Perceived Usability of Arabic System Usability Scale (A-SUS): Faculty Members' Perspective of Smart PAAET Application

Waheeda Almayyan, PhD
The Public Authority for Applied Education and Training, Kuwait

Manal Alsuwaidi, BSc
The Public Authority for Applied Education and Training, Kuwait

Bareeq A. AlGhannam, PhD
The Public Authority for Applied Education and Training, Kuwait

ABSTRACT
The recent past has seen increased usability of computer and mobile phones applications (apps) for education and applied skills services. Nevertheless, the usability of such apps, particularly in languages such as Arabic, remains a concern of theoretical understanding and practical implementation. This study applies the Arabic System Usability Scale (A-SUS) survey in Kuwait to understand the faculty' perspective of utilizing the Smart PAAET, that is an app utilized by the Public Authority for Applied Education and Training (PAAET) with various faculty members of Basic Education, Business Studies, Technological Studies, and Administrative Services. The A-SUS results suggest that the i) Faculties of Business and Technology have relatively higher usability ii) particularly among the faculty of age between 30-39, and iii) having a master qualification. The data is further reduced into two composite components of usability and learnability to explore the current theoretical critique on the factoring structure of SUS. These findings enhance the data theory of SUS in general and A-SUS in particular into an essential area of learning management applications from the faculty perspective.

Keywords
Arabic, Applications, Usability, Scale, Education, and Training.

1. INTRODUCTION
In the last few decades, the development of Information and Communication Technologies (ICT) has introduced many essential benefits in various areas [1,2], including higher education and its systems. Higher educational institutions have benefited from ICT, particularly in the Covid-19 pandemic, where they transformed teaching through ICT [3,4]. The integration between ICT and education is further enhanced with smartphone advancements and associating applications (apps). Smartphones and applications are the recent shapes of the systems that are growing fast. The 2020 statistics show 2.7 billion smartphone users and 1.35 billion tablet users [5].

In the ICT and systems research, the System Usability Scale (SUS) is mostly used to scale the relative usability of the systems and associated apps [6]. The standard version of the SUS has ten items, usually negatively and positively worded, usually with 5-point Likert scales [7]. Literature shows that the SUS has been applied to various systems such as Cognitive Behavioral Therapy System, Health Care System, Automated Teller Machines, Insulin Pens, Social Media Sites, Smartphones, Mobile Application for Housing, Email & Word Processors [10,11,12,13,14,15,16].

The use of SUS is vital for continuous evaluation of systems and the scale itself, to compare results with the previous studies to ensure that system quality in terms of usability and thus sustain or remain use [17]. The ICT-based learning management system is a growing area, and evaluating the systems' usability from various perspectives such as students [18] and faculty is essential. Such systems and allied applications for educational institutions are advancing the sustainability of the overall academic management and thus can cause satisfaction of concerned stakeholders if the perceived usability is constantly checked and compared. This usability, therefore, requires urgent research [17] from various perspectives.

This paper is a part of a series of usability sets conducted on various applications applied in an academic environment of higher education. This paper is explicitly piloted on 54 Using the A-SUS, which is applied to Smart PAAET, an application utilized by the Public Authority for Applied Education and Training (PAAET). Such research enhances the overall knowledge of A-SUS in terms of its factors segregation as the published research has consistently failed to replicate that factor structure of SUS [7].

2. LITERATURE REVIEW
Usability Definition
The (International Organization for Standardization, 2018) p5 defines usability as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." Since usability is understood and practice contextually, it is not easy to come up with same absolute sense and measures for its items. Therefore, both usability and measures require operational definitions from their context [19].

