







Accessibility or Usability of the User Interfaces for Visually Impaired Users? A Comparative Study

Kamran Khowaja^(✉) , Dena Al-Thani , Aboubakr Aqle ,
and Bilikis Banire 

Division of Information and Computing, Technology,
College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar
kamran.khowaja@gmail.com, dalthani@hbku.edu.qa,
{aaqle, bbanire}@mail.hbku.edu.qa

Abstract. The use of the Internet to search and browse information has increased in the recent past from the younger children to adults and older adults with or without any disability. The access to the information is becoming more common through the use of mobile apps as it is ubiquitous. This requires an evaluation to ensure that it can be used by everyone including people with disability. However, there is no generic set of guidelines or heuristics that can be used by anyone to evaluate an interface in terms of usability and accessibility. This paper provides expert evaluation of two apps i.e. Accessible Qatar and LinkedIn using a set of heuristics by Gómez and colleagues, and Web Content Accessibility Guidelines (WCAG) 2.0 in terms of both usability and accessibility evaluation. The data analysis was carried out based on the usability problems and the opinion of experts using the system usability scale (SUS) questionnaire. The usability problems were analysed based on the number of usability problems found and the average severity ratings of the usability problems. The results show that the most frequently violated heuristics from the usability heuristics are *visibility* and *recognition*. The average severity rating of all the problems found using usability heuristics set is a mix of both minor and major. The results also show that the most frequently violated WCAG 2.0 guideline alternatives; the average severity rating of all the problems found using WCAG 2.0 guidelines is major.

Keywords: Visually impaired user · Heuristic evaluation · Accessibility guidelines · Web Content Accessibility Guidelines (WCAG) · Mobile heuristic · Interface evaluation

1 Introduction

There is a growing trend to make the desktop applications available for the mobile phones/web browsers so that they can be accessed anytime, and anywhere. It is because, the use of the internet for all types of users (children, adults, older adults) with/without disability/impairment has increased in recent years. The use of the website or any application (desktop/mobile phone) to browse user-specific information may

become difficult for visually impaired users among other users if the website/application has not followed the accessibility guidelines. The Web Accessibility Initiative (WAI) of World Wide Web Consortium (W3C) often referred to as W3C-WAI develop standards for accessibility (W3C Web Accessibility Initiative [W3C WAI] 2019). They also develop supporting materials for an individual to understand and implement accessibility. The web standards produced by W3C are known as “W3C Recommendations”. WAI has developed a number of recommendations, and web content accessibility guidelines (WCAG) is one of them. These guidelines are the most popular guidelines since 2000 when the first set of guidelines were made available. WCAG aims to address the accessibility of web pages and make web interaction available for people with disabilities (Web Content Accessibility Guidelines [WCAG] 2019).

Besides accessibility, for the developer working on an application, it is important to identify and fix the usability problems in the application before it is released for the potential users to use (Yeratziotis and Zaphiris 2017). The potential users are typically involved towards the last stage of the application development, therefore, the domain experts are usually recruited to evaluate the applications on behalf of the potential users. Heuristic evaluation (HE) is one of the many ways in which expert evaluation of an application can be performed (Nielsen and Molich 1990). The HE is quick, cheap and easy to perform over other methods of expert evaluation which include cognitive walkthrough (Polson et al. 1992; Wharton et al. 1994), goals, operators, methods and selection (GOMS) (Card et al. 1983), keystroke-level model (Card et al. 1980), or using results of previous study as a basis to prove or disprove different aspects of the design.

This led to an interesting question that if accessibility or usability at their own is enough? And can they replace each other? From the search, a number of research studies were found in which researcher have emphasised that accessibility and usability cannot substitute each other (Horton and Leventhal 2008; Hudson 2004; Shneiderman 2000); this shows that while following the accessibility guidelines like WCAG 2.0, one can ensure that an application (standalone or web-based) is accessible by everyone including the person with any disability/impairment. On the contrary, following the usability guidelines like Nielsen’s heuristics (Nielsen 1995; Nielsen and Molich 1990) or any other user interface guideline like eight golden rules of interface design by Shneiderman et al. (2009), one can ensure that usability problems have been fixed before anyone including the person with any disability/impairment starts using that application.

A study conducted by Correani et al. (2004) has highlighted that there are websites which fulfil the success criteria of accessibility guidelines but lack to meet the usability guidelines due to which users may face difficulty in terms of interaction with the websites.

To the best of our knowledge, there is no specialised set of heuristics or guidelines that can be used to evaluate interfaces for visually impaired users in terms of both the usability as well as accessibility.

