic. Social, cultural, and economic affairs were also disrupted. To manage the situation, authorities across the world such as the national health service in the United Kingdom and the ministry of health in the KSA focused on developing mobile applications that could guide social contact and thereby limit the spread of the disease. Mobile technologies have indeed played an important role in managing and monitoring the spread of COVID-19. Mobile application development has only increased as governments have recognized its numerous purposes in healthcare, monitoring, and contact tracing.

The government of Saudi Arabia developed several applications during the pandemic to serve various purposes. Tawakkalna, a geo-localization application, was developed to electronically authorize citizens to move around the city. 'Mawid' was developed to enable users to locate hospitals and make appointments for check-ups. 'Tabaud', another geo-localization application, was developed to let Saudi users report any contact with an infected person. 'Sehha' and 'Tataman' were also among the applications developed by the Government of Saudi Arabia to control the spread of COVID-19 in the KSA. The Saudi Data and Artificial Intelligence Authority (SDAIA) claimed that Tawakkalna had more than 22 million users. Almufarij and Alharbi (2022) confirmed that users of Tawakkalna represent more than 96% of the population in KSA, making it the most popular and frequently used application for COVID-19 in the KSA.

Usability testing of the Tawakkalna application is necessary to promote its use and ensure that users obtain its intended benefits. Users should be able to use these applications with convenience. A usable product must allow users to perform functions easily with little to no external facilitation. According to the International Organization for Standardization (ISO-9241-11), usability is the degree to which a service can be used by its intended users to achieve their goals in an efficient (less resource intensive), effective (task completeness), and satisfactory manner. Usability testing provides user feedback to developers, allowing them to improve the product or service. Communication between users and developers can take place in several ways. Conducting interviews, experiments, focus groups and surveys are all options, albeit resource-intensive methods. The planning, arrangements, and conducting of such activities demand large amounts of time, money, and human resources. Furthermore, these methods require users to visit a particular place and dedicate a certain amount of time and effort to meet a schedule. In restricted times, such as during COVID-19 lockdowns, online surveys were the best tool available. Online surveys are flexible, i.e., do not require respondents to be physically present at the same place and time as the researcher. Surveys in general also allow administrators to collect responses from a large number of the target population in an efficient and less resource-intensive manner (Lazar *et al.*, 2017). In human computer interaction research, several survey tools have been repeatedly validated and verified as reliable by researchers. Two of these tools, the System Usability Scale (SUS) and the Usability Metric for User Experience (UMUX), were adopted in this research.

This study presents a usability evaluation of the Tawakkalna application (COVID-19 KSA). The study aims to evaluate the perceived usability of Tawakkalna with respect to users of all age groups, especially children under 12 and adults older than 50 years of age. In addition, this study aims to identify any usability issues and to provide recommendations to address those issues. An online survey was conducted using SUS and UMUX questionnaires. Statistical analysis was then performed on the quantitative data to provide an in-depth understanding of the survey data. This study is beneficial for Tawakkalna users since it identifies usability issues that they may be facing. This study is also beneficial for developers, who can utilize user feedback and pragmatic recommendations to further improve the application. Apart from users and developers, this study also supports the Saudi Data and artificial intelligence authority and the ministry of health since improved usability of Tawakkalna may result in prolonged use and may contribute to the long-term vision of digitization per the 2030 program. Specifically, this study makes the following contributions:

- A usability evaluation of Tawakkalna (COVID-19 KSA) that identifies the application's perceived usability for all age groups
- An online survey conducted using SUS and UMUX
- Descriptive statistical analysis is performed on quantitative data to provide evidence about data variables and their relationships
- Inferential statistical analysis was performed on quantitative data to provide an in-depth understanding of the data obtained from the survey
- Comparative analysis was performed on the results of the SUS and UMUX scores to increase confidence in the evaluation results
- Insights for Tawakkalna developers and the ministry of health, including user feedback and pragmatic recommendations for the application's improvement and long-term use

Background

Research in the field of Human-Computer Interaction (HCI) has been progressing since the early 1980s (Whig *et al.*, 2022). HCI research is driven in part by the daily advancement of technology in today's systems, products, and services. However, this complex field also holds human beings as central to its research. Empirical research is one of the most popular tools of HCI research, which makes use of both quantitative and qualitative data (Lazar *et al.*, 2017). A survey, which is part of empirical research, is the method adopted in this study.

This section presents the background of the survey tools used in this study, namely SUS and UMUX, followed by an overview of the Tawakkalna application.

System Usability Scale

The System Usability Scale (SUS), originally developed by Brooke (1996), is a popular standardized questionnaire to evaluate perceived usability. It is a simple and easy-to-use method for both participants and researchers. One of the many advantages of the SUS is that it is technology independent, which makes it applicable to a wide range of systems. For instance, it has been used to evaluate the perceived usability of a clinical decision support system (Karajizadeh *et al.*, 2022) and chatbots based on artificial intelligence (Borsci *et al.*, 2022). The SUS is also widely used in industrial studies that wish to assess the usability of a product or technology. Lewis (2018b) strongly recommended SUS as a tool of the future for measuring perceived usability. Since its inception, the SUS has gone international. It has been translated and published in several languages other than English, including Portuguese (Martins *et al.*, 2015); Danish (Hvidt *et al.*, 2020); Polish (Borkowska and Jach, 2017); Arabic (Alghannam *et al.*, 2018); Persian (Dianat *et al.*, 2014); and Italian (Borsci *et al.*, 2009).

The SUS questionnaire consists of 10 items and uses a five-point Likert scale, where 1 represents 'strongly disagree' and 5 represents 'strongly agree'. The items alternate between positive and negative assessments. Even-number items have a positive tone. Odd-number items have a more negative or disapproving tone. Figure 1 presents the original SUS. Lewis and Sauro (2009) observed that, in this version, items 4 and 10 are evaluating different factors as compared to the other questions. Indeed, items 4 and 10 are related to learnability whereas the remaining questions are related to usability.

The SUS results in a single value that can be easily interpreted by researchers from diverse fields, for instance, human resource managers as well as software developers. As compared with other surveys like Software Usability Measurement (Kirakowski, 1996) and the Web Site Analysis and Measurement Inventory (WAMMI) (Kirakowski *et al*., 1998), the SUS is not protected by copyright or trademark and is much more concise, meaning it is less likely to yield fewer or incomplete responses.

Usability Metric for User Experience

The usability metric for user experience (UMUX) was developed at Intel by Finstad (2010). The aim was to produce a shorter yet more detailed evaluation tool to determine the perceived usability of a product. The UMUX has only four items, as compared with the 10 items in the SUS. These four items are selected from an initial pool of 12 items. Each of the four items represents one aspect of the ISO usability definition. The UMUX was developed with the ISO usability standards in mind. It thus has a close correspondence with the ISO usability definition (i.e., efficiency, satisfaction, and effectiveness). Figure 2 presents the items present in a standard UMUX questionnaire. Unlike the SUS, the UMUX uses a seven-point Likert scale, ranging from 'strongly agree' (1) to 'strongly disagree' (7). This scale generates a more distributed and detailed score. Finstad (2010) claimed that users are more likely to score in points, for instance, 3.5 instead of 3, when presented with a seven-point scale as compared with a five-point scale. The SUS's five-point scale does not allow for this score interpolation. Several studies have found the UMUX to be a reliable, valid, and sensitive tool for evaluating the perceived usability of a product or service (Berkman and Karahoca, 2016).