2.1 Standard Usability Tools
The term usability, which bases the SUS, has been entirely contextual since its emergence and is therefore controversial in both its science and the practice of experimental psychology, measurement, and statistics [8]. For scoring usability, the origin of standardized usability questionnaires appeared in the late 19th century [9,21,22] App usability can also be tested through psychometric tests, including usability inspection methods like heuristic evaluation and user testing [23]. Heuristic evaluation is based upon the expert's opinion who inspects the application and identifies usability problems from potential users' points of view. With user testing, a sample user is selected and asked to use the app and report things they think did not work or are not appropriate. On the other hand, SUS, on the other hand, being a psychometric measurement tool, uses the Likert scale to collect the user's opinions to be used as basis for scoring the usability and

provide essential insights for the scale's factor structure.

2.2 SUS Specifically

The SUS primarily works as an opinion-based questionnaire that has been introduced by Brooke [9]. It is a very popular standardized questionnaire for assessing the perceived usability of systems and applications in various fields [7]. The instrument is coined as very reliable, quick, easy to use, and low cost, thus made it ideal for large-scale global surveys in which the companies are usually interested. SUS has ten questions, positively and negatively worded, with either 5 or 7 points Likert scale, and it scores an application from 0-100. Recent attempts seem to segregate the ten items of SUS into a more composite structure for two elements, namely i) usability and ii) learnability [7]. Item 4 and 10 are kept into learnability, as they are claimed to be more focused upon scoring learning than usability. At the same time, the rest of the eight items are kept in usability.

2.3 Translated SUS into Arabic Language

Recent research has attempted to translate the standard SUS into many languages such as Arabic, Chines, French, German, Hindi, and Spanish following systematic translation and validation procedures [6]. These continue attempts to translate the usability questionnaires and surveys, primarily aim to remain contextually relevant and approach the usability community with greater confidence as they are feeling natives when using a translated SUS in their native language [20] and also to capture the contextual details and rhetoric [6] thus leading to words greater compliance. It is because verification of factor structure of such a translated SUS, as aimed in this study, can bring compliance of Adapted SUS with SUS to build confidence and eventually standardize them, from which the A-SUS investigation rises.

However, usability evaluation of A-SUS for the Arabic region is slight and lacks sufficient confidence [20] as usability is contextual and cannot be measured with exact numbers [6]. Therefore, the faculty context in the Arabic region can bring interesting dimensions, particularly the verification of a composite structure (usability and learnability) [7] and how the A-SUS scores vary among faculty with different ages, subjects, and qualifications in the region. The following section outlines a methodology for achieving these objectives.

3. METHODOLOGY

To help advance the Arabic region in usability studies, a standard usability tool is administered on the Smart PAAET application. This study presents faculty members' perspectives of the smart PAAET application as an academic management app. It is administered using the Arabic System Usability Scale (A-SUS) tool, an Arabic adaptation of the standard System Usability Scale. The "Smart PAAET" serves educational management system, where it is specifically developed for users of the Public Authority for Applied Education and Training (PAAET) in Kuwait. The research conducts psychometric evaluation on the perceived usability data of Smart PAAET. The following section outlines the methodology of this research.

SUS studies use the survey method to collect and analyze data. The study has adopted an A-SUS scale that contains ten questions as outlined in[7]. Studies can be found to use a 5-point scale starting with a strong degree to agree strongly. A minimum score of 68 on a scale of 0-100 signifies that usability is above average [24,25]. Since the study aims to build a faculty perspective, 54 active faculty members were selected who used PAAET to fill the A-SUS. Ideally, a sample size greater than 200 participants is considered robust and Sufficient [26]. However, studies can be found with samples lower than 200 and generated great arguments, such as [27] have used 50 responses. Test for Sampling Adequacy, Kaiser-Meyer-Olkin (KMO), is also run. For interpreting KMO values between 0.8 and 1, a rule of thumb indicates the sampling is adequate [28]. In this study, the KMO value is returned as 0.855, which shows the data is adequate for the component analysis[29].

