Screen size effect on usability of Arabic forms for smartphones

Rihabb Salamin
Deanship of Graduate Studies and Scientific Research
Palestine Polytechnic University
Hebron, Palestine
rihabbs@ppu.edu

Dr. Ghassan Shahin
Department of Administrative Science
Palestine Polytechnic University
Hebron, Palestine
gshahin@ppu.edu

***Abstract*— Researchers have established a concrete and solid ground for web forms design and their usability testing since the nineties. However, since the dramatic shift towards devices that carries a varied set of properties, such as, smartphones and tablets, developers started to build forms to suit these devices. As a consequence, a radical shift in research and usability testing needed to strike as well. We attempt to work with Arabic forms that are designed particularly for mobile devices (smartphones to be exact). Our goal, was to come up with empirically tested usability guidelines for Arabic mobile forms through achieving a set of objectives. Firstly, we have investigated the effect of smartphones screen size on error rate. In other words, do smaller screens increase the number of errors made by the users during form filling tasks? Secondly, we have addressed the following question: what is the best location for an error message to appear on smartphones' forms? The above issues as come to our knowledge never been addressed before for Arabic smartphones' forms. We have reached that conclusion after a detailed literature review. Our results have revealed the existence of a negative relationship between screen size and number of errors, users are less likely to make errors when the screen is big. As to the best location to display an error message, we tested a variety of locations. Our findings indicated that for a smartphone that has a small screen, below text box is the best location for a message to be shown. Pop up messages comes next but it scored the worst regarding subjective preferences.**

Keywords—usability; usability testing; forms; error message presentation; error message design

I. INTRODUCTION

People consider form filling tedious and boring. A form that is designed poorly will definitely worsen the situation. Moreover, form filling in that case generates frustration. Frustrated users in turn means higher web sites' turnover. On the other hand, a usable web form implies user satisfaction and better responsiveness. Researchers and practitioners broached form usability testing from different angles. For example, they have examined label location, inline help, error messages and others [4]. However, since the emergence of hand held devices, the research trends needed to be shifted as will. Researchers started to direct their efforts towards forms designed for mobile devices (such as smartphones and tablets).

Mobile devices are associated by a set of properties that might hamper users' ability to complete their online tasks that include form filling [6]. The diverse set of functionalities associated with mobile devices keeps in altering each and every day. One of the most common and popular properties associated with palm sized devices is the ability of accessing web on the go, due to their relative small screen real estate and other features [3]. Those features have made forms that are designed for desktops not suitable to be used on mobile devices. Researchers need a do over regarding addressing forms usability testing. In this paper, we tested error messages location. In addition, we have studied the effect of smartphones screen size on form filling error rate. We wanted to check if the small screen size has an effect on the number of errors made by the users during the process of form filling.

II. THEORITICAL BACKGROUND

Researchers covered aspects regarding screen size and forms usability. In [5] authors addressed the following question: Do small screens affect web usability? The results of their experiments have shown that those users who have used the smaller screen were 50% less effective in their performance than those who used larger screens. They have recommended that a direct research feature should be provided, and reduce scrolling as much as possible. Authors of [8] have inspected the possible impact of screen size on learning that is based on videos. The results of their research highlighted two interesting points. First, that users no matter the size of the screen used to view the video for learning have revealed a positive attitude towards the process. Second, despite the users positive attitude there is some sort of a tradeoff when it comes to the effectiveness of the learning process. The smaller the device screen the less effective the process. Authors of [1] orientated towards three main activities: navigation, searching, and form filling. The participants in the conducted experiments in this work were asked to interact with the same application, but on three different devices. The previous devices have different screen sizes. The results proved that larger screens lead to higher efficiency, usability, and effectiveness. In [9] The

researchers wanted to ensure if the screen- size of small devices affect the way search results are displayed, and if yes how to provide an appropriate summery for each result. Their investigation has approved that the screen size affect the summery of the viewed results. The presented summery on those devices has to be limited, specific, and representative as much as possible. They also suggested the use of contextual mechanisms to filter the retrieved search results.