In our previous research (Aqle et al. 2018), a search engine (SE) currently being developed as a part of our ongoing research for visually impaired users, was evaluated for both accessibility and usability. The SE has limited functionalities (no images and videos). The SE was evaluated by the experts using WCAG 2.0 and a set of heuristics by Nielsen (1995). Both of these sets have different purposes and cannot replace each

other but can complement each other. This was seen in the results as well. Although the results suggested that both sets contributed to the identification of usability issues that had to be fixed before an experimental evaluation with the visually impaired users was carried out. It was realised that one application is not enough to generalise the results on what to do with the evaluation of the systems. Therefore, this research builds on our existing research with an aim to present an expert evaluation of two mobile apps namely Accessible Qatar and LinkedIn. The mobile apps have been carefully chosen in a way that they may be useful for visually impaired users in their daily life. Nielsen's set is generic and not preferred for the evaluation of mobile apps, therefore, a specialised set of heuristic by Gómez et al. (2014) and WCAG 2.0 guidelines are used in this research to evaluate mobile apps in terms of both usability and accessibility.

2 Related Work

2.1 Web Content Accessibility Guidelines (WCAG)

The WCAG 1.0 was released back in 1999 and mainly provided accessibility for the static web pages. The W3C recommendation of WCAG 2.0 was published in December 2008 so that it can be applied on different web technologies and can be tested using a combination of automated tests through an application (web accessibility evaluation tool list¹) and manual evaluation by a human subject. The WCAG 2.1 is a recent W3C recommendation and was released in June 2018.

The foundation of WCAG 2.0 is based on four principles namely, perceivable, operable, understandable, and robust (Web Content Accessibility Guidelines [WCAG] 2019). Each principle has one or more guideline; there are 12 guidelines altogether. The researchers, developer among others should target each of these guidelines to make content accessible by people with different disabilities (Ellis and Mike 2011). It is to be noted that these guidelines at their own are not testable but each of these guidelines has one or more success criterion from a total of 61 success criterion. Each success criterion is individually testable in terms of the requirements and three levels of conformance: (1) A is the lowest, (2) AA, and (3) AAA is the highest (Caldwell et al. 2008). WCAG 2.1 includes 17 new success criterion which is part of 12 guidelines of WCAG 2.0.

Since WCAG 2.1 has been recently released and it may take time for companies to adopt new success criterion and ensure that their web content complies with the WCAG 2.1 guidelines. Therefore, WCAG 2.0 is used in this research. The conformance of WCAG 2.0 is widely used for web accessibility evaluation (Vigo et al. 2013). The use of the automated tool for the web accessibility evaluation is always subject to criticism as they can produce incorrect or misleading results (Abou-Zahra et al. 2017; Ivory and Chevalier 2002; Vigo et al. 2013). The automated tool can only assist than determining the accessibility. Therefore, human judgment is needed for the evaluation. The human subject referred to as an expert can manually check web content against all the guidelines as a part of the review process.

¹ <http://www.w3.org/WAI/ER/tools/>.

Following is a list of 12 guidelines from WCAG 2.0; these guidelines would be referred to as “accessibility guidelines” throughout the manuscript. One word in the complete name of each guideline is written inside a square bracket which represents the shorter names of the guidelines. These shorter names will be referred to in the subsequent sections.

1. Text [Alternatives]
2. [Time]-based Media
3. [Adaptable]
4. [Distinguishable]
5. [Keyboard] Accessible
6. [Enough] Time
7. Seizures and Physical Reactions [S and PR]
8. [Navigable]
9. [Readable]
10. [Predictable]
11. Input [Assistance]
12. [Compatible]

2.2 Set of Heuristics by Gómez et al. (2014)

Gómez et al. (2014) have compiled a set of heuristic evaluation checklists readapted for the mobile interface. The authors used existing heuristics from the desktop heuristic evaluation as a base, rearranged and expanded with the best practices and recommendations for a mobile interface which were missing in the existing heuristics. Following is a list of 13 heuristics which include 230 subheuristics; these heuristics would be referred to as “usability heuristics” throughout the manuscript. The first 10 heuristics are from desktop heuristics, while, the remaining 3 heuristics are taken from the mobile interfaces. From the subheuristics, 158 are based on the compilation of subheuristics from traditional general heuristic checklists and 72 are based on the compilation of mobile-specific subheuristics. One word in each complete name of the heuristic is written in a square bracket which represent the shorter names of the heuristics. These shorter names will be referred to in the subsequent sections.

1. [Visibility] of system status
2. [Match] between system and the real world
3. User [control] and freedom
4. [Consistency] and standards
5. [Error] prevention
6. [Recognition] rather than recall
7. [Flexibility] and efficiency of us
8. Aesthetic and [minimalist] design
9. Help users recognize, diagnose, and [recover] from errors
10. Help and [documentation]
11. [Skills]
12. Pleasurable and respectful [interaction] with the user
13. [Privacy]

2.3 Screen Readers

There are different screen readers available; some of them are free to use while the other needs to be purchased before it can be used.

Job Access with Speech (JAWS). JAWS is the world's most popular screen reader, developed for visually impaired users (Job Access With Speech [JAWS] 2019). Unfortunately, the trial version can only be installed and used on one personal computer (PC) for a week/month time. For the personal use, a visually impaired user can buy an annually 'home annual license' for \$90 or buy a perpetual version for \$1000 for personal/non-commercial purpose. Both paid versions of the licenses can be installed and used on three PCs only which means it cannot be used on the smartphones.