The statistical analysis included descriptive statistics, Cronbach's alpha, and principal component analysis (PCA). While doing the PCA, 10 SUS items are factorized into two main components (PC1=Usability, all items except 4 and 10, PC2=Learnability (Q4 and Q10) [7]. Cronbach's alpha is used to test whether the instrument has internal consistency and is reliable or not. Whereas the descriptive statistics were used to see how the scores vary by the faculty's age, subject, and qualification as it was aimed to find out. The principal component analysis enabled the research to verify the factor structuring. The principal component analysis enables the study to transform and verify all ten items of A-SUS into two composite constructs of usability and learnability, as factor structure of SUS [7,15].

4. RESULTS AND INTERPRETATIONS

4.1 Single-Test Reliability Analysis

Table 1: Frequentist Scale Reliability Statistics

Estimate	Cronbach's α	mean	sd
Point estimate	0.912	3.544	0.205
95% CI lower bound		0.407	
95% CI upper bound		0.629	

Table 1 shows Cronbach's α=0.912, which is much higher than the minimum required standard value of 0.70 alpha coefficient. As a rule of thumb, Cronbach's α>.90 represents "excellent" internal consistency (Glen., 2020). This means that A-SUS items have a higher internal consistency, and thus, the instrument is highly reliable.

Table 2: Frequentist Individual Item Reliability Statistics

	If item dropped			
Item	Cronbach's α	Item-rest correlation	mean	sd
Q1	0.910	0.579	3.574	1.075
Q2	0.897	0.804	3.537	0.862
Q3	0.896	0.812	3.796	0.939
Q4	0.909	0.582	3.796	1.016
Q5	0.905	0.654	3.389	1.036
Q6	0.901	0.719	3.296	0.983
Q7	0.897	0.779	3.685	0.987
Q8	0.895	0.821	3.667	0.932
Q9	0.900	0.748	3.519	0.947
Q10	0.922	0.379	3.185	1.029

Note. The following items were reverse scaled: Q2, Q4, Q6, Q8, Q10.

Table 2 shows item-wise co-efficient α. Among the ten items of A-SUS, six items (Q1,Q4, Q5,Q6,Q9,Q10) have Cronbach's α greater than 0.90, whereas the rest of the 4 items

(Q2,Q3,Q7,Q8) Cronbach's α a little lower than 0.90. Thus, all items are individually reliable, too, as they scored higher than the standard co-efficient value of 0.70. This also means that six items fall under the category of "excellent" and four items fall in the category of "Good" internal consistency [28].

4.2 Demographics

Table 3: Demographics of the study

	[ALL] N=54	N
College:		54
Basic Education College	26 (48.1%)	
College of Business Studies	9 (16.7%)	
College of Technological Studies	4 (7.41%)	
Higher Institute for Administrative Services	15 (27.8%)	
Education:		54
Bachelor	22 (40.7%)	
Master	7 (13.0%)	
Ph.D	25 (46.3%)	
Age:		54
30-40	13 (24.1%)	
41-59	38 (70.4%)	
60 and above	3 (5.56%)	

Table 3 demonstrates inclusive demographics of the respondents being from 4 colleges, with three different academic levels and three different age categories. Firstly, from a college point of view, 48.1% (n-26) participants were from Basic Education College, 27.8%(n-15) were from Higher Institute for Administrative Services, 16.7% (n=9) were from College of Business Studies, and 7.41% (n=4) were from College of Technological Studies. Secondly, from an education level's point of view, 46.3%(n=25) of the respondents were Ph.D. qualified, followed by 40.7% (n=22) have had bachelor's degrees, and the remaining 13.00% (n=7) were qualified as Masters. Thirdly, from an age point of view, most of the participants, 70% (n=38), were of age between 41-59 years, whereas 24.1% (n=13) were of age between 30-40 years, and only 5.56% (n=3) of the respondents were of age above 60 years.