SN	Items In standard System Usability Scale (SUS) Version	Strongly Disagree 1	2	3	4	5	Strongly Agree 6
1.	I think I would like to use this system frequently	o	o	o	o	o	o
2.	I found the system unnecessarily complex	o	o	o	o	o	o
3.	I thought the system was easy to use	o	o	o	o	o	o
4.	I think I would need the support of a technical person to be able to use this system	o	o	o	o	o	o
5.	I found the various functions in the system were well integrated	o	o	o	o	o	o
6.	I thought there was too much inconsistency in this system	o	o	o	o	o	o
7.	I would imagine that most people would learn to use this system very quickly	o	o	o	o	o	o
8.	I found the system very cumbersome to use	o	o	o	o	o	o
9.	I felt very confident using the system	o	o	o	o	o	o
10.	I needed to learn a lot of things before I could get going with this system	o	o	o	o	o	o

Fig. 1: Standard version of system usability scale

SN	Standard Usability Metric For User Experience (UMUX) Items	Strongly Disagree 1	2	3	4	5	6	Strongly Agree 7
1.	This system capabilities meet my requirements	o	o	o	o	o	o	o
2.	Using this system is a frustrating experience	o	o	o	o	o	o	o
3.	This system is easy to use	o	o	o	o	o	o	o
4.	I have to spend too much time correcting things with this system	o	o	o	o	o	o	o

Fig. 2: Standard version of usability metric for user experience

Tawakkalna Application (COVID-19 KSA)

The Tawakkalna application won the United Nations public service award 2022 for institutional resilience and innovative responses to the COVID-19 pandemic (SPA, 2022). Tawakkalna was developed by the Saudi data and artificial intelligence authority to manage and control the spread of COVID-19. 'Tawakkalna' is an Arabic word that means 'We Trust'. The primary goal of the application was to automatically authorize citizens and residents to move around the city to procure essentials and goods during restricted periods. Tawakkalna allotted e-permits to users who requested to travel within their city. Every individual was allowed to use the application whether they belonged to the public or private sector and whether they were citizens or resident. One of the application's functionalities was to alert users to any violations made during curfew time. The Saudi government imposed heavy fines on curfew violators, a necessary step to minimize the spread of the disease. A user could also report a suspected COVID-19 case to authorities and raise an objection regarding any gathering that violated the standard operating procedures. Some people were exempted from curfew in order to let them travel to and from their workplaces. These exempted parties included healthcare, pharmaceutical, and security workers and certain officials and government staff. These people received special work permits which they could show in case of investigation. In addition to healthcare workers, special cases such as students and drivers also received emergency permits. The Saudi government also made allowances to protect the physical and mental health of its people. It allowed special jogging permits so that people could breathe fresh air and improve their physical and mental health. The application allowed relevant entities to submit their employees' work licenses so that these employees could continue to perform their crucial duties during the pandemic. In addition, Tawakkalna generated live QR codes that security personnel could check upon a user entering any public place. Figure 3 presents some of the numerous functionalities of Tawakkalna.

Even now that the pandemic is under better control, Tawakkalna continues to offer numerous services to its users. One of its functions is the reservation of appointments related to COVID-19. It allows users to book appointments to get vaccinated and to get themselves tested for COVID-19. The Tawakkalna application effectively monitors COVID-19 results afterward and changes users' health status accordingly. It also alerts users to exposure to other infected individuals. Furthermore, the application connects several services simultaneously to the internet, thus enabling man-to-machine and machine-to-machine interactions. For example, it tracks digital documents such as the 'Iqamah' (residence card) and other important official documents by exploiting the 'Absher' service. It also tracks a person's immunity record with respect to the vaccine doses and generates coloured QR codes. One recent feature added to Tawakkalna is the ability for users to authorize a gathering and get QR codes for all invited guests and visitors.

Muslims from all over the world visit Saudi Arabia to perform the Haj (the holy pilgrimage) and Umrah. The Haj is annual, while Umrah is performed throughout the year (Al-Ajarma, 2021). Tawakkalna enables citizens and residents of Saudi Arabia to apply for 'Eatmarna' permits for performing the Haj and Umrah. It also offers 'Traffic' services by showing committed traffic violations. Moreover, it shows children's education status and results through 'Noor' services.

Tawakkalna performs still more functions in addition to the aforementioned. In short, it is a complicated yet important application for its users. It is crucial to determine its usability and user experience so that it can be improved further.

Usability Research and User Experience

Kaya *et al*. (2019) evaluated the usability of several mobile applications among Turkish users with respect to two popular operating systems. The applications were Facebook, WhatsApp, and YouTube. The two operating systems were android and apple IOS. The researchers conducted a SUS-based survey in which 222 responses were received. The results indicated that there was no significant difference in terms of users using android or an apple IOS operating system. In general, the usability scores of all applications were accepted. In comparison, WhatsApp's usability score was the highest, whereas Facebook was categorized as the application with the lowest usability. Klug (2017) highlighted the areas where the SUS has been successfully applied. These areas included print media, websites, systems, and medical technology. In addition to the applicability of the SUS, researchers have also outlined the benefits, limitations, and best practices of using the SUS. Mclellan *et al*. (2012) highlighted an important factor related to the level of user rating. Namely, users' experience with a product or service can severely affect how users rate it. A new user might feel more frustrated with a product as compared with an experienced user. Furthermore, the same users, even after they gain sufficient experience with an application, might still provide different ratings on the same application. This factor must be considered when evaluating a product or service. Longitudinal studies e.g., (Kjeldskov *et al*., 2005; Mclellan *et al*., 2012) have evaluated these factors. Lewis (1995) stated that users who have used an application previously tend to rate that application useable up to 11% more often that users who have not used the application at all. Tullis and Stetson (2004) evaluated the reliability of five questionnaires

in terms of conclusive abilities. The results indicated that the SUS led to a correct conclusion 90% of the time when the sample size was greater than 12 (Tullis and Stetson, 2004). Lewis and Sauro (2009) verified that the SUS is not one-dimensional. Rather, the SUS supports two factors. The first factor is usability and the second is learnability. The authors emphasized that practitioners and researchers can acquire additional information by using the same SUS. Namely, two items on the SUS, i.e., items 4 and 10, evaluate learnability, while the remaining eight items evaluate usability.

Lewis (2018a) investigated whether the SUS and the Computer System Usability questionnaires evaluate the same or different factors. Lewis also examined how different versions of the UMUX, such as the UMUX-Lite (two-point regression adjusted), relate to the SUS, as both surveys ultimately evaluate perceived usability. Scores of the three independently developed questionnaires were normalized to a common 0-100-point scale. The results showed a close correlation among all three, confirming that all three measure the same perceived usability. In other words, one could be used as an alternative to another. However, the SUS remained the most popular of the three.

Lewis *et al.* (2013) presented a smaller version of the UMUX, the aforementioned UMUX-lite, which has only two items. Psychometric analysis of the two-item UMUX yielded positive results. Furthermore, as an experimental treatment, two surveys (i.e., the SUS and UMUX-Lite) were conducted on two systems, one with good and one with poor perceived usability. Both the SUS and UMUX produced the same results, leading the authors to conclude that in cases where the SUS is too lengthy to conduct, the UMUX-lite is a good shorter alternative.

Berkman and Karahoca (2016) evaluated the sensitivity and reliability of the UMUX and its shorter version the UMUX-lite. A psychometric analysis was conducted to assess the subjectivity present in different individuals. The results of the experiment indicated that the metric conforms to ISO standards and supports the ISO definition's attributes. However, the results of this study do not align with the results of the previous research; instead, it suggests that rather than the UMUX-lite, the UMUX corresponds better with the SUS. Furthermore, the authors recommended that practitioners use both the UMUX and the UMUX-lite to measure perceived usability, as both surveys' validity sensitivity, and reliability are proven.
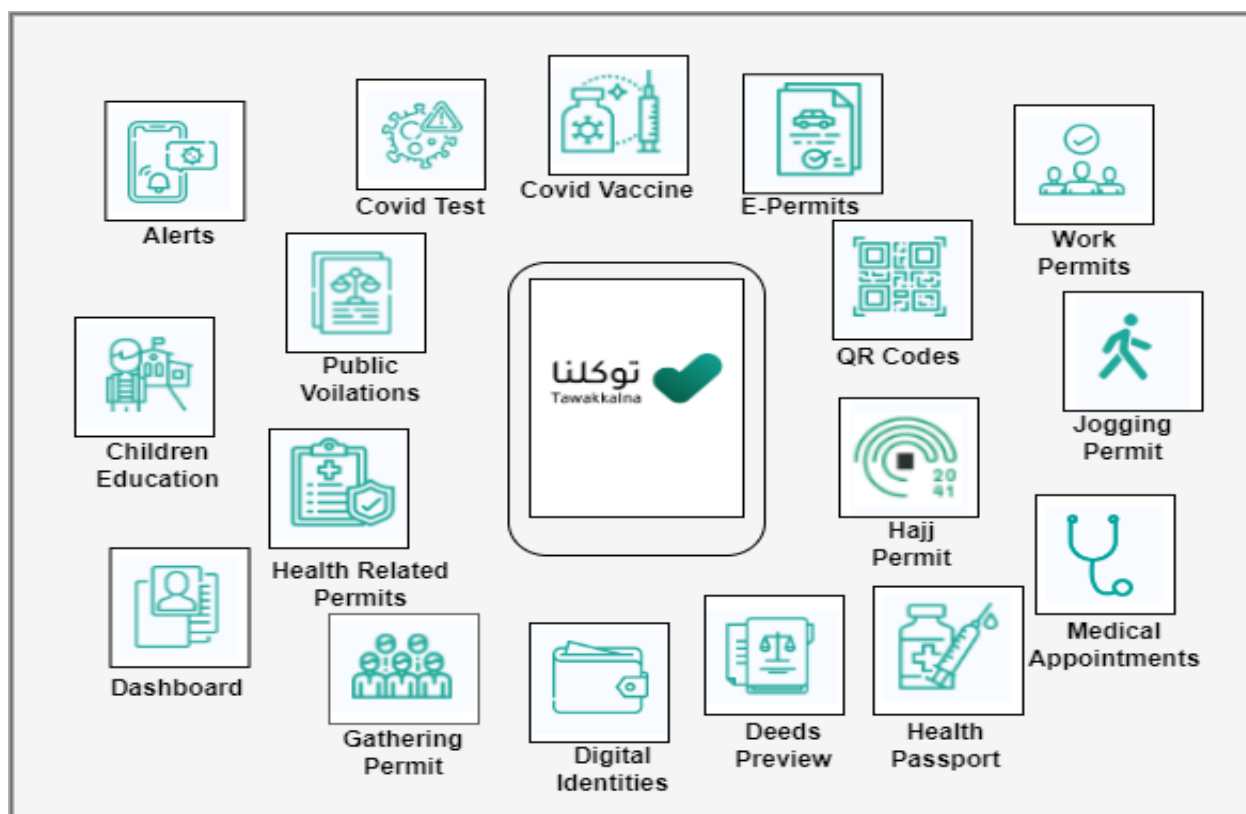