NonVisual Desktop Access (NVDA). NVDA allows visually impaired users and blind users to interact with the Windows operating system and many third-party applications for free (NonVisual Desktop Access [NVDA], 2019). As mentioned on their website, "We're free by principle, not by merit! We strive for a world where EVERYONE has equal access to the life-changing benefits of technology—not just the privilege." However, they do provide two options for those who are interested to willingly contribute to the better cause. The first option is 'one-off donation' of a fixed amount in Australian currency which include \$30, \$50, \$100, \$250 or other amount based on your preferences. The second option is 'monthly donation' of a fixed amount in Australian currency which include \$5, \$10, \$20, \$50 or other amount based on your preferences. Although, NVDA is free for users to use it can only be used on the PCs but not on the smartphones.

TalkBack Accessibility Service. TalkBack is developed by Google and it is being used as a screen reader on the smartphones. TalkBack service provides spoken feedback to the users so it allows them to operate its smartphone without having a look at the screen. The advantages of using TalkBack screen reader are: (1) it is pre-installed with Android operating system (OS), (2) it is free, and (3) it is easy to turn on/off TalkBack screen reader in device settings.

3 Study Design

3.1 Participants and Recruitment

The number of experts to be recruited should be at least three to identify 75% or more problems in the system to be evaluated (Nielsen and Molich 1990). Considering, this requirement and for the consistency in terms of the experts' profile, all experts who participated in our earlier research were invited from which two accepted our invitation. We used an approach of Snowballing to identify more participants and two more experts confirmed their participation. The recruited participants include researchers and faculty members at the university level and have used or use Nielsen's heuristics and WCAG guidelines for the interface design and evaluation (Table 1).

Table 1. Demographic information of the experts

Characteristic	Values
Gender	2 Female, 2 Male
Education	2 PhD, 2 Master's degree
Occupation	2 Employed, 2 PhD Student
HCI Courses	All
HCI experience	2 Expert, 2 Advanced
Screen readers	3 JAWS, 3 NVDA

3.2 Mobile Apps Used

Two Mobile Apps were to be selected for use by the experts to perform an evaluation; the criteria used to select apps is as follow:

1. The app is available on Apple App Store and Google Play Store
2. The app can be downloaded for free and used without any trial period
3. The app is of interest to visually impaired users (Hollier 2012)
4. Visually impaired users can use the app with minimal guidance

Based on the above-mentioned criteria, two apps are selected; the first app is Accessible Qatar. Accessible Qatar is a smartphone app and website where the disabled community is able to view the public and touristic locations and outlets in Qatar and see whether they are accessible and in what way. This research was carried out in Qatar and this app is particularly important in the local context as the number of people with disabilities (Qatari and non-Qatari) is increasing (Gulf Times, 2018).

The second app is LinkedIn mobile app which provides a faster way to tap into your professional world. It allows the user to get the latest news and updates related to their profession, provide a daily brief about users connected with you on the network, and an easy to get in touch with all of them. It may also be useful to establish connections with friends and companies to get updates regarding available jobs and their requirements (Hollier 2012).

3.3 Study Protocol

Following protocol was used as a part of this research:

1. An email was sent to the participants from our existing studies; they were informed that evaluation data submitted as a part of this research will remain anonymous. They were requested to respond to the email and inform about their willingness to be part of the evaluation process.
2. To recruit more participants in the replacement of declined request, an email was sent to new participants using snowball sampling.
3. The experiments were carried out face-to-face on campus. Each participant was given a briefing about the purpose of evaluation, introduction to both apps, and usability heuristics and accessibility guidelines to be used for the evaluation as a part of this research. They were informed about the identification of usability

problems in both the apps and the reporting process of those problems. For each usability problem they identify, they were asked to rate it from 0 to 4 (0 for 'not a problem', 1 for 'cosmetic', 2 for 'minor', 3 for 'major', and 4 for 'usability catastrophe').

4. They were asked to download, install and apps on their smartphones and enable 'TalkBack' accessibility feature on their smartphones to act as screen reader before using any of the mentioned apps. They were given two files; one to report usability problems based on the heuristics by Gomez, and another to report usability problems based on the WCAG guidelines. They were also given a softcopy of the System Usability Scale (SUS) questionnaire and were informed to fill-in once they have gone through both apps in detail and reported all the usability problems in the respective files. They were asked to carry out an evaluation at their own and submitted all three files through an email.
5. The SUS contains 10 statements to gather the opinion of an expert about the app they used. For each statement, expert need to select the best option from the scale of 1 referred to as 'strongly disagree' to 5 referred to as 'strongly agree'.