Descriptions of SUS Items

Table 4: Descriptive Statistics for SUS items

Item	Definition	Mean	SD	SEM
Q1	I think I like to use this app constantly.	3.57	1.075	0.146
Q2	I found this app too complicated	2.46	0.862	0.117
Q3	I think this app is very easy to use.	3.8	0.939	0.128
Q4	I think I need help from a technical person to use this app.	2.2	1.016	0.138
Q5	I found the multiple functions in this application consistent with each other.	3.39	1.036	0.141
Q6	I thought there was a lot of conflict in using this app.	2.7	0.983	0.134
Q7	I imagine a lot of faculty / training will learn to use this app easily.	3.69	0.987	0.134
Q8	I found this app strange to use.	2.33	0.932	0.127
Q9	I felt totally confident using this app.	3.52	0.947	0.129
Q10	You must know many things to facilitate the use of this application.	2.81	1.029	0.14
SUS		**63.611**	**18.363**	**2.498**

Table 4 shows that Item 3, which states " I think this app is straightforward to use", has the highest mean of 3.8 on the scale of 5, which indicates that "Smart PAAET" is mainly liked for its easiness, followed by Item 7 (3.69) which states "I imagine a lot of faculty/training will learn to use this app easily" and then Item 1 with score 3.57 stating "I think I like to use this app constantly". The overall mean score is, however, 63.611. This score is, though, lower than the SUS average benchmark score of 68[24] and also lower than the benchmark of software products, which is 72 [20]. However, further segregation of the scores from various parameters perspectives brings exciting variations, as reported in the next section.

Table 5: College-wise Scores of A-SUS

	Basic Education College N=26	College of Business Studies N=9	College of Technological Studies N=4	Higher Institute for Administrative Services N=15	p-value	N
Q1	3.31 (1.05)	4.33 (1.00)	4.00 (0.00)	3.47 (1.13)	0.072	54
Q2	2.69 (0.84)	2.11 (0.93)	2.25 (0.50)	2.33 (0.90)	0.275	54
Q3	3.46 (0.95)	4.22 (1.09)	4.25 (0.50)	4.00 (0.76)	0.074	54
Q4	2.62 (0.94)	2.00 (1.22)	1.75 (0.50)	1.73 (0.88)	0.030	54
Q5	3.12 (0.91)	3.89 (1.17)	3.50 (1.00)	3.53 (1.13)	0.239	54
Q6	3.08 (0.69)	1.89 (1.27)	2.50 (1.00)	2.60 (0.99)	0.012	54
Q7	3.27 (0.96)	4.22 (1.30)	3.75 (0.50)	4.07 (0.59)	0.018	54
Q8	2.65 (0.80)	2.00 (1.22)	2.50 (1.00)	1.93 (0.80)	0.063	54
Q9	3.27 (0.87)	3.89 (1.27)	3.75 (1.26)	3.67 (0.72)	0.290	54
Q10	3.19 (0.90)	2.44 (1.24)	1.75 (0.50)	2.67 (0.98)	0.021	54
SUS	**55.5 (13.5)**	**75.3 (26.9)**	**71.2 (14.4)**	**68.7 (15.5)**	**.011**	**54**

Table 5 demonstrates that the 10 A-SUS items are crossed with four colleges to generate 40 observations through the matrix table 5. In this respect, the overall average score of the College of Business Studies is the highest (75.3), followed by the College of Technological Studies (71.2). Both these colleges demonstrate that there is above average (68) usability of the "Smart PAAET". Higher Institute for Administrative

Services (68.7) has average usability equal to the SUS benchmark (68). College of Basic Education, however, has shown the lowest usability (55.5).