Fig. 3: Functionalities of the Tawakkalna application

Algothami and Saeed (2021) performed a user experience evaluation of Tawakkalna. They conducted a User Experience Questionnaire (UEQ), gathering completed surveys from 87 participants in Saudi Arabia. The UEQ was translated into Arabic and English to serve various ethnic populations. The UEQ was distributed using email and social media applications. The researchers' main goal was to evaluate user experience by examining several attributes: Dependability, novelty, efficiency, stimulation, and self-expression. The results of the survey indicated that user experience with the COVID-19 KSA mobile application was good in terms of attractiveness and practical qualities such as dependability. In comparison to attractiveness and other qualities, the authors found that the Saudi population did not recognize the novelty aspect of this application, so developers could ruminate on improvements in this area. The present study built on this study in particular by increasing the sample size and applying other usability evaluation metrics, namely the SUS and UMUX.

Booday and Albesher (2021) evaluated the usability of five mobile-based applications developed by the Saudi government to manage the spread of COVID-19. To evaluate usability, the authors applied smart heuristics developed specifically for mobile-based applications. The authors argued that mobile-and web-based applications are different in terms of computation capabilities since mobile devices have limited resources as compared to computers. As a result, the authors used smart heuristics with certain amendments to evaluate the usability of mobile applications in the KSA. Five expert evaluators were employed to judge the usability of these applications. The results indicated that the mobile-based applications that Saudis have used during the pandemic pose certain usability issues. The authors highlighted usability issues and offered recommendations on how to improve the usability of these mobile applications. Only five HCI experts determined the usability issues in mobile applications. However, a small group of experts does not represent a complete population. A good sample size and users from all age groups must be employed to evaluate the usability of a mobile application.

Mzoughi and Garrouch (2021) conducted a survey to investigate the resistance expressed by the Saudi population towards adopting technology-based preventive measures such as the Tawakkalna and Tabaud applications against the spread of COVID-19. The authors hypothesized that media coverage plays a positive role in spreading risk awareness among the public. The results obtained after statistical analysis indicated that the media has a positive impact on spreading risk awareness of COVID-19. Furthermore, advertisements and other awareness campaigns also reduced the resistance to the use of geo-localization applications such as Tawakkalna.

Alharbi *et al.* (2021) conducted a survey and interviews with older adults living in the KSA. The aim of this research was to identify the difficulties that adults face while using geo-localization and other pandemic-related applications. The authors argued that the older population often hesitates to adopt technology, which prevents this population from performing routine activities on their own without seeking help from others. As a result, this population becomes more vulnerable to encountering infected persons. A total of 397 survey responses were received. The results indicated that older adults and people with less understanding of technology face many challenges on a daily basis. The results of the surveys and interviews also indicated that the older population in the KSA requires help when using COVID-19 tracking applications such as Tawakkalna and Sehatay. The authors emphasized the need to enhance the usability, interactivity, and design of such applications to facilitate the older population.

Alanzi (2022) investigated the satisfaction level of users using m-health applications in the KSA before and after Covid-19 restrictions were lifted. The authors aimed to determine whether the increase in the use of m-health applications occurred due to users' own interests or because they were forced to use the applications due to pandemic restrictions. For this reason, the authors performed a survey using the mHealth application usability questionnaire in which 318 people participated. The results indicated that, overall, m-health applications have specifically provided ease of use, usefulness, and effectiveness. At the same time, more than 75% of the population reported being less likely to use m-health applications once the pandemic restrictions are lifted. In another study, Almufarij and Alharbi (2022) evaluated the perception, awareness, and use of m-health applications developed by the Saudi government to manage the pandemic situation. A total of five applications were evaluated: Tawakkalna, Tabaud, Mawid, Seha, and Tataman. The authors designed a survey and statistically analyzed the results of 876 respondents. The results indicated that in general the KSA population was aware of and satisfied with all five m-health applications and perceived the applications as offering clear benefits. The results also indicated that, as compared to the other four applications, Tawakkalna was the most popular and most frequently used application, adopted by more than 96% of the population.

Research Questions

RQ 1: What is the usability level of the Tawakkalna application?

Motivation: This research question aims to evaluate the perceived usability of Tawakkalna based on two types of existing questionnaires and their respective scores.

RQ 2: What is the usability level of Tawakkalna with respect to users of all age groups?

Motivation: This research question aims to identify the usability level of Tawakkalna with respect to six age

groups: <12, 12-17, 18-30, 31-40, 40-50, and >50. To answer this research question, SUS and UMUX scores were calculated for every age group. Mean ratings and standard deviations for each item are presented for both surveys. Furthermore, inferential statistical analysis was performed to identify the differences in the mean scores of both questionnaires. This research question helped the study to dig deeper into the issues faced by each age group.

Materials and Methods

This study adopted a cross-sectional survey methodology to evaluate the perceived usability of the Tawakkalna application. The online survey was used as a tool to collect data from a large number of geographically dispersed users. The online survey, thanks to the anonymity it can offer, was a better choice than the face-to-face survey because the participants were more likely to self-disclose honestly and to share their honest negative opinions.

Questionnaire

Two well-known tools were used to evaluate the perceived usability of the Tawakkalna application: The System Usability Scale (SUS) and the Usability Metric for User Experience (UMUX). The estimated time required to complete the questionnaire was approximately 3-5 min. The questionnaire was developed using Microsoft forms and distributed through email and social media applications. Arabic and English versions of the questionnaire were developed to facilitate users of different ethnicities.

The questionnaire was composed of three sections. All the items in the survey were closed-ended. The first section contained general information about the user, including gender, age, citizenship, education, and operating system used for accessing the Tawakkalna application. The second section was based on the 10 items of the SUS. The items in the SUS alternate between positive and negative tones, with even-numbered items positive and odd-numbered items more negative or disapproving. Figure 1 presents the original SUS. The third section consisted of the four items of the UMUX. The UMUX items closely correspond with the ISO usability standard. As in the SUS, the UMUX contains positive and negative items. Figure 2 presents the standard UMUX items. The standard versions of the two tools were used, with the phrase 'the system' replaced by 'Tawakkalna'. It is a common practice to alter a generic version to apply it to a more specific context.

A five-point Likert scale was used to collect responses to each item. In the five-point Likert scale, there were five options for every question as follows: Strongly disagree (1), Disagree (2), Neither agree nor disagree (3), Agree (4), and strongly agree (5).

Scoring SUS and UMUX

Scores were calculated by converting the responses to their respective numeric representation. For instance, 'strongly agree' was replaced by 5, and 'agree' and 'disagree' were replaced by 4 and 2, respectively. This calculation was done for both SUS and UMUX responses. For the SUS, the second step was calculating the differences, for which the scores of odd items were subtracted by 1 and the scores of even items were subtracted by 5. This difference calculation was done to make the scores comparable, as the original scores were not comparable due to the alternating positive and negative tone in even-and odd-numbered items. This step brought all the values within the range of 0-4, making 4 the maximum positive score irrespective of question tone. In the third step, the sum of the newly calculated scores was computed. Next, the sum was multiplied by 2.5 to bring all the scores within the range of 0-100. As a result, score interpretation became much easier. Finally, the average of the scores was calculated to produce the final SUS score. A score in the range of 40-50 earned an F grade. The range of 50-60 earned a D grade, the range of 60-70 earned a C grade, the range of 70-80 earned a B grade and the range of 80-100 earned an A grade.