3.4 Data Analysis

Two data were gathered as a part of evaluation; the first is related to the usability problems in both apps, whereas the second one is related to their opinion using system usability scale (SUS) questionnaire. The data related to usability problems were analysed based on the following two parameters:

Number of Usability Problems Found. It is calculated as a sum of all the usability problems found by all the experts using each heuristic of Gómez and guidelines of WCAG.

Average Severity Ratings of the Usability Problems Found. It is calculated as an average of all the usability problems found by all the experts using each heuristic of Gómez and guidelines of WCAG.

System Usability Scale (SUS). The second analysis is carried out based on the SUS; it includes ten statements which are based on the Likert scale ('strongly disagree' as 1 to 'strongly agree' as 5) and provides an overall subjective assessment of any system. The usability measurements included in SUS cover the effectiveness, efficiency and user satisfaction. Following are the ten statements of SUS:

1. I think that I would like to use the app frequently.
2. I found the app unnecessarily complex.
3. I thought the app was easy to use.
4. I think that I would need the support of a technical person to be able to use the app.
5. I found the various functions in the app were well integrated.
6. I thought there was too much inconsistency in the app.
7. I would imagine that most people would learn to use the app very quickly.
8. I found the app very cumbersome to use.
9. I felt very confident using the app.
10. I needed to learn a lot of things before I could get going with the app.

4 Results

The results of evaluations based on usability heuristics and accessibility guidelines are discussed in the following sub-sections.

4.1 Overall Results

A total of 134 usability problems were found using usability heuristics and accessibility guidelines for Accessible Qatar and LinkedIn apps. The total number of usability problems found using usability heuristics for both apps is (N = 81, 60%), while the total number of usability problems found using accessibility guidelines for both apps is (N = 53, 40%).

Tables 2 and 3 shows the number of usability problems identified, its percentage within severity and percentage within usability heuristics (referred to as “GH” in the table) and accessibility heuristics (referred to as “WCAG”) for Accessible Qatar and LinkedIn app respectively.

Table 2. Comparison of usability problems identified in Accessible Qatar app using usability heuristics (GH) and accessibility guidelines (WCAG)

		Accessible Qatar		Total issues
		GH	WCAG	
Severity Rating (SR)	Count of SR 4	5	4	9
	% within Severity	55.56	44.44	100.00
	% within Set	9.43	11.43	10.23
	Count of SR 3	21	9	30
	% within Severity	70.00	30.00	100.00
	% within Set	39.62	25.71	34.09
	Count of SR 2	23	22	45
	% within Severity	51.11	48.89	100.00
	% within Set	43.40	62.86	51.14
	Count of SR 1	4	0	4
	% within Severity	100.00	0.00	100.00
	% within Set	7.55	0.00	4.55
Total	Total count	53	35	88
	% covered	60.23	39.77	100.00

Table 3. Comparison of usability problems identified in the LinkedIn app using usability heuristics (GH) and accessibility guidelines (WCAG)

		LinkedIn		Total issues
		GH	WCAG	
Severity Rating (SR)	Count of SR 4	4	0	4
	% within Severity	100.00	0.00	100.00
	% within Set	14.29	0.00	8.70
	Count of SR 3	17	16	33
	% within Severity	51.52	53.33	100.00
	% within Set	60.71	88.89	71.74
	Count of SR 2	7	2	9
	% within Severity	77.78	4.44	100.00
	% within Set	25.00	11.11	19.57
	Count of SR 1	0	0	0
	% within Severity	0.00	0.00	0.00
	% within Set	0.00	0.00	0.00
Total	Total count	28	18	46
	% covered	60.87	39.13	100.00

4.2 Usability Heuristics

Figures 1 and 2 show the usability problems identified and average severity ratings of the identified problems for each heuristic proposed by Gómez et al. (2014). The number of usability problems identified is represented using the vertical axis on the left side, while the average severity ratings of all the usability problems identified are represented using the vertical axis on the right side.

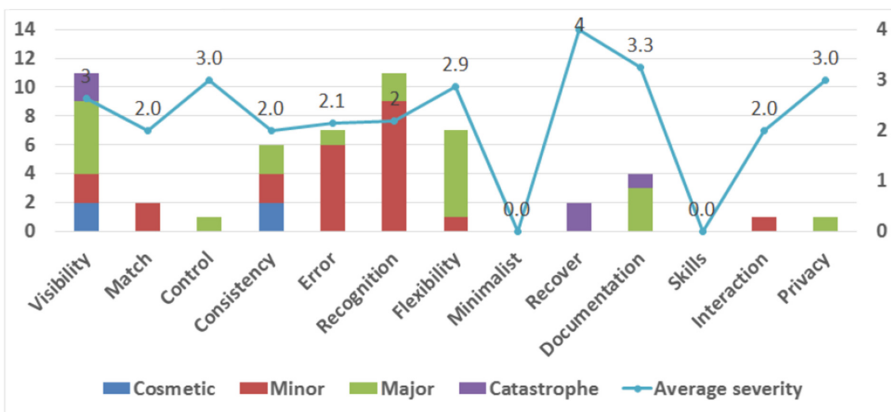


Fig. 1. Number of usability problems and average severity ratings found using usability heuristics in Accessible Qatar app