4.3 Qualification-wise scores of A-SUS

Table 6: Usability by Faculty Qualification

	Bachelor N=22	Master N=7	PhD N=25	p-value	N
Q1	3.55 (0.96)	4.00 (1.00)	3.48 (1.19)	0.529	54
Q2	2.50 (0.86)	1.86 (0.69)	2.60 (0.87)	0.127	54
Q3	3.82 (0.80)	4.43 (0.79)	3.60 (1.04)	0.117	54
Q4	2.18 (0.96)	1.71 (0.76)	2.36 (1.11)	0.335	54
Q5	3.41 (1.01)	3.86 (1.21)	3.24 (1.01)	0.383	54
Q6	2.86 (0.94)	2.29 (0.95)	2.68 (1.03)	0.402	54
Q7	3.82 (0.73)	4.14 (0.69)	3.44 (1.19)	0.180	54
Q8	2.27 (0.88)	1.57 (0.53)	2.60 (0.96)	0.030	54
Q9	3.45 (0.86)	4.29 (0.49)	3.36 (1.04)	0.065	54
Q10	2.86 (1.17)	2.86 (0.90)	2.76 (0.97)	0.938	54
SUS	**63.4 (17.6)**	**76.1 (13.3)**	**60.3 (19.3)**	**0.132**	**54**

Table 6 shows that faculty members with a qualification of Master have demonstrated the highest usability score (76.1), which is above the benchmark of SUS (68) as well as the benchmark of SUS for software (72) [20]. The faculty with qualifications of Bachelor (63.4) and Ph.D (60.3) have shown the usability a little lower than the said standards.

4.4 Age-wise scores of A-SUS

Table 7: A-SUS scores by Age

	30-40 N=13	41-59 N=38	60 and above N=3	p-value	N
Q1	4.00 (0.58)	3.45 (1.18)	3.33 (1.15)	0.261	54
Q2	2.08 (0.49)	2.55 (0.92)	3.00 (1.00)	0.123	54
Q3	4.23 (0.44)	3.71 (1.01)	3.00 (1.00)	0.070	54
Q4	2.08 (0.64)	2.18 (1.06)	3.00 (1.73)	0.364	54
Q5	3.69 (0.85)	3.32 (1.04)	3.00 (1.73)	0.429	54
Q6	2.31 (0.75)	2.79 (1.02)	3.33 (1.15)	0.164	54
Q7	3.92 (0.76)	3.66 (0.99)	3.00 (1.73)	0.334	54
Q8	1.92 (0.64)	2.39 (0.95)	3.33 (1.15)	0.043	54
Q9	3.69 (0.85)	3.50 (0.92)	3.00 (1.73)	0.517	54
Q10	2.77 (1.01)	2.82 (1.01)	3.00 (1.73)	0.943	54
SUS	**71.0 (10.9)**	**62.2 (18.6)**	**49.2 (33.6)**	**0.125**	**54**

Table 7 shows that faculty members with ages between 30-40 demonstrated the highest usability score (71.0), which is above the benchmark of standard SUS (68) and a little lower than the benchmark of SUS for software (72) [20]. The faculty aged between 41-59 (62.2) and of age 60 and above (49.2) have shown the usability lower than the said standards.

4.5 Principal Component Analysis

Table 8: Chi-squared Test

	Value	df	p
Model	59.521	26	< .001

Kaiser-Meyer-Olkin Measure of Sampling Adequacy = 0.855.

Table 8 shows the principal component analysis. The results show that the 10 A-SUS items can be factorized into two main components (Principal Component 1, PCI=Usability, all items except 4 and 10, Principal Component, PC2=Learnability, item 4 and 10). In the comparative analysis of PC1 and PC2, the results show enough evidence, as the $p<0.05$, to conclude that PC1 and PC2 are distinctive components. Kaiser-Meyer-Olkin (KMO) test is used, which is a measure of how suited the data is. KMO returns values between 0 and 1. For interpreting KMO values between 0.8 and 1, a rule of thumb indicates the sampling is adequate [25]. In this study, the KMO value is 0.855, which shows that the data was adequate for the component analysis.

Table 9: Component Loadings

	PC1	PC2	Uniqueness
Q1	0.860		0.254
Q2	0.785		0.246
Q3	0.886		0.160
Q4		0.867	0.206
Q5	0.839		0.292
Q6	0.464	0.699	0.296
Q7	0.776		0.270
Q8	0.639	0.605	0.225
Q9	0.710		0.347
Q10		0.873	0.235

Note. Applied rotation method is varimax.