For the UMUX, the first step was the same, i.e., to code the responses into numeric data. In the second step, differences were calculated, with 1 subtracted from odd items and even items subtracted from 5. In the third step, the sum of the newly calculated values was computed and divided by 16. In the fourth step, the values were multiplied by 100 to bring them into the range of 0-100. Finally, the average of the scores was calculated to produce the final UMUX score.

Sampling

Probability sampling was adopted in this survey. This method involves the random selection of individuals where everyone in the population has an equal (greater than zero) chance of getting selected. Probability sampling allows researchers to infer strong relations for a whole population (Lazar *et al*., 2017). Since Tawakkalna is being used by almost every citizen and resident in Saudi Arabia, the inclusion criteria treated all users (Saudis and residents) from different age groups as part of the target population. It was advised to supervise children under 18 while they filled out the form. According to Krejcie and Morgan (1970), for a population of 10,000,000, the sample size must be 384. The form was active from 8th April to 27th June 2022. At the end of the survey activation period, a total of 395 responses were received.

Ethical Considerations

At the beginning of the questionnaire, a consent form was presented to the participant. The form contained necessary information about the research such as the research title and the purpose for which the collected data would be used. Ethical approval for the research was secured by the Saudi electronic university research ethics committee. Participation of users was voluntary. To maintain anonymity, no personal information (e.g., name, contact number, etc.) was collected.

Statistical Analysis

Data were analyzed using Excel, R, and R-studio. The scores for the SUS and UMUX were calculated using Excel. Before analysis, the data were cleaned by the researcher to ensure that there was no invalid, incomplete, or repeated responses. Invalid responses included those where the individual did not meet the qualification criteria. Incomplete responses were those where some items were left unanswered. Repeated responses were those where an individual submitted more than one response. Descriptive statistical analysis was performed by calculating average ratings and standard deviations for each individual item.

Results

A total of 395 participants responded to the survey. This section presents the results. Descriptive statistical analysis was conducted to calculate the frequencies, mean ratings, and standard deviation of responses. According to the results, participants of all age groups participated in the survey. Regarding gender distribution, 47.3% of participants were female and 52.6% of respondents were male. For age-related data, the majority of the sample population, i.e., 37%, belonged to the 18-30 age group. The representation of children under 12 was 5%. Adults in the 41-50 age group constituted 17% of the total sample. Children in the 12-17 age group represented 16% of the sample and adults in the greater than 50 age group constituted 10% of the sample. Figure 4 presents the age distribution. Regarding the operating system used by respondents, the majority of the participants, i.e., 73.6%, used IOS Apple devices to run Tawakkalna and 26% used an Android operating system. One of the items in the questionnaire was about the nationality of the user. The results indicated that the majority of the respondents, i.e., 87.8%, were Saudi nationals while only 10% were residents. Considering education levels, 15% of the respondents mentioned a Ph.D. as their highest obtained degree and 14.6% reported their highest qualification as a master's degree. Around 38% of the sample reported a bachelor's as their highest degree, around 5.5% of the population reported intermediate,

19% secondary, and 15.5% primary school. Figure 5 presents the level of education of participants. Table 1 presents the information about the sample overall:

RQ 1: What is the usability level of the Tawakkalna application?

To answer this question, the SUS and UMUX scores were computed. Inferential statistical analysis was performed to compare the scores of the two tools. The overall SUS score for the complete population was 71.1, which means usability lies within a good range and can be assigned grade B. Table 2 presents the mean rating and standard deviation for the complete population. Here it can be seen that 95% of the users reported that Tawakkalna is easy to use. Around 71% of the users felt very confident while using Tawakkalna. Furthermore, 76% of the population perceived that other people could quickly learn to use this application and 69% of the population agreed that most of the functionalities of Tawakkalna are well integrated. Finally, 71% of respondents believed that Tawakkalna is not complex and 76% believed that they did not need any technical support while using Tawakkalna. However, 24% of the users believed that they did need technical support while using this application and 65% thought they would like to use Tawakkalna often. Only 12% of the population found that there was a lot of inconsistency in Tawakkalna. In contrast, 63% did not agree that there are inconsistencies in Tawakkalna. In addition, 71% of the population was of the opinion that they did not need to learn a lot of skills to use Tawakkalna, whereas 19% believed they would need to learn a lot of skills before they could use this application. Lastly, 64% of the users believed that Tawakkalna is not awkward to use.
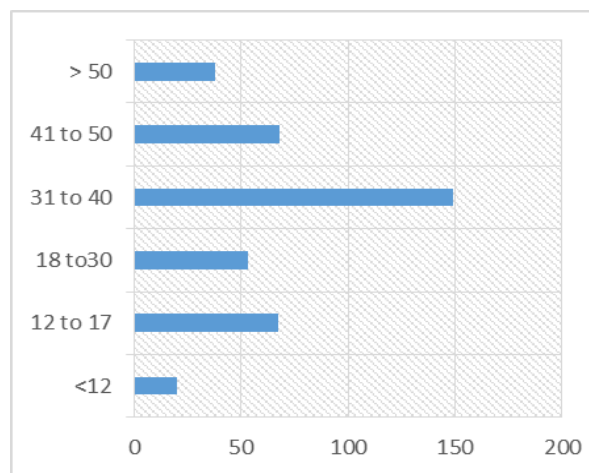

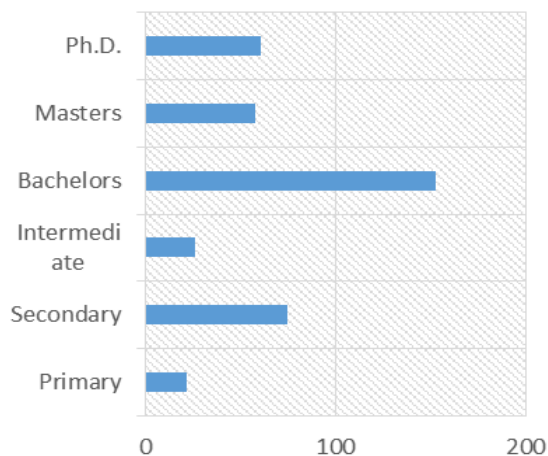**Fig. 4:** Number of respondents in each age group

Fig. 5: Number of respondents corresponding to each education category

Table 1: Information about participants

Characteristics	Frequency	Percentage
Gender		
Male	187	47
Female	208	52
Age (years)		
<12	20	5
12-17	67	17
18-30	53	13
31-40	149	37
41-50	68	17
>50	38	10
Operating system		
IOS apple	291	73
Android	104	26
Population		
Saudi	347	87
Resident	48	12
Application		
The mobile application	355	90
Both	40	10
Education		
Primary school	22	6
Secondary school	75	19
Intermediate school	26	7
Bachelor's degree	153	39
Master	58	15
Ph.D.	61	15

Table 2: Overall mean ratings of SUS scores

Standard SUS items	Mean rating	Standard deviation
Item1	3.6	3.3
Item 2	2.1	1.9
Item 3	4.3	3.8
Item4	2.0	1.8
Item5	3.7	3.3
Item 6	2.3	1.9
Item 7	3.9	3.4
Item 8	2.3	2.0
Item 9	3.9	3.5
Item 10	2.2	1.9

Focusing on the UMUX score, the overall score is 72.0, which also indicates the good perceived usability of the Tawakkalna application. Table 3 presents the mean ratings and standard deviations of the UMUX items. More than 90% of the respondents believed that Tawakkalna is easy to use, whereas only 2% found Tawakkalna to be a difficult application to use. Of all the respondents, 71% believed Tawakkalna fulfills their requirements and only 6% believed otherwise. Around 71% of the respondents did not consider Tawakkalna usage to be an unsatisfying experience and 9.3% believed it is frustrating to use. Around 53% of respondents felt they spent a lot of time fixing things with Tawakkalna, whereas 16% believed they had to spend too much time fixing things with this application and around 31% of respondents felt neutral, i.e., neither agreed nor disagreed, about this item (i.e., item 4 of the UMUX).

Inferential statistical analysis was also performed to compare the UMUX and SUS scores. First, normality tests such as Shapiro-Wilk and plots such as boxplots and histograms were developed to identify whether the data were parametric or non-parametric. These tests confirmed that the data were not normally distributed (non-parametric). This led to the application of paired Mann Whitney U test. The alpha was set to 0.05. The null hypothesis was as follows: H_0 = The median difference between pairs of SUS and UMUX is zero. The alternate hypothesis was as follows: H_1 = The median difference between pairs of SUS and UMUX is not zero. The result was p-value = 0.002<0.05. Hence, the null hypothesis was rejected.