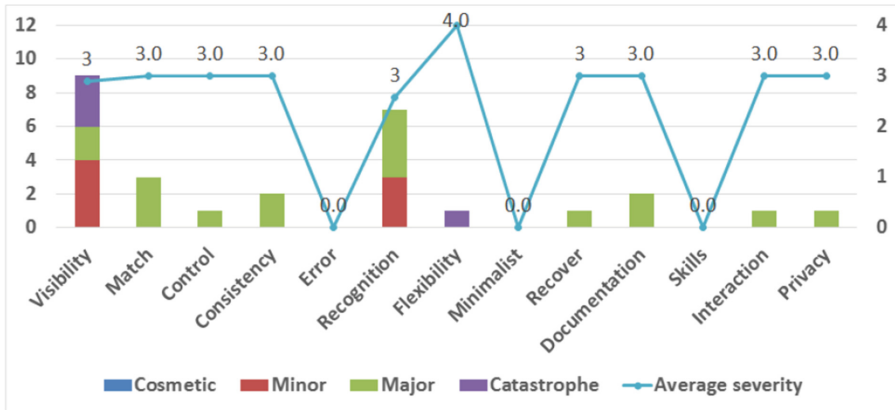


Fig. 2. Number of usability problems and average severity ratings found using usability heuristics in the LinkedIn app

Each heuristic in a usability heuristics is associated with a stacked column which represents the number of usability problems identified for one or more of the four severity ratings using different colours (blue for cosmetic, red for minor, green for major, and purple for catastrophe). The line connecting markers in a circle and running on a horizontal axis represents the average severity ratings of all the usability problems identified.

Number of Usability Problems Found. The most commonly broken heuristics for Accessible Qatar app are *visibility* and *recognition* (each has $N = 11$) followed by *error* and *flexibility* heuristics (each has $N = 7$) and *consistency* heuristic ($N = 6$). Some comments of the most commonly broken heuristics are presented in Table 4. The most commonly broken heuristics for the LinkedIn app are *visibility* heuristics ($N = 9$) followed by *recognition* heuristic ($N = 7$). Some comments of the most commonly broken heuristics are presented in Table 5.

Table 4. Examples of problems found in Accessible Qatar app using usability heuristics

Heuristic broken	Problem description	Severity rating
Visibility of system status	When using the search feature, the system doesn't give any feedback once the results are displayed. Hence, the screen reader user will not be able to know that the results are displayed on the screen	4
Recognition rather than recall	Under specific news detail screen, there is an icon to view photos related to the news. When clicks on the icon, it's read as "unlabelled button"	2
Error prevention	Under "Sign up" screen, the textbox of "Email" is read as "Editing, edit box"	2

(continued)

Table 4. (continued)

Heuristic broken	Problem description	Severity rating
Flexibility and efficiency of use	Under the favourites screen, there is no way for the user to select all the filters at one	3
Consistency and standards	There are two type filters in list view and map view. However, the naming can confuse the users as one is called 'filter' and the other is called 'category filter'	3

Table 5. Examples of problems found in the LinkedIn app using usability heuristics

Heuristic broken	Problem description	Severity rating
Visibility of system status	When clicking a profile photo there is no feedback	4
Visibility of system status	When message alert comes, the voiceover doesn't read it. Thus, the user will not be able to know that an alert message has popped up on the screen	4
Recognition rather than recall	Clicking on "Kudos username" button is read as "Quick comment button". It's confusing	2
Recognition rather than recall	Clicking on an image of those who viewed your profiles in a private, it is read as "group people image" which does not indicate anything about what it is	2

Average Severity Ratings of the Usability Problems Found. The visual analysis of the average severity ratings for Accessible Qatar app shows that the majority are problems are a combination of minor (N = 23, 43%) and major (N = 21, 40%) respectively. The visual analysis of the average severity ratings for the LinkedIn app shows that the majority are problems are major (N = 17, 61%).

4.3 Accessibility Guidelines

Figures 3 and 4 show the usability problems identified and average severity ratings of the identified problems for each accessibility guideline. For the consistency, the format of presenting the usability problems and their severity ratings are the same as of Figs. 1 and 2. Although each expert classified the identified usability problem into 1 of the 51 success criterion, space is limited to show all the information at once. Thus, all the related success criterion were grouped together and instead usability problems are shown using 12 guidelines.

Number of Usability Problems Found. The most commonly broken heuristics for Accessible Qatar app are assistance (N = 11) followed by *predictable* (N = 8) and *alternatives* (N = 7). Some comments of the most commonly broken heuristics are presented in Table 6. The most commonly broken heuristics for the LinkedIn app are *alternatives* (N = 4) followed by *time* and *navigable* (each has N = 3). Some comments of the most commonly broken guidelines are presented in Table 7.