In principal component analysis, component loadings are the correlation coefficients between the variables (Q1-Q2) and components (Usability & Learnability). The component loading is the percent of the variance in that variable explained by the component. PC1 (usability) explains the highest variance in Q3, which is 0.886, and lowest in Q6, 0.464. In comparison, PC2 (learnability) explains the highest variance in Q4, followed by Q10. These results confirm that previous studies claim and interpret the learnability to be used as a separate construct with usability [7].

Table 10: Component Characteristics

	Eigenvalue	Proportion var.	Cumulative
PC1	5.825	0.583	0.583

Table 10: Component Characteristics

	Eigenvalue	Proportion var.	Cumulative
PC2	1.642	0.164	0.747

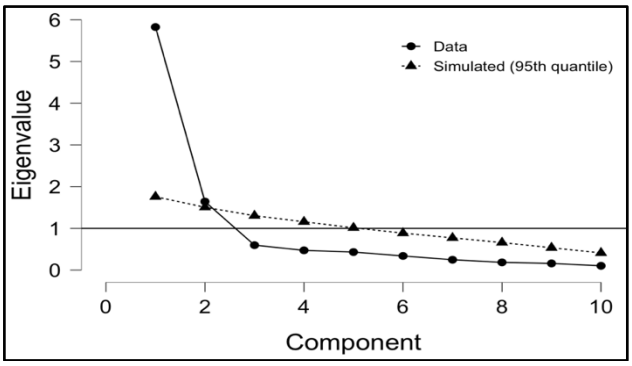

Fig 1: Scree plot

A scree plot (Table 10 & Figure 1) shows the eigenvalue of each of the ten questions in A-SUS. The eigenvalue, which is greater than one, is rotated through the varimax rotation for interpretation. The consequent varimax factors which are greater than 0.75 are considered as "strong", the values range from 0.50-0.75 is considered as "moderate", and the values range from 0.30-0.49 is considered as "weak" factor loadings. Both PC1 (usability) fall in the category of moderate loading with an eigenvalue of 0.583, whereas loading of PC2 (learnability) eigenvalue is merely 0.164, thus falls below weak factor loading.

5. DISCUSSION

In this study, the A-SUS instrument is applied to a mobile application, Smart PAAET, to score the system's usability by the faculty members. The overall usability score is below the average usability standard; however, further analysis indicated that the usability score is higher than standards for faculty from Colleges of Business and Technology, particularly among the young faculty aged between 30-40 and with Master's qualification. This finding indicates that the usability is contextual to the user's characteristics as they create a context for the application [30] and subjective usability [31]. This finding also confirms that usability cannot be standardized for all types of users and therefore needs significant attention to detail [6]. This study found that this contextuality particularly emerges from the age of the users, their qualifications, and the nature of the subject they teach.

The contextuality of usability increases the dimensionality of the A-SUS scores. Therefore, more composite constructs are required to have a more halitotic understanding of the system usability [32]. Principle component analysis has signified how the usability and learnability, as two principal components explain variations in various items of the A-SUS scores. Similar to the recent discussions in literature [7], Q4 and Q10 have shown higher loading to learnability and thus supporting the argument of having learnability as a unique and independent construct to be studied and applied into the systems' studies, specifically to the novice users of Smart PAAET. This complies with the literature as researchers state that as users get more experienced with the system, the SUS score becomes a single scale of perceived usability.

6. CONCLUSION AND FUTURE RESEARCH

The comprehensive research on SUS has a recent history of it being a quick, cheap, and, more importantly, reliable tool for measuring the usability of systems at a large scale. However, the translated A-SUS needs continued evaluation into various fields and stakeholders to capture the essential contextuality that the Arabic language and culture bring in. This study brings in the faculty perspective of A-SUS. It brings in an exciting variety of scores bases upon the nature of the subject, age, and qualification of the faculty and how their scores composite back into the usability and learnability as standalone constructs among the faculty. These results will bring more theoretical power to the A-SUS to capture more in-depth details of variations in usability scores, mainly with teaching and learning applications such as the Smart PAAET. The results are significant for usability practitioners, such as Public Authority for Applied Education and Training in Kuwait and other such authorities to cautiously implement education and training applications with particular care for customization for the teachers' age, the subjects they teach, and the qualifications possess.