To summarize, the perceived usability of Tawakkalna is good. The majority of the population rated it positively and did not report major problems. Furthermore, the inferential statistical analysis found that there is a difference between the ratings of the UMUX and SUS scores:

RQ 2: What is the usability level of Tawakkalna with respect to users of all age groups?

This research question is answered by calculating the SUS and UMUX scores for each of the six age groups. In addition, the scores were statistically analyzed by comparison to evaluate the perceived usability of Tawakkalna with respect to every age group. The SUS score of the under-12 age group was 69.1. Table 4 presents the mean ratings and standard deviations of the responses for this age group. In this age group, 75% of the respondents did not like to use Tawakkalna frequently. In response to the second item, 25% rated Tawakkalna as complex, whereas 75% avoided making a choice. In response to the statement that they need technical assistance while using this application, 95% of the respondents disagreed. Similarly, 95% believed that Tawakkalna did not have many inconsistencies. In response to the fifth item, which was a positive-tone statement about the functional integration of Tawakkalna, 80% of the respondents remained indecisive. Interestingly, 100% of the under-12 population felt that Tawakkalna is

easy to use and that others could learn to use it efficiently. All respondents in this age group also felt confident while using the application and needed no technical assistance for using it.

The UMUX score of the under-12 age group was 77.8, which indicates the good perceived usability of Tawakkalna. Table 5 presents the mean rating and standard deviation of the under-12 age group. The mean rating of the first item was 4.15, which indicates that about 95% of this age group believed that Tawakkalna meets their requirements.

Table 3: Overall mean ratings of UMUX score

Standard UMUX items	Mean rating	Standard deviation
Item 1	3.8	3.3
Item 2	2.0	1.7
Item 3	4.2	3.8
Item 4	2.5	2.1

Table 4: SUS score for under 12 age group

Standard SUS items SUS = 69.1 N = 20	Mean rating for <12 age group	Standard deviation
Item 1	2.4	2.0
Item 2	2.7	2.2
Item 3	4.1	3.6
Item 4	2.1	1.5
Item 5	3.2	2.7
Item 6	2.1	1.5
Item 7	4.2	3.7
Item 8	2.7	2.2
Item 9	4.2	3.7
Item 10	1.0	0.3

Table 5: UMUX score for under 12 age group

Standard UMUX items N = 20 UMUX = 77.8	Mean rating for <12 age group	Standard deviation
Item 1	4.15	3.6
Item 2	1.15	0.7
Item 3	3.35	2.8
Item 4	1.9	1.3

Table 6: SUS score for under 12-17 age group

Standard SUS items SUS = 71.3 N = 67	Mean rating for 12-17 age group	Standard deviation
Item 1	3.6	3.1
Item 2	2.0	0.8
Item 3	4.4	4.0
Item 4	2.3	1.6
Item 5	3.8	3.6
Item 6	2.4	1.4
Item 7	4.0	3.3
Item 8	2.4	1.9
Item 9	4.0	3.7
Item 10	2.2	1.5

Likewise, 95% believed that the use of Tawakkalna is not an unsatisfying experience for them. However, only 25% of these respondents found Tawakkalna easy to use. Finally, 100% of the under-12 age group believed that they do not need to spend a lot of time fixing things with this application.

For the 12-17 age group, the SUS score was 71.3, which indicates good perceived usability by this age group. Table 6 presents data for the 12-17 age group. Around 60% of the respondents in this age group would like to use Tawakkalna repeatedly. Around 71% disagreed with the second item. The third item has the highest mean rating, indicating that 97% of these users believed that Tawakkalna is easy to use. Around 69% of the users in this age group felt very confident while using Tawakkalna, though 25% remained indecisive about this item (i.e., Item 9). Furthermore, 98.5% perceived that other people could quickly learn to use this application. Around 80% agreed that most of the functionalities of Tawakkalna are well integrated. In this age group, 48% of the respondents believed that Tawakkalna is not complex, but 25% believed that it is unnecessarily complex. Though 73% felt that they do not need any technical support while using Tawakkalna, 27% believed that they do need technical support while using this application. Similarly, 27% felt that there is a lot of inconsistency in Tawakkalna. Finally, 67% of these respondents were of the opinion that they do not need to learn a lot of skills to use Tawakkalna and 70% believed that Tawakkalna is not awkward to use.

The UMUX score for the 12-17 age group was 73.5, which is an indicator of good perceived usability from users of this age group. Table 7 presents the mean ratings and standard deviations of this group. Here it can be seen that the highest mean rating is 4.4, which was calculated in response to item 3. In this age group, 97% of the respondents believed that Tawakkalna is easy to use.

For the 18-30 age group, the SUS score was 74.6, which is an indicator that the perceived usability of Tawakkalna is good. Table 8 presents the mean ratings and standard deviations of the SUS scores of this age group. Here it can be seen that the highest mean rating is for the third item. This result indicates that around 98% of this age group believed that Tawakkalna is easy to use. Around 87% believed it is easy for other people to quickly learn and use this application. In this age group, 66% of the respondents would like to use Tawakkalna frequently while 34% would prefer not to. Around 7% of these respondents felt that Tawakkalna is complex. Likewise, 7.5% felt they would need technical support to use this application. Around 40% believed that using Tawakkalna is an awkward experience. Around 89% felt confident while using this application. Only 6% felt that there are inconsistencies in Tawakkalna. Finally, 90% of the respondents in this age group felt that Tawakkalna is well integrated.

For the 18-30 age group, the UMUX score was 79.5 ≈ 80, which indicates that perceived usability is excellent. More than 94% of the respondents in this age group felt that Tawakkalna meets their requirements. In this age group, 98% of respondents believed that it is easy to use. Regarding the second item, a statement that using Tawakkalna is a frustrating experience, only 5.6% agreed while 60% disagreed and 40% remained indecisive. Regarding the fourth item, a statement that users need to spend a lot of time correcting things with Tawakkalna, only 9% agreed while 45% disagreed and another 45% remained indecisive. Table 9 presents the details for this age group.

For the 31-40 age group, the SUS score is 69.3, which is an indication that the perceived usability of this application is fair. The highest mean rating is 4.3, which was calculated by computing the response scores for the first item. This indicates that 96% of the respondents in this age group believed that Tawakkalna is easy to use. In this age group, 66% of respondents believed they would like to use Tawakkalna frequently, 68% disagreed with the statement that Tawakkalna is complex and 78% believed they do not need any technical assistance while using Tawakkalna. The mean rating of the fifth item is 3.5, which indicates that 54% of the respondents were of the opinion that the functions in Tawakkalna are well integrated. Around 11% believed that there are inconsistencies in Tawakkalna. Around 64% believed that it would be easy for other people to learn how to use Tawakkalna. Around 62% felt confident in using Tawakkalna. Only 5% of this age group believed that Tawakkalna is cumbersome to use. Finally, 69% disagreed with the notion that they needed to learn a lot of skills to be able to use Tawakkalna. Table 10 presents the details for this age group.

The UMUX score for the 31-40 age group was 69.8, which indicates that the perceived usability of Tawakkalna for this age group is fair. Table 11 presents the descriptive statistics for this age group. In this age group, 64% of the respondents were of the opinion that Tawakkalna meets their requirements. Only 6% agreed with the statement that using Tawakkalna is an awkward experience and 93% felt Tawakkalna was easy to use. Finally, 48% thought that they need to spend too much time fixing things with Tawakkalna.

For the older than 50 age group, the SUS score was 70.3, which is an indication of good perceived usability. Table 12 presents the descriptive statistics of this age group. In this age group, 68% of the respondents felt that they would like to use Tawakkalna frequently and 87% disagreed with the notion that the application is complex. In addition, 92% believed that Tawakkalna is easy to use. While 73.6% felt they do not need technical assistance while using the application, 10.5% believed they do require technical assistance. Around 89% of the respondents in this age group thought that the functions are well integrated. Around 55% disagreed with the notion that there are inconsistencies in Tawakkalna and 76% believed that others could efficiently learn to use this application. Only 10.5% agreed that using Tawakkalna

is an awkward experience. Finally, around 66% of these respondents were of the opinion that they do not need to learn a lot of skills to use this application.