Authors of [2]][7][10][11] evaluated error messages location. Authors of [7] looked in depth to six different locations: right, left, below, top, and bottom. Their results indicated that the best location for an error message in a web form is to be placed to the right of the erroneous field. In [11] the usability of web forms was tested in terms of form completion time, error recognition time, the number of saccades, and error correction time. They have tested four different locations: to the right of input field, above it, top of the form, and the bottom of the form. They have used an eye tracking device to analyze users' interaction with the forms through eye movement. The eye tracking test has shown users were able to notice error messages faster when it was displayed on the right side of input field, which in turn implies less effort required by the user to understand the given message. Authors of [10] have investigated the appearance of error messages at the top of the page after completing the entire form, below the input field after completing the entire form, and below the input field after completing it. The results showed that the third option was the worse among them. Users missed the message with a total of 40%. On the other hand, in the above two options error messages were missed less often by the users. Efforts in [2] are concentrated basically on the timing of errors. However, they have investigated several locations of error messages. Their results show that error messages might be embedded inside a form or to have the shape of a dialogue and dialogue boxes must be used to present error messages one by one and not all at once. All of the previous studies have investigated error messages in respect to web forms. However, research regarding error messages for mobile forms have been given very little forethought. Moreover, researchers only investigated error messages for English forms. So, we have carried our own research and experiments to investigate error messages regarding Arabic forms for smartphones.

III. METHODOLOGY

A. Experimental Design

We have designed 2 main forms. The requested information are simple and easy to remember, the user doesn't need to take time to remember the answer, the answer already in his head. So, no time is required for thinking which might increase the time needed to complete the form. All of the involved fields are required with validation. Obstacles used in experiment 2 are summarized in table I.

A hundred student participated in experiment one. Both genders were recruited, 50% females and 50% males. The average age was 21 years and SD was 1.84; range: 18 – 27. In experiment two 48 students were selected. The average age was 22 years (SD = 1.56; range: 18 – 26). The analysis of demographic characteristics (age, gender, educational level, internet skills and others) of recruited students revealed no significant differences between the experimental groups, therefore effects due to the demographic characteristics were excluded.

B. Experimental Procedure

Testing was carried out in Palestine Polytechnic University in March 2017. Each of the involved subjects was introduced to the device that is used in the testing. The experimental settings and purposes was explained to them as well before they start their assigned task.

In experiment one where the effect of screen size on error rate is questioned, each student is required to fill a questionnaire that contains some demographical aspects, such as age, gender, students' educational level and others. After that the participant was asked to fill the designed form. Two devices were involved, 50 students were assigned the form completion process using the small device, the other 50 were supposed to use the big device to fill the form with the required data. Experiment one was targeting only the measurement of the number of errors that the user has made during form filling, in order to check if the number of errors differs among the two devices or not. The user fills the form and after he finishes, he hits the submit button, and that's the end of it. Then the observer counts the number of errors made by the user.

In experiment 2, the user is required to correct the errors. If the user enters wrong data in a specific field, he will be informed by an error message. We have tested four different locations: above the label (AL), below the label (BL), below the input field (BT), and pop up (PU). The participants were divided into groups, each group will test a form that contains the same data, but the error message is placed in a different location than the other forms assigned to other groups. Each group assigned 12 members. After submitting the form successfully, the user was asked to fill a second questionnaire that is related to the location error messages. In addition, we have utilized retrospective think aloud (RTA) as a mean to gain more insightful information from users about their experience with the designed forms. The process was videotaped and the observant latter analyzed each video for extracting the required information, such as, number of errors made by the user, type of the error, number of consecutive errors, time required to correct the errors and the time required to complete the form successfully. Authors and Affiliations

C. Apparatus

Android studio 2.3 was used to design the forms involved in the testing process. Fig 1 represent one of the developed forms. Two Samsung devices were involved in the testing process. Samsung Galaxy Star 2 plus of a screen size 480 x 800 pixels, this device is the one that we refer it as the small device in this paper. Samsung Galaxy A7 which we refer it as the big device in this paper, of a display size 1080 x 1920 pixels. The used operating systems, android KitKat 4.4.2 and marshmallow 6.0.1 respectively. For analyzing the collected data we have used SPSS 24.0 and alpha for all the tests was set to 0.05.