For the future, this study provides multiple directions to enhance the overall theoretical and practical debates of usability. For example, it would be interesting to study how much divergence or convergence of usability score happens among various strata of the users. This can be done through rigorous cross-sectional research designs. It could also enhance usability by approaching usability as a spiral issue of time-space, where the usability of apps, such as the Smart PAAET, may lose or gain its scores among the same or similar group of users over multiple points in time and/or location. A time-wise score of lapses and gains can provide essential insights related to usability and thus the sustainability of the applications and systems.

7. REFERENCES

[1] Nuere, S., & de Miguel, L. (2020). The Digital/Technological Connection with COVID-19: An Unprecedented Challenge in University Teaching. Technology, Knowledge and Learning, 1–13.

[2] Blair, I. (2020). Mobile App Download and Usage Statistics (2020). https://buildfire.com/app-statistics/

[3] Gao, M., Kortum, P., & Oswald, F. L. (2020). Multi-Language Toolkit for the System Usability Scale. International Journal of Human-Computer Interaction, 00(00), 1–19. https://doi.org/10.1080/10447318.2020.1801173

[4] Lewis, J. R., &Sauro, J. (2017). Revisiting the Factor Structure of the System Usability Scale. Journal of Usability Studies, 12(4), 183–192.

[5] Lewis, J. R. (2014). Usability: Lessons Learned. and Yet to Be Learned. International Journal of Human-Computer Interaction, 30(9), 663–684. https://doi.org/10.1080/10447318.2014.930311

[6] Brooke, J. (1996). SUS: A "quick and dirty'usability. In Usability evaluation in industry (p. 189). CRC press.

[7] De Angeli, A., Athavankar, U., Joshi, A., Coventry, L., & Johnson, G. I. (2004). Introducing ATMs in India: a contextual inquiry. Interacting with Computers, 16(1), 29–44.

[8] Frandsen-Thorlacius, O., Hornbæk, K., Hertzum, M., &Clemmensen, T. (2009). Non-universal usability? A survey of how usability is understood by Chinese and Danish users. Proceedings of the SIGCHI Conference on

Human Factors in Computing Systems, 41–50.

[9] Hsieh, H. C. L. (2011). Exploring the impact of cultures on web usability test. International Conference on Human Centered Design, 47–54.

[10] Jonsson, O., Haak, M., Tomsone, S., Iwarsson, S., Schmidt, S. M., Mårtensson, K., Svensson, T., &Slaug, B. (2016). Cross-National Usability Study of a Housing Accessibility App: Findings From the European InnovAge Project. Journal of Usability Studies, 12(1).

[11] Li, Z. (2014). Critical evaluation of cross-cultural impact on usability studies: analysis and case study data collection. California State University, Northridge.

[12] Mol, M., Van Schaik, A., Dozeman, E., Ruwaard, J., Vis, C., Ebert, D. D., Etzelmueller, A., Mathiasen, K., Moles, B., Mora, T., Pedersen, C. D., Skjøth, M. M., Pensado, L. P., Piera-Jimenez, J., Gokcay, D., Ince, B. Ü., Russi, A., Sacco, Y., Zanalda, E., … Smit, J. H. (2020). Dimensionality of the system usability scale among professionals using internet-based interventions for depression: A confirmatory factor analysis. BMC Psychiatry, 20(1), 1–10. https://doi.org/10.1186/s12888-020-02627-8

[13] Walsh, T., Nurkka, P., & Walsh, R. (2010). Cultural differences in smartphone user experience evaluation. Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia, 1–9.

