

Usability Testing of Marking Menus

Darko Kuzmanovski¹

¹ Quod Financial,
20-22 Jean Jaurès Puteaux, France

Darko.kuzmanovski@gmail.com

Abstract. In this paper I investigate the usability of the multi-touch marking menus in multi-touch devices through three prototype applications that support multi-touch marking menus and I suggest a general area where they seem fit to be used. For the usability testing four testing methods were used. These methods are described further in this paper along with the resources used and the produced results. At the end of the paper the conclusions are given with the recommended area of usage of the multi-touch marking menus.

Keywords: usability, usability testing, multi-touch, marking menu.

1 Introduction

A very valuable and reliable input in designing marking menus [1] (two-handed marking menus for multi-touch devices [4]) is through a well-defined and executed usability testing [5]. The usability testing should point out the possible problems and positive aspects of usage, as well as suggesting means to overcome the weaknesses. Conducting a usability testing includes a choice of testing methods, resources, data gathering and data analysis. The usability testing is conducted to investigate the usability of the multi-touch marking menus in different kind of mobile applications and to define a general area where they seem fit to be used. The usability testing gives answers to the questions:

1. How easy and successfully the users are adopting the multi-touch marking menus;
2. How easy is to complete a task by using the multi-touch marking menus;
3. Can the novices use the marking menus without any user guide or help;
4. The most common issues encountered while using the multi-touch marking menus.

2 Multi-touch Marking Menus

The marking menus are gesture-based menus that extend the radial menu where the users can select an item from the radial menu by drawing a directional stroke [2]. Multi-touch devices can detect multiple points of contact. When these devices are put in the context of the multi-stroke [3] marking menus, means that these menus can be extended to use both hands for drawing the strokes in parallel or to overlap the hand

movements. By introducing the multi-touch devices, the next type (evolution) of marking menus are the multi-touch marking menus [4]. An example of the multi-touch marking menu can be seen on **Error! Reference source not found.**

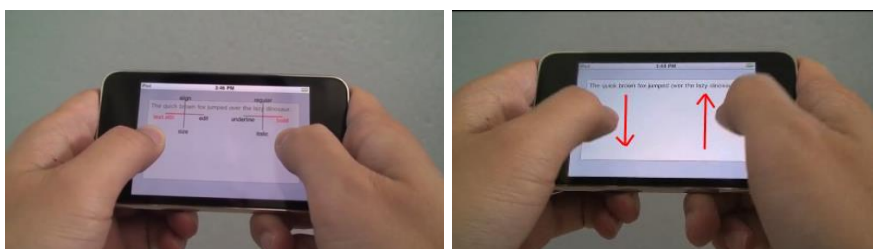


Fig. 1. Multi-touch marking menu

The marking menus have two distinctive operating modes. The first mode is when the user presses the screen with a lag and the radial menu pops up. This mode is called “menu” mode. This mode can be seen on the left in **Error! Reference source not found.** The second mode is called “stroke” or “mark” mode. This mode is when the user does not wait for the menu to popup and the item from the menu is selected with simple drawing of strokes (**Error! Reference source not found.** right).

3 Testing Methods and Resources

There is a variety of testing methods that can be used and they answer a different range of questions [6]. The methods used for the usability testing of the marking menus are:

- Usability questionnaire – this method contains a pre-test questionnaire, preferences measurement questionnaire, Mobile Phone