IV. RESULTS

A. *Efficiency*

Efficiency in experiment two was measured by the time needed to correct the errors and time needed to complete the forms. The results for those two variables are shown in table II and table III.

Participants who were involved in testing the location of error messages (AL, BL, BT, and PU) using the small screen have corrected the errors in average of: 15.61, 14.99, 4.01, and 6.96 seconds respectively. The longest time to correct the errors was associated with error messages presented above the label. The shortest time was associated with error messages displayed below the text box. Differences between group means were diffidently significant as it was shown by one-way ANOVA test, $F(3, 44) = 6.080$, $p = 0.001$. As a result of this significance between the tested groups, we have continued the analysis using LSD to determine the group that gave the best result. LSD post hoc analysis revealed that some of the groups vary significantly than others, in other words there is a statistically significant differences between some of the groups as the following: AL and BT where $p = 0.001$, AL and PU where $p = 0.012$, BL and BT where $p = 0.002$, BL and PU where $p = 0.019$. However, we have observed that LSD post-hoc analysis revealed no significance whatsoever when it comes to AL and BL where $p = 0.852$, as well as BT and PU where $p = 0.392$.

TABLE I. OBSTACLES FOR EXPERIMENT 2

Field	Arabic Field	Obstacle	Example
First name	الاسم الأول	- In English. - Lower case.	ahmad.
DOB	تاريخ الميلاد	D/M/Y	5/12/1996.
Phone number	رقم الهاتف	Jawwal number only.	0591278655
Email	البريد الالكتروني	Hotmail or gmail.	Ah_ty@gmail.com
Password	كلمة المرور	- Starts with %. - Followed by any 4 letters. - Ends with 2 digits.	%uhyg12
Confirm password	تأكيد كلمة المرور	- Matches the password.	%uhyg12

Fig.1 Error message displayed above label.

TABLE II. TIME TAKEN TO CORRECT ALL ERRORS

Method	M	SD
AL	1.3008	.84583
BL	1.2492	.80747
BT	.3417	.37868
PU	.5800	.55829

TABLE III. TIME COMPLETE THE FORM SUCCESSFULLY

Method	M	SD
AL	3.0492	1.10840
BL	2.4900	.87483
BT	2.8133	.82876
PU	2.3333	1.26664

The users who filled the form with the error messages displayed above the label (AL) completed the form with a total time 36.59 seconds, which is the longest time among the groups. The shortest time was delivered by those who have completed the form (PU) which is 28 seconds. However, one-way ANOVA test have revealed no significant differences between the tested groups regarding form completion time, $F(3, 44) = 1.158$, $p = 0.337$.

B. *Effectiveness*

Effectiveness in experiment one was measured by the number of errors made by the users. In experiment two, effectiveness was measured by the number of errors and number of consecutive errors. The results for those two variables are shown in table IV, table V and table VI.

As it has been determined by a one-way ANOVA that in experiment 1 there was a statistically significant difference between the two groups (small and big screen) regarding error rate. $F(1, 98) = 6.671$, $p = 0.011$. The analysis showed that users who filled the form using the small screen device have made much more errors than those who used the big screen. The results proved that devices with small screens can lead to a

higher error rate in form filling. Most of the participants even expressed their thoughts regarding the matter, especially those who used the small device, they have voiced their opinions after finishing the task by saying: "the device is small, it's frustrating and derived me to make more errors during filling due its screen size".

A one-way ANOVA have confirmed that the error rate in experiment 2 for the small device where the location of error messages was tested, is not statistically different for the tested, is not statistically significant between the experimented groups.