[14] AlGhannam, B. A., Alsuwaidi, M., & Almayyan, W. (2018). Perceived Usability Using Arabic System Usability Scale (A-SUS): Student Perspective of Smart PAAET App. International Journal of Computer Science and InformationSecurity, 16(7), 66–78.

[15] Binyamin, S., Rutter, M., & Smith, S. (2016). the Utilization of System Usability Scale in Learning Management Systems: a Case Study of Jeddah Community College. ICERI2016 Proceedings, 1, 5314–5323. https://doi.org/10.21125/iceri.2016.2290

[16] Brooke, J. (2018). System usability scale (SUS). Iron and Steel Technology. https://doi.org/10.5948/upo9781614440260.011

[17] 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

[18] Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 213–218.

[19] Lewis, J. R., Utesch, B. S., & Maher, D. E. (2015). Measuring Perceived Usability: The SUS, UMUX-LITE, and AltUsability. International Journal of Human-Computer Interaction, 31(8), 496–505. https://doi.org/10.1080/10447318.2015.1064654

[20] Brajnik, G. (2000). Automatic web usability evaluation: what needs to be done. Proc. Human Factors and the Web, 6th Conference. http://users.dimi.uniud.it/~giorgio.brajnik/papers/hfweb00.html

[21] Al Ghurair, N., &Alnaqi, G. (2020). Schemes of eStories for Children with Social Communication Difficulties. Advances in Human-Computer Interaction, 2020, 1–9. https://doi.org/10.1155/2020/9530218

[22] Finstad, K. (2006). The system usability scale and non-native English speakers. Journal of Usability Studies, 1(4), 185–188.

[23] Mol, M., Van Schaik, A., Dozeman, E., Ruwaard, J., Vis, C., Ebert, D. D., Etzelmueller, A., Mathiasen, K., Moles, B., Mora, T., Pedersen, C. D., Skjøth, M. M., Pensado, L. P., Piera-Jimenez, J., Gokcay, D., Ince, B. Ü., Russi, A., Sacco, Y., Zanalda, E., … Smit, J. H. (2020). Dimensionality of the system usability scale among professionals using internet-based interventions for depression: A confirmatory factor analysis. BMC Psychiatry, 20(1), 1–10. https://doi.org/10.1186/s12888-020-02627-8

[24] Binyamin, S., Rutter, M., & Smith, S. (2016). the Utilization of System Usability Scale in Learning Management Systems: a Case Study of Jeddah Community College. ICERI2016 Proceedings, 1, 5314–5323. https://doi.org/10.21125/iceri.2016.2290

[25] Glen., S. (2016). Kaiser-Meyer-Olkin (KMO) Test for Sampling Adequacy. https://www.statisticshowto.com/kaiser-meyer-olkin/

[26] Glen., S. (2020). Cronbach's Alpha_ Simple Definition, Use and Interpretation. https://www.statisticshowto.com/contact/

[27] Putnam, C., Puthenmadom, M., Cuerdo, M. A., Wang, W., & Paul, N. (2020). Adaptation of the System Usability Scale for User Testing with Children. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, 1–7.

[28] Kortum, P., & Peres, S. C. (2014). The relationship between system effectiveness and subjective usability scores using the System Usability Scale. International Journal of Human-Computer Interaction, 30(7), 575–584.

[29] Drew, M. R., Falcone, B., &Baccus, W. L. (2018). What does the system usability scale (SUS) measure? International Conference of Design, User Experience, and Usability, 356–366.

[30] Putnam, C., Puthenmadom, M., Cuerdo, M. A., Wang, W., & Paul, N. (2020). Adaptation of the System Usability Scale for User Testing with Children. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, 1–7.

[31] Kortum, P., & Peres, S. C. (2014). The relationship between system effectiveness and subjective usability scores using the System Usability Scale. International Journal of Human-Computer Interaction, 30(7), 575–584.

[32] Drew, M. R., Falcone, B., &Baccus, W. L. (2018). What does the system usability scale (SUS) measure? International Conference of Design, User Experience, and Usability, 356–366.