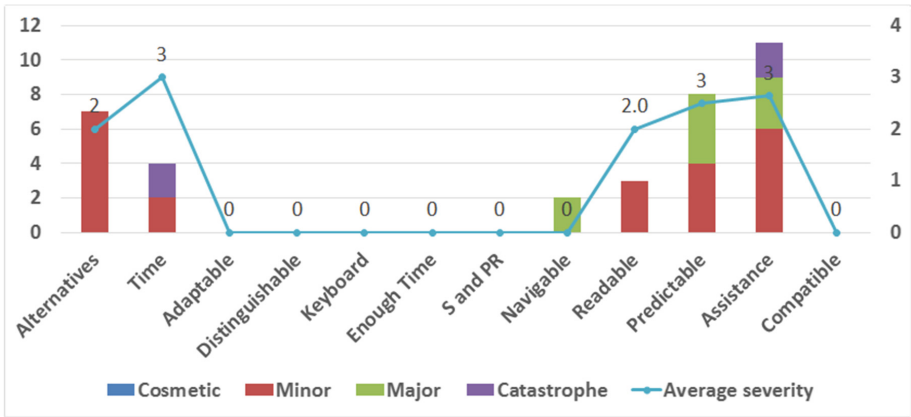


Fig. 3. Number of usability problems and average severity ratings found using WCAG 2.0 in Accessible Qatar app

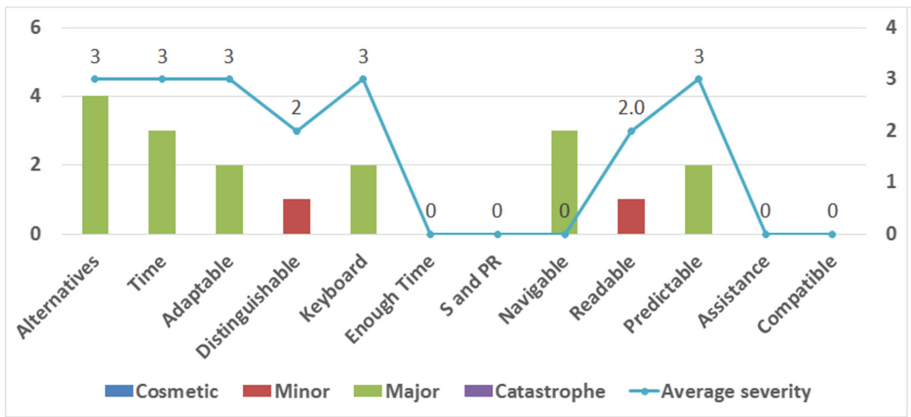


Fig. 4. Number of usability problems and average severity ratings found using WCAG 2.0 in the LinkedIn app

Table 6. Examples of problems found in Accessible Qatar app using WCAG 2.0 guidelines

Guideline broken	Problem description	Severity rating
Predictable	Some navigation buttons are not working when clicked. Such as ‘alphabetical and Distance links’ and this makes the user to be confused	3
Alternatives	When viewing the photos related to the news, there is no description of each photo which makes it difficult for the user to understand	3

Table 7. Examples of problems found in the LinkedIn app using WCAG 2.0 guidelines

Guideline broken	Problem description	Severity rating
Alternatives	There are no alternative texts for images, videos or diagrams added by users	3
Time-based Media	There are a number of shared audio clips which do not have sign language interpretation for those who are hearing-impaired	3
Navigable	There is no link description for the user to know the purpose of the link	3

Average Severity Ratings of the Usability Problems Found. The visual analysis of the average severity ratings for Accessible Qatar app shows that the majority are problems are minor (N = 22, 63%). The visual analysis of the average severity ratings for the LinkedIn app shows that the majority are problems are major (N = 16, 89%).

4.4 System Usability Scale (SUS)

In this paper, SUS was used for usability evaluation of the two mobile apps: Accessible Qatar and LinkedIn. The SUS guidelines by Brooke were used for the calculation base of the participants’ feedback to the survey questions (Brooke 1996).

The value of the SUS score for Accessible Qatar SUS score is distributed between 45% and 55% with an average value of 50%. While the value of the SUS score for LinkedIn is distributed between 50% and 77.5% with an average of 61%. The SUS scores for each participant and the mean scores for both apps are shown in Fig. 5.