TABLE IV. ERROR RATE IN EXPERIMENT 1

Screen Size	M	SD
Big	2.82	1.870
Small	3.76	1.768

TABLE V. ERROR RATE IN EXPERIMENT 2

Method	M	SD
AL	2.00	1.044
BL	2.08	.900
BT	1.33	.651
PU	2.08	.669

(AL, BL, BT, PU). F(3, 44) = 2.284, and p = 0.092. In other words, the number of the errors made by the users don't vary among the groups using the small device.

TABLE VI. CONSECUTIVE ERROR RATE

Method	M	SD
AL	1.42	1.379
BL	1.17	1.403
BT	.58	1.240
PU	.33	.651

A one-way ANOVA test was carried to determine the significance between tested groups. The number of consecutive errors is not statistically significant between the experimented groups for the small device. F(3, 44) = 2.076, p =0.117. Therefore, we concluded that the different locations of error messages using the small screen don't affect the consecutive error rate.

C. Satisfaction

Satisfaction was measured by the subjective ratings provided by the users. Results of subjective ratings are provided by table VII.

Regarding subjective preferences and satisfaction where the location of error messages and users have tested with the small screen, the users were asked to answer the three questions that is implemented for example in table 32. For question 1, "EL was not annoying", as come to those who tested the forms with errors being displayed above the label (AL), 2 (6.25%) answered with strongly agree, 6 (50%) answered with agree, 1 (6.25%) answered with disagree and 3 (25% with strongly disagree.

For BL group, 2 (16.67%) have answered with strongly agree, 8 (66.67%) with agree, 1 (6.25%) with disagree, and 1 (6.25%) with strongly disagree. BT group, 3 (25%) have answered with strongly agree, 8 (66.67%) with agree, 1 (8.33%) with disagree, and none answered with strongly disagree. And finally, for PU participants, the results was as for BT group. A one-way ANOVA has confirmed that there is no significant differences between the tested groups (AL, BL, BT, and PU), F(3, 44) = 1.485, p = 0.232.

The second question: "Error message was helpful", 14 (29.67%) answered strongly agree, 26 (54.17%) agree, 6

TABLE VII. SUBJECTIVE RATINGS IN EXPERIMENT 2

Questions	AL M SD	BL M SD	BT M SD	PU M SD
Error location isn't disturbing	2.58 1.08	2.92 .793	3.17 .577	3.17 .58
Message was helpful	3.25 .754	2.83 .937	3.00 .426	3.25 .866
Message was easy to correct	3.08 .669	3.08 .669	3.00 .603	3.58 .515

(12.5%) disagree, and 2 (4.17%) strongly disagree, which suggests that the involved users found the error messages helpful. However, ANOVA test has shown no significant differences between the tested groups, F(3, 44) = 0.841, p = 0.479. The third question: "the error is easy to correct", 15 (31.25) answered with strongly agree, 27 (56.25%) have answered with agree, 6 (12.5%) have answered with disagree and none of the participants have answered with strongly disagree. The previous numbers indicate that most of the participants found that the error easy to correct. A one-way ANOVA has shown that there is no significant differences among the groups, F(5, 42) = 1. 211, p = 0.321. As to the preferred location according to the participants who have tested small device: 11 (22.91%) reported that they preferred the message to be placed above the label (AL), 15 (31.25%) answered they consider BL as the best location, 13 (27%) voted for BT location, and 9 (18.75%) have chosen PU as the best presentation.

V. DISCUSSION

Smartphones come with different capabilities and sizes. We wanted to examine if there is a relationship between the size of the screen and the number of errors the user make during form filling. The results has proven the existence of such relationship. The number of errors made by the users of the big screen is much less than the number of errors made by the users of the small screen. The feedback that we got from the participants supported our findings. The users' feedback included statements regarding screen size and its impact on the

number of errors made by them. The majority of the small screen participants expressed their frustration that resulted from having to fill a form with a small device. The small real estate of the device required more focus and users' attention to read the instructions provided, as well as the presented error. In addition, the keyboards small buttons generated more typos.