The UMUX score of the over-50 age group was 68, which indicates that the perceived usability of Tawakkalna is fair. Around 79% of this age group was of the opinion that Tawakkalna meets their requirements. Around 84% believed it is very easy to use and 87% disagreed with the statement that using Tawakkalna is a frustrating experience. Around 34% believed that they need to spend too much time correcting things with Tawakkalna. Table 13 presents the details for this age group.

Table 7: UMUX score for under 12-17 age group

| Standard UMUX items | | |
| N = 67 | Mean rating for | Standard |
UMUX score = 73.5	12-17 age group	deviation
Item 1	3.8	3.4
Item 2	2.0	1.9
Item 3	4.4	3.9
Item 4	2.4	2.1

Table 8: SUS score for the 18-30 age group

| Standard SUS items | | |
| SUS = 74.6 | Mean rating for | Standard |
N = 53	18-30 age group	deviation
Item 1	3.6	3.3
Item 2	1.9	1.6
Item 3	4.4	3.9
Item 4	1.9	1.6
Item 5	4.0	3.5
Item 6	2.2	1.9
Item 7	4.5	4.0
Item 8	2.8	2.9
Item 9	4.1	3.6
Item 10	1.9	1.6

Table 9: SUS score for the 18-30 age group

| Standard UMUX items | | |
| N = 53 | Mean rating for | Standard |
UMUX score = 79.5	18-30 age group	deviation
Item 1	4.3	3.8
Item 2	2.1	1.9
Item 3	4.7	4.2
Item 4	2.4	2.1

Table 10: SUS score for the 31- 40 age group

| Standard SUS items | | |
| SUS = 69.3 | Mean rating for | Standard |
N = 149	31-40 age group	deviation
Item 1	3.4	3.1
Item 2	2.3	2.1
Item 3	4.3	3.8
Item 4	1.8	1.6
Item 5	3.5	3.1
Item 6	2.3	1.9
Item 7	3.7	3.3
Item 8	2.2	1.8
Item 9	3.8	3.4
Item 10	2.3	2.1

Table 11: UMUX score for the 31-40 age group

Standard UMUX Items		
N = 149	Mean rating for	Standard
UMUX score = 69.8	31-40 age group	deviation
Item 1	2.0	2.3
Item 2	2.1	1.8
Item 3	4.1	3.7
Item 4	2.1	3.2

Table 12: SUS score for greater than 50 age group

Standard SUS items		
SUS = 70.3	Mean rating for	Standard
N = 38	>50 age group	deviation
Item 1	3.8	3.5
Item 2	1.8	1.5
Item 3	4.2	3.8
Item 4	2.2	1.9
Item 5	4.0	3.5
Item 6	2.3	1.9
Item 7	3.7	3.3
Item 8	2.1	1.8
Item 9	3.4	3.0
Item 10	2.5	2.2

Table 13: SUS score for greater than 50 group

Standard UMUX items		
N = 38	Mean rating for	Standard
UMUX = 68.0	>50 age group	deviation
Item 1	3.8	3.3
Item 2	2.0	1.6
Item 3	4.1	3.6
Item 4	3.1	2.7

In addition to the descriptive analysis, an inferential statistical analysis was conducted to compare the SUS and UMUX scores across each age group to find any differences between the mean ranks. First, a comparison of the SUS scores across the six age groups was performed. For this reason, a null and alternate hypothesis was formulated. The null hypothesis was as follows: H_0 = The mean ranks of SUS scores across all age groups are the same. The alternate hypothesis was as follows: H_1 = The mean ranks of SUS scores across all age groups are not the same. The normality of the data was then analyzed by applying the Shapiro-Wilk test and developing boxplots and histograms. The results indicated that the data were not normally distributed. Therefore, a non-parametric test, namely the Kruskal-Walli's test, was applied with alpha = 0.05. The result was p-value = 0.84, which is greater than 0.05, hence the null hypothesis was accepted. Thus, there is no difference in the mean ranks of the SUS scores across the six age groups.

The same procedure was followed for the UMUX scores. The results indicated that the data distribution was not normal. Figures 6-7 present box plots of the SUS and

UMUX scores with respect to each age group. Figure 8 presents the UMUX and SUS scores of all age groups. Here the letters A through F represent the age groups from under 12 to over 50 in increasing order. The non-parametric Kruskal-Walli's test was applied for the UMUX scores. The null and alternate hypotheses were formulated as follows: H_0 = The mean ranks of UMUX scores across all age groups are the same and H_1 = The mean ranks of UMUX scores across all age groups are not the same. The resultant p-value = 0.0001, which is less than 0.05, hence the null hypothesis was rejected. The results of this test indicated a significant difference in the mean ranks of the UMUX scores across the six age groups. Since the Kruskal-Walli's test does not identify which group or groups differ from others, to further identify which groups were responsible for the differences in the mean ranks, a post hoc test called the Dunn test was applied. The results of the Dunn test revealed that the mean rank of the UMUX scores for the 18-30 age group was significantly different from that of all other age groups.
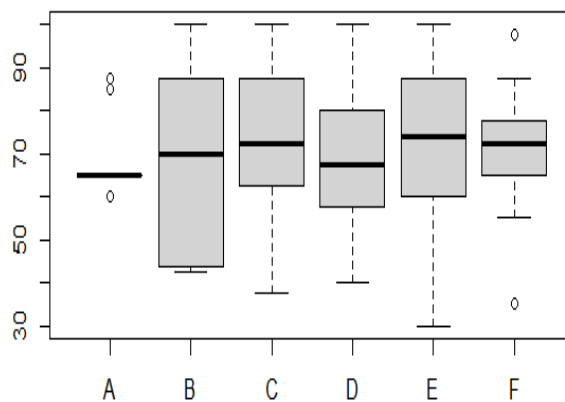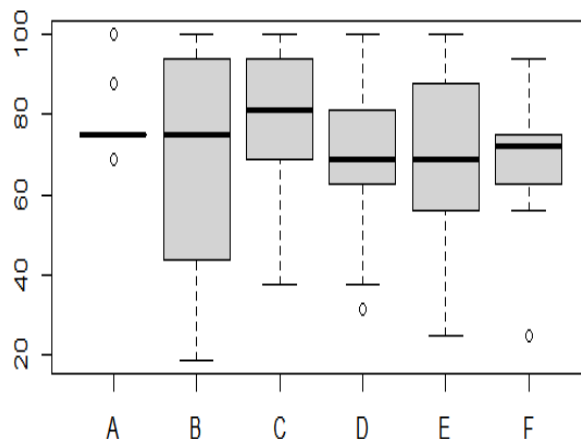

Fig. 6: Box plot for SUS score
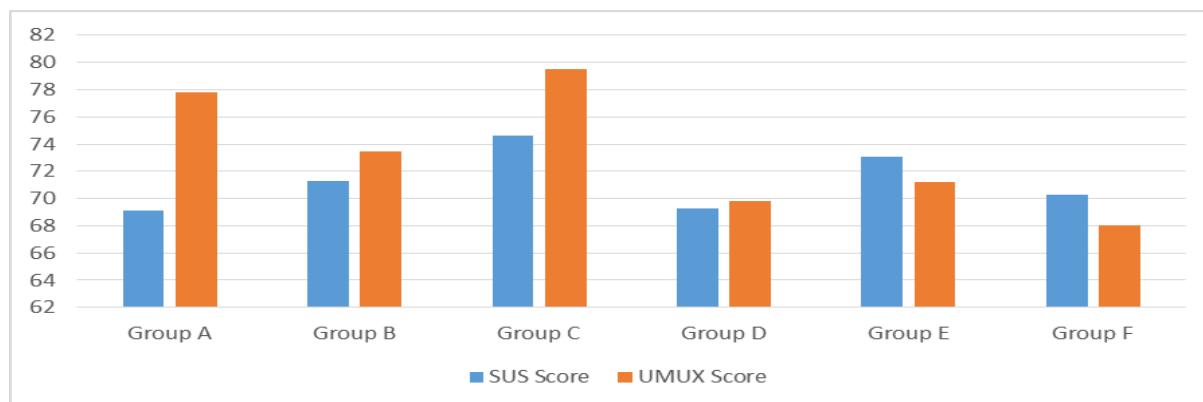

Fig. 7: Box plot for UMUX scores

Fig. 8: SUS and UMUX scores for all age groups

Discussion

This section discusses the outcomes of this study as well as its relevance to previous studies. The goal of this study was to evaluate the usability of the Tawakkalna application from the perspective of users belonging to various age groups. The study used both the SUS and UMUX in its questionnaire. These two survey tools were used together to increase the confidence levels of the evaluated results. While the SUS is reliable and widely accepted as a standard to calculate perceived usability, the UMUX is more general and closer to the ISO's definition of usability (Finstad, 2010; Peres *et al*., 2013; Berkman and Karahoca, 2016). In addition, the UMUX is shorter than the SUS and takes only a minute of respondents' time to complete. Consequently, a comparative analysis was performed to compare the results of both usability evaluations.