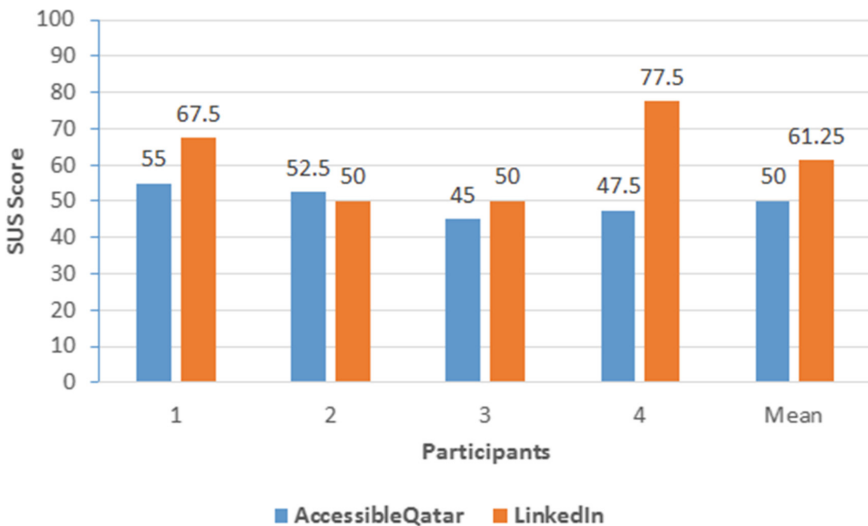


Fig. 5. Accessible Qatar and LinkedIn SUS Scores by participants

The average SUS scores for both mobile apps i.e. Accessible Qatar and LinkedIn are shown in Fig. 6. There are different ways to interpret the SUS scores (Sauro 2018); these include percentiles, grades, adjectives, acceptability, and promoters and detractors. In this research, SUS scores are interpreted using percentiles which also include the grade as well.

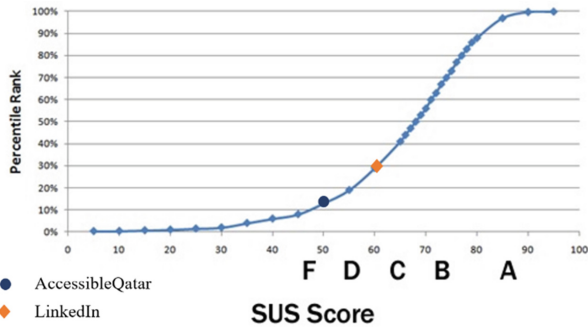


Fig. 6. Mean SUS Scores of Accessible Qatar and LinkedIn

The average SUS score of Accessible Qatar is 50, while the average SUS score of LinkedIn is 61.12. According to the grading scale interpretation of SUS scores by Sauro and Lewis (2016), Accessible Qatar and LinkedIn have “D” and “F” grade respectively. This shows that Accessible Qatar and LinkedIn mobile apps need major improvements and enhancements to meet the minimum usability scale requirements.

5 Conclusion

This research presented an expert evaluation of two apps i.e. Qatar Accessibility and LinkedIn using a set by heuristics by usability heuristics and accessibility guidelines. It was carried out as an extension of our previous study in which a web search interface (currently being developed as a part of our ongoing research) for visually impaired users was evaluated using a set of heuristics by Nielsen and accessibility guidelines. The key findings are as follows:

1. The analysis of results revealed that they were similar to our previous research (Aqle et al. 2018). The usability heuristics and accessibility guidelines both contributed to the identification of usability problems which needs to be fixed in the apps for better interaction experience of the user with both of the apps.
2. In comparison to Nielsen’s heuristics, the usability heuristics allowed identification of usability problems especially based on three new heuristics namely, Skills, Pleasurable and respectful interaction with the user, and Privacy. Although, these three heuristics contributed less but supported experts in finding major issues which had otherwise remained undetected.

3. The *visibility* and *recognition* usability heuristics were more commonly broken heuristics in both apps. The ignorance of these and other heuristics means the important usability problem would remain undetected.
4. The *alternatives* guideline was highly broken guideline in both apps.
5. There are a usability heuristics like “error prevention”, and “help users recognize, diagnose, and recover from errors” which are similar to accessibility guidelines like “error identification”, “error suggestion”, and “error prevention”. Such heuristics/guidelines revealed similar usability problems in both apps.
6. The mean SUS score shows that both apps need major improvements and enhancements to meet the minimum usability requirements.

While working on the design and development of an app, it is often possible to give more emphasis on either a set of usability heuristics or a set of accessibility guidelines such that other set may be overlooked. Based on the evaluation carried out as a part of our previous and this research, both usability heuristics and accessibility guidelines are equally important. The ignorance of either of them means critical usability and accessibility problems which should have been fixed would remain undetected.

WCAG is the standard guidelines to make the web accessible to people with disabilities when it comes to the implementation/evaluation of the accessibility of the content. However, the set of usability heuristics varies from the domain to domain; for instance, heuristics for deaf web user experience (Yeratziotis and Zaphiris 2017), heuristics for children with ASD (Khowaja et al. 2015), heuristics for child e-learning (Alsumait and Al-Osaimi 2009) among others. Thus, it is important for researchers to identify an appropriate set of heuristics to identify relevant usability problems in the user interface.

In future, the researchers can evaluate more apps especially in their local context. They can also use WCAG 2.1 guidelines for the evaluation of web content.