The detailed literature review that was carried for this paper has demonstrated that forms error messages can be placed in different locations. Popular locations include above the label, below text boxes, placeholders, above the forms, dialogue boxes and others. We have chosen to evaluate a set of them through empirical testing. We have picked the following locations: above the label, below label, pop up, and below text boxes. We have evaluated the locations of error messages regarding efficiency, effectiveness, and satisfaction. The findings the location of error messages with respect to the completion time indicated that when error messages appeared to the users as a pop up message has achieved the best results Our guessing is that the reason refer to the fact that error messages of BL, BT, and AL groups appeared to the users all at once while PU was shown one by one and since the message stay in front of the users, they probably were kept staring at the message. However, that was actually against the assumptions we have made at the beginning of the study. We have expected PU to achieve the worst results since the error message disappear in PU group, and it is displayed far away from the fields, plus if the users want to reread the message in PU groups he need to hit the submit button again in order to see the message which implies some additional time to the filling process. Anyway, the analysis of the collected data shown that the differences between the means of tested groups are not statistically significant. Which in turn implies that the time to complete the forms is close for all of the tested groups. Our findings matches to some level the findings of [10] regarding error messages in web forms. However, our testing of the pop up messages involved the presentation of error messages after clicking the submit button, while their work involved immediate feedback.

The results regarding the time taken to correct all errors made by each group indicated the following: the analysis have shown that there is statistically significant differences between the tested groups, which contradict the findings of [2] that declared that error correction time is not statistically significant between the different error messages location for web forms. BT has scored the shortest time among the tested groups.

Effectiveness was measured by error rate and consecutive error rate. As for error rate which is the first time errors the analysis have shown no significant differences among the tested groups. The group with the fewest number of errors is BT. Consecutive error rates don't t vary greatly between the groups as it was suggested by the analysis. The group that scored the fewest number of consecutive errors is PU.

Satisfaction was measured by subjective rating and feedback given by the participants. The small device users voted for BL as the best location to present the error messages. While PU was chosen by the users as their least favorite. As we discussed before, PU involves presenting the error message as a pop up box that appears for some time then disappears quickly which requires to show the message again if the user wants recheck the content of the error message. In addition, the message appear at the end of the screen, which is quite the distance between the first few boxes and the error messages. While BL shows the message between the label and its related erroneous input field. As it was sated by the users, this way led them to determine the location of the text box that needs correction quickly.

In general, most of the tested dependent variables yielded close results for all tested groups. We attribute that to the fact that error messages were short, few number of fields were involved, and error messages were displayed close to the input

fields.

VI. Conclusion

The proliferation of mobile devices has led the cyber age to a new era. This breakthrough altered research efforts and directed them to new routes. Practitioner now need to come up with new design trends to accommodate the diversity associated with these devices. Mobile devices such as tablets, smartphones, and PDAs properties are different. Our work concentrates on smartphones.

Experiment one findings demonstrated that the small screen generated more errors than the big screen. This signifies an important issue that must be considered during the design process. Designers should start considering to involve shorter forms for screens with small real estate and to provide assistance as much as possible for users during form filling. Experiment two answered the question, regarding the best location. Small device: We have found that BT achieved the best results regarding error rate and time to correct errors, while PU achieved superior results regarding the number of consecutive error rate. We consider BT as the best among the groups, since PU has close results to BT.

We found no evidence to prove the effect of the culture on error messages representation. We attribute that to fact that we only involved text boxes and we have allowed the use of both languages (English and Arabic) to enter the data. We received zero complaints regarding the used languages from the users. They have shifted back and forth between the languages whenever it was required. We provide the following set of mobile forms design guidelines based on the findings of this dissertation:

- Avoid long forms with crowded components, especially for small screens.
- Small devices best error location is below the text field.
- Presenting error messages with a pop up massages is recommended when the context of the error message is not very long. When the input involves many restrictions as it is required for the password field, PU are not favorable. Since the user will more likely to make multiple mistakes while filling such a box.
- Use PU to present error messages one by one. PU are not recommended for displaying errors all at once.