The first research question focused on evaluating the perceived usability of Tawakkalna with respect to users of all age groups. A total of 395 responses were received. The mean SUS score calculated was 71.1 and the mean UMUX score was 72.0. These scores indicate that, per both survey tools, the perceived usability overall was 'good'. The difference between the scores of the two tools being minor indicates that the tools complement each other well and raise the confidence level in the study's evaluation results. Comparing the results with those of Algothami and Saeed (2021), the results are consistent in terms of the overall ranking. Algothami and Saeed (2021) found that the Tawakkalna user experience ranked as 'above average' and according to the present study, the perceived usability of Tawakkalna is good. The results of the present study also align with those of Alanzi (2022),

namely that Tawakkalna users are satisfied and well aware of its functionalities and usage.

The second research question focused on evaluating the perceived usability of Tawakkalna with respect to different age groups. This question is intended to identify the usability scores given by each age group to determine whether a particular age group is facing problems using this application. The results of the inferential analysis indicated that there is no difference in the mean ranks of the SUS scores across different age groups. However, the analysis did find that the mean rank of the UMUX scores of the 18-30 age group is different as compared to the mean rank of other age groups. This age group has the highest UMUX scores and the perceived usability of Tawakkalna by this age group was excellent. The 18-30 age group's SUS and UMUX scores are similar, with both scores being the highest among all age groups.

A detailed analysis was performed by analyzing the scores of every item and comparing the results with those of other age groups. Considering the mean rating of item 1 of the SUS scale, the lowest score was assigned by the under 12 age group, which indicates that children under 12 are less likely to use the system frequently. Furthermore, as compared to other age groups, the under-12 age group gave the lowest score to item 3 of the SUS scale and remained generally indecisive on item 2 of the SUS. These results indicated that out of all age groups, children under 12 face the most difficulties in using Tawakkalna. In addition, the majority of the under-12 age group remained indecisive about item 5 of the SUS, which indicates that they are confused about the system's functionality integration. Similarly, the majority of the under-12 age group were indecisive about item 3 of the UMUX, which confirms their confusion about the application's ease of use. Figures 9-10 present the mean ratings for the UMUX and SUS scores for each item and age group.
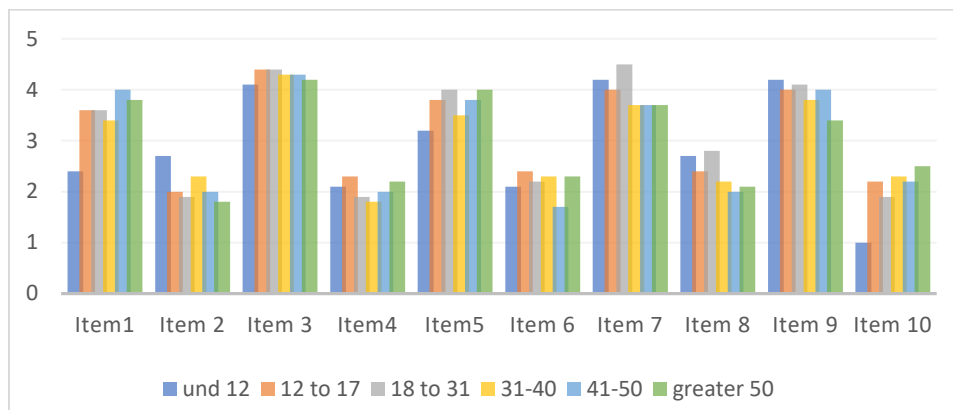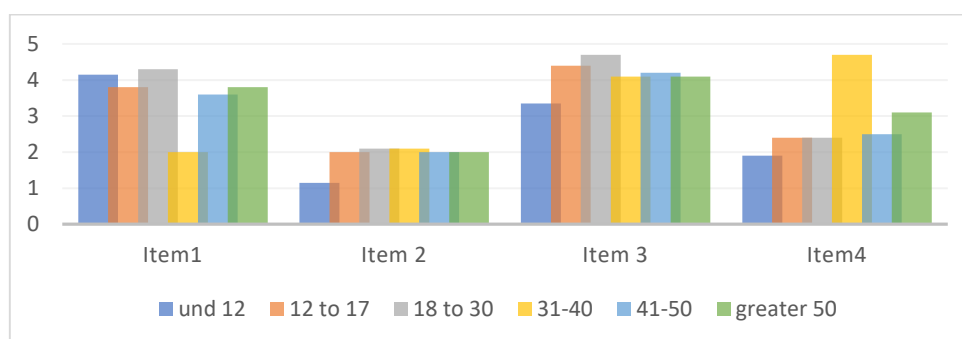
Fig. 9: Mean SUS score

Fig. 10: Mean UMUX score

The results also indicated that adults above age 50 face challenges with using Tawakkalna. In the present study, this age group felt generally hesitant about the frequent use of the application. The results also indicated that out of all age groups, adults above age 50 felt least confident while using Tawakkalna. Respondents greater than 50 felt they need to spend too much time fixing things with Tawakkalna. This result further confirms that the greater than 50 age group is facing difficulties with using Tawakkalna. These challenges may arise from this group's limited cognitive, motor, and perceptual skills. People over age 50 tend to have diminished visual and hearing abilities, fine motor skills, and information processing abilities. Moreover, their passion and motivation to learn new technical skills tend to be reduced. Thus, if they do not perceive the benefits of using technology, they will not adopt it (Liu *et al*., 2021). Awan *et al*. (2021) found that elderly people typically face two main types of issues when considering the usage of smartphones. First, as Liu *et al*. (2021) stated, one type of issue consists of age-related issues, such as memory decline, diminished cognitive and physical abilities, and weaker mental models and sensory functions. The second type of issue concerns the design of the applications. Designers tend to design software for young people, neglecting the needs of older users. Software designers must pay attention to their older users and produce solutions that can assist elderly users with interacting with new technologies using smartphones.

Based on the results, this study makes the following recommendations:

- Adopt simpler terms and graphical symbols to improve the usability of Tawakkalna for children under 12
- Ensure that guardians and family members have access to children's and adults' data so they can oversee and support the application's use
- Improve the application's interface to increase its attractiveness. Possibilities include offering different themes, colour combinations and font sizes for different age groups
- Allow and assist the relevant authorities with developing short and informative video tutorials for each task and make those tutorials available for all users
- Create an interactive helpline to facilitate users who need assistance
- Develop a 'help' feature to give tips to users while performing tasks

The findings of this study are beneficial for Tawakkalna users since it has identified several usability issues. In addition, this study is beneficial for developers since it offers user feedback and pragmatic recommendations on ways to further improve the application. Finally, this study is also useful for the saudi authority for data and artificial intelligence and the ministry of health, as improved usability of Tawakkalna may result in prolonged use of the application and even contribute to the long-term vision of digitization in the 2030 program.

Conclusion

In this study, the researcher has assessed the usability of the Tawakkalna application developed by the Saudi authority for data and artificial intelligence to manage and control the spread of COVID-19. Even now that the pandemic is more controlled, Tawakkalna continues to offer numerous services to users of different age groups. This study has evaluated the perceived usability of the Tawakkalna (COVID-19 KSA) application from the perspective of users belonging to all age groups. Two existing, well-known survey tools, namely the System Usability Scale (SUS) and usability metric for user experience (UMUX) were used to collect data from 395 respondents. The results indicate that the tools complement each other and that the overall perceived usability of Tawakkalna is good. The results of the comparative analysis highlight that the mean UMUX ranking of the 18-30 age group is different from that of other age groups. The 18-30 age group is the only age group that ranked Tawakkalna's usability as excellent. The results further indicate that children under 12 and adults over 50 face challenges with using Tawakkalna. Finally, this study has presented a list of recommendations to improve the usability of Tawakkalna. As a part of future work, a longitudinal study evaluating the usability of the application with respect to new and old users could be conducted. Additionally, evaluating other factors such as the learnability of Tawakkalna would be interesting.

Acknowledgment

Funding Information

Data Availability

The data used to support the finding of this study are available from the corresponding author upon request.

Ethics

This article is original and contains unpublished material. The authors declare no conflict of interest.