References

- Job Access With Speech [JAWS] (2019). <http://www.freedomscientific.com/products/software/jaws/>
- Abou-Zahra, S., Steenhout, N., Keen, L.: Selecting web accessibility evaluation tools (2017). <https://www.w3.org/WAI/test-evaluate/tools/selecting/>
- Alsumait, A., Al-Osaimi, A.: Usability heuristics evaluation for child e-learning applications. Paper presented at the Proceedings of the 11th International Conference on Information Integration and Web-Based Applications and Services (2009)
- Aqle, A., Khowaja, K., Al-Thani, D.: Accessibility or usability of InteractSE? A heuristic based approach to evaluate proposed search engine for the visually impaired users. Paper presented at the Eighth International Conference on Innovative Computing Technology (INTECH 2018), London, UK (2018)
- System Usability Scale (SUS): a quick-and-dirty method of system evaluation user information, 19172650 (1996)
- Caldwell, B., Reid, L.G., Vanderheiden, G.: Web Content Accessibility Guidelines (WCAG) 2.0 (2008)

- Card, S.K., Moran, T.P., Newell, A.: The keystroke-level model for user performance time with interactive systems. *Commun. ACM* **23**(7), 396–410 (1980)
- Card, S.K., Newell, A., Moran, T.P.: The psychology of human–computer interaction (1983). <https://dl.acm.org/citation.cfm?id=578027>
- Correani, F., Leporini, B., Paternò, F.: Supporting Web Usability for Vision Impaired Users. Springer, Berlin (2004). https://doi.org/10.1007/978-3-540-30111-0_20
- Ellis, K., Mike, K.: Disability and New Media. Routledge, New York (2011). https://doi.org/10.1007/978-1-84800-050-6_21
- Gómez, R.Y., Caballero, D.C., Sevillano, J.-L.: Heuristic evaluation on mobile interfaces: a new checklist. *Sci. World J.* **2014**, 434326 (2014). <https://doi.org/10.1155/2014/434326>
- Hollier, S.: SociAbility: Social Media for People with a Disability. Media Access Australia, Ultimo (2012)
- Horton, S., Leventhal, L.: Universal usability. In: Harper, S., Yesilada, Y. (eds.) *Web Accessibility*, pp. 346–355. Springer, London (2008)
- Hudson, W.: Inclusive design. *Interactions* **11**(4), 55–56 (2004). <https://doi.org/10.1145/1005261.1005278>
- Ivory, M.Y., Chevalier, A.: A Study of Automated Web Site Evaluation Tools. Department of Computer Science. University of Washington, Washington, DC (2002)
- Khowaja, K., Salim, S.S., Asemi, A.: Heuristics to evaluate interactive systems for children with Autism Spectrum Disorder (ASD). *PLoS ONE* **10**(7), e0132187 (2015). <https://doi.org/10.1371/journal.pone.0132187>
- Nielsen, J.: Ten usability heuristics (1995). <http://www.nngroup.com/articles/ten-usability-heuristics/>
- Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People (1990)
- NoName (2018). Qatar’s efforts to help people with disabilities highlighted. Retrieved from <https://www.gulf-times.com/story/591963/Qatar-s-efforts-to-help-people-with-disabilities-h>
- NonVisual Desktop Access [NVDA], NVDA (2019). <https://www.nvaccess.org/about-nvda/>
- Polson, P.G., Lewis, C., Rieman, J., Wharton, C.: Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *Int. J. Man Mach. Stud.* **36**(5), 741–773 (1992)
- Sauro, J.: MeasuringU: 5 ways to interpret a SUS Score (2018). <https://measuringu.com/interpret-sus-score/>
- Sauro, J., Lewis, J.R.: Chapter 8: Standardized usability questionnaires. In: Sauro, J., Lewis, J.R. (eds.) *Quantifying the User Experience*, 2nd edn, pp. 185–248. Morgan Kaufmann, Boston (2016)
- Shneiderman, B.: Universal usability. *Commun. ACM* **43**(5), 84–91 (2000). <https://doi.org/10.1145/332833.332843>
- Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S.: *Designing the User Interface: Strategies for Effective Human–Computer Interaction*. Addison-Wesley Publishing Company, Reading (2009)
- Vigo, M., Brown, J., Conway, V.: Benchmarking web accessibility evaluation tools. Paper presented at the Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility—W4A 2013, Rio de Janeiro, Brazil (2013)
- W3C Web Accessibility Initiative [W3C WAI]: W3C-WAI (2019). <https://www.w3.org/WAI/>
- Web Content Accessibility Guidelines [WCAG]: WCAG (2019). <https://www.w3.org/WAI/standards-guidelines/wcag/>
- Wharton, C., Rieman, J., Lewis, C., Polson, P.: The cognitive walkthrough: a practitioner’s guide. In: *Usability Inspection Methods*. Wiley, New York (1994)
- Yeratziotis, A., Zaphiris, P.: A heuristic evaluation for deaf web user experience (HE4DWUX). *Int. J. Hum. Comput. Interact.* **34**(3), 195–217 (2017). <https://doi.org/10.1080/10447318.2017.1339940>