- Displaying error messages above the label is not recommended.

VII. LIMITATIONS

There are quite number of limitations that have to be addressed. First, there are several number of possible locations where the error message might appear, we have addressed only a portion of them. We have neglected for example, testing error messages where they appear beside the label or inside the input field (also known as placeholders). Second, our tests were targeting computer major students. Students that surly have some experience regarding computer, internet, forms, and their components. In our tests we have neglected novice users. This segment of users might yield different types of results. Third, our samples were selected from the same university that is located at Palestine which implies that the groups of people we have involved in our tests belong to the same culture. As a consequence, the findings of this study can't be applied to other cultures. However, our investigations were directed towards Arabic forms in the first place, so the applicability to other cultures doesn't matter at this point. Fourth, the tested form in experiment two were short, only few input fields were involved. Long forms might lead to different sorts of findings. Fifth, we have involved text input boxes, other type of input boxes include calendars, or radio buttons might affect the results. Finally, the users provided their ratings of the best location without truly having the experience with the different locations of error messages. So, may be if they tried them all will provide a better comparative process, and perhaps they change their mind about their answers regarding their ratings of location error messages.

VIII. FUTURE WORK

Our future work will include overcoming the limitations imposed on this work and we are planning to test the timing of error messages. In addition, we are planning to examine the effect of graphics and whether they lead to achieve higher affordance (error message noticed faster). We look forward to deploy an eye tracking device. These devices gain the researchers a better insight and knowledge regarding the human behavior, like what is first thing that captures the human eye, did the participant notice the error message immediately or not.

REFERENCES

[1] D. Raptis, N. Tselios, J. Kjeldskov, and M. Skov, "Does size matter? Investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency," Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services. ACM, 2013.

[2] J. Bargas-Avila, G. Oberholzerb, P. Schmutza, M. de Vitoa, and K. Opwis, "Usable error message presentation in the World Wide Web: Do not show errors right away," Elsevier, vol. 19. 2007, pp. 330–341.

[3] J. Bargas-Avila, O. Brenzikofer, S. Roth, A. Tuch, S. Orsini, and K. Opwis, "Simple but crucial user interface in the world wide web: Introducing 20 guidelines for usable web form design," 2010. Rita Matrai (Ed.), ISBN: 978-953-307-084-1.

[4] J. Bargas-Avila, S. Orsini, H. Piosczyk, D. Urwyler, and K. Opwis, "Enhancing online forms: Use format specifications for fields with format restrictions to help respondants," Elsevier, 2010.

[5] J. Harms, M. Kratky, C. Wimmer, K. Kappel, and T. Grechenig, "Navigation in long forms on smartphones: Scrolling worse than tabs, menus, and collapsible fieldsets," Proceedings of the 15th IFIP TC 13 International Conference Human-Computer Interaction, vol. 9296. Springer, 2015.

[6] M. Al-Saleh, A. Al-Wabil, E. Al-Attas, A. Al-Abdulkarim, M. Chaurasia, and R. Alfaifi, , "Inline immediate feedback in Arabic web forms: An eye tracking study of transactional tasks," International Conference on Innovations in Information Technology (IIT), 2012.

[7] M. Seckler, A. Tuch, K. Opwis, and J. Bargas-Avila, "User-friendly locations of error messages in web forms: Put them on the right side of the erroneous input field," Elsevier, vol. 24, 2012.

[8] N. Maniar, "The effect of mobile phone screen size on video based learning," Journal of software, vol. 3, no. 4, April 2008.

[9] W. Mockovak, "Comparing the effectiveness of alternative approaches for displaying edit-error messages in web forms," Bureau of Labor Statistics, Statistical Survey Papers, 2005.

[10] S. Sweeney, and F. Crestani, "Effective search results summary size and device screen size: Is there a relationship?," Elsevier, vol. 42, July 2006, pp. 1056-1074.

[11] Y. Inal, "User-friendly locations of error messages in web forms: An eye tracking study," Journal of eye movement research, 2016, pp. 1-11.