References

Al-Ajarma, K. (2021). After Hajj: Muslim pilgrims refashioning themselves. *Religions*, *12*(1), 36. https://doi.org/10.3390/rel12010036

Alanzi, T. M. (2022). Users' satisfaction levels about mHealth applications in post-Covid-19 times in Saudi Arabia. *Plos One*, *17*(5), e0267002. https://doi.org/10.1371/journal.pone.0267002

Alghannam, B. A., Albustan, S. A., Al-Hassan, A. A., & 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. https://doi.org/10.1080/10447318.2017.1388099

Algothami, S. S., & Saeed, S. (2021). Digital Transformation and Usability: User Acceptance of Tawakkalna Application during COVID-19 in Saudi Arabia. *Pandemic, Lockdown and Digital Transformation: Challenges and Opportunities for Public Administration, NGOs and Businesses*, 95-109. https://doi.org/10.1007/978-3-030-86274-9_6

Alharbi, R. A., Altayyari, F. T., Alamri, F. S., & Alharthi, S. A. (2021, October). Pandemic-driven technology during COVID-19: Experiences of older adults. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (pp. 5-9). https://doi.org/10.1145/3462204.3481769

Almufarij, A., & Alharbi, A. (2022). Perceptions of using mobile health apps (mHealth) during COVID-19 pandemic in Saudi Arabia: A cross-sectional study. *Journal of Health Informatics in Developing Countries*, *16*(1). https://www.jhidc.org/index.php/jhidc/article/view/352

Awan, M., Ali, S., Ali, M., Abrar, M. F., Ullah, H., & Khan, D. (2021). Usability barriers for elderly users in smartphone app usage: An analytical hierarchical process-based prioritization. *Scientific Programming*, *2021*, 1-14. https://doi.org/10.1155/2021/2780257

Berkman, M. I., & Karahoca, D. (2016). Re-Assessing the Usability Metric for User Experience (UMUX) Scale. *Journal of Usability Studies*, *11*(3).

Booday, M., & Albesher, A. (2021, December). Evaluating the Usability of Mobile Applications: The Case of COVID-19 Apps in Saudi Arabia. In *2021 22ⁿᵈ International Arab Conference on Information Technology (ACIT)* (pp. 1-7). IEEE. https://doi.org/10.1109/ACIT53391.2021.9677200

Borkowska, A., & Jach, K. (2017). Pre-testing of polish translation of System Usability Scale (SUS). In *Information Systems Architecture and Technology: Proceedings of 37ᵗʰ International Conference on Information Systems Architecture and Technology–ISAT 2016-Part I* (pp. 143-153). Springer International Publishing. https://doi.org/10.1007/978-3-319-46583-8_12

Borsci, S., Federici, S., & Lauriola, M. (2009). On the dimensionality of the System Usability Scale: A test of alternative measurement models. *Cognitive Processing*, *10*, 193-197. https://doi.org/10.1007/s10339-009-0268-9

Borsci, S., Malizia, A., Schmettow, M., Van Der Velde, F., Tariverdiyeva, G., Balaji, D., & Chamberlain, A. (2022). The Chatbot Usability Scale: The design and pilot of a usability scale for interaction with AI-based conversational agents. *Personal and Ubiquitous Computing*, *26*, 95-119. https://doi.org/10.1007/s00779-021-01582-9

Brooke, J. (1996). Usability Evaluation In Industry (B. W. Patrick W. Jordan, B. Thomas, Ian Lyall McClelland (Ed.); 1ˢᵗ ed., p. 6). CRC Press. https://doi.org/10.1201/9781498710411

Dianat, I., Ghanbari, Z., & AsghariJafarabadi, M. (2014). Psychometric properties of the persian language version of the system usability scale. *Health Promotion Perspectives*, *4*(1), 82. https://doi.org/10.5681/hpp.2014.011

Finstad, K. (2010). The usability metric for user experience. *Interacting with Computers*, *22*(5), 323-327. https://doi.org/10.1016/j.intcom.2010.04.004

Hvidt, J. C. S., Christensen, L. F., Sibbersen, C., Helweg-Jørgensen, S., Hansen, J. P., & Lichtenstein, M. B. (2020). Translation and validation of the system usability scale in a Danish mental health setting using digital technologies in treatment interventions. *International Journal of Human-Computer Interaction*, *36*(8), 709-716. https://doi.org/10.1080/10447318.2019.1680922

Karajizadeh, M., Zand, F., Vazin, A., Sharifian, R., & Bayati, N. (2022). Informatics and Technology in Clinical Care and Public Health. In *Informatics and Technology in Clinical Care and Public Health* (pp. 220-223). https://doi.org/10.3233/SHT1289

Kaya, A., Ozturk, R., & Altin Gumussoy, C. (2019). Usability measurement of mobile applications with system usability scale (SUS). In *Industrial Engineering in the Big Data Era: Selected Papers from the Global Joint Conference on Industrial Engineering and Its Application Areas, GJCIE 2018, June 21-22, 2018, Nevsehir, Turkey* (pp. 389-400). Springer International Publishing. https://doi.org/10.1007/978-3-030-03317-0_32

Kirakowski, J. (1996). The software usability measurement inventory: background and usage. *Usability evaluation in industry*, 169-178. https://doi.org/10.1201/9781498710411

Kirakowski, J., Claridge, N., & Whitehand, R. (1998, June). Human centered measures of success in web site design. In *Proceedings of the Fourth Conference on Human Factors & the Web*.

Kjeldskov, J., Skov, M. B., & Stage, J. (2005). Does time heal?: A longitudinal study of usability. In *Proceedings of the Australian Computer-Human Interaction Conference 2005 (OzCHI'05)*. Association for Computing Machinery.

Klug, B. (2017). An overview of the system usability scale in library website and system usability testing. *Weave: Journal of Library User Experience*, *1*(6). https://doi.org/10.3998/weave.12535642.0001.602

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, *30*(3), 607-610. https://doi.org/10.1177/001316447003000308

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, *7*(1), 57-78. https://doi.org/10.1080/10447319509526110

Lewis, J. R. (2018a). Measuring perceived usability: The CSUQ, SUS and UMUX. *International Journal of Human Computer Interaction*, *34*(12), 1148-1156. https://doi.org/10.1080/10447318.2017.1418805

Lewis, J. R. (2018b). The system usability scale: past, present and future. *International Journal of Human Computer Interaction*, *34*(7), 577-590. https://doi.org/10.1080/10447318.2018.1455307

Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. In *Human Centered Design: First International Conference, HCD 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009 Proceedings 1* (pp. 94-103). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02806-9_12

Lewis, J. R., Utesch, B. S., & Maher, D. E. (2013, April). UMUX-LITE: When there's no time for the SUS. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems* (pp. 2099-2102). https://doi.org/10.1145/2470654.2481287

Liu, N., Yin, J., Tan, S. S. L., Ngiam, K. Y., & Teo, H. H. (2021). Mobile health applications for older adults: A systematic review of interface and persuasive feature design. *Journal of the American Medical Informatics Association*, *28*(11), 2483-2501. https://doi.org/10.1093/jamia/ocab151

Martins, A. I., Rosa, A. F., Queirós, A., Silva, A., & Rocha, N. P. (2015). European Portuguese validation of the system usability scale (SUS). *Procedia Computer Science*, *67*, 293-300. https://doi.org/10.1016/j.procs.2015.09.273

McLellan, S., Muddimer, A., & Peres, S. C. (2012). The effect of experience on system usability scale ratings. *Journal of Usability Studies*, *7*(2), 56-67.

Mzoughi, M. N., & Garrouch, K. (2021). Does Social Risk Amplification Theory Explain the Resistance to COVID-19 Geo-Localization Applications?. *TEM Journal*, *10*(3), 1202. https://doi.org/10.18421/TEM103

Peres, S. C., Pham, T., & Phillips, R. (2013, September). Validation of the system usability scale (SUS) SUS in the wild. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 57, No. 1, pp. 192-196). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/1541931213571043

SPA. (2022). In a New Saudi Achievement, Tawakkalna Wins the United Nations Public Service Award 2022. *Saudi Press Agency*. https://www.spa.gov.sa/2364888

Tullis, T. S., & Stetson, J. N. (2004). A Comparison of Questionnaires for Assessing Website Usability ABSTRACT: Introduction. *Usability Professional Association Conference*, June 2006, 1-12.

Whig, P., Velu, A. R. U. N., & Sharma, P. (2022). Demystifying Federated Learning for Blockchain: A Case Study. In *Demystifying Federated Learning for Blockchain and Industrial Internet of Things* (pp. 143-165). IGI Global. https://doi.org/10.4018/978-1-6684-3733-9.ch006

WHO. (2023). World Health Organization. https://COVID-19.who.int/region/emro/country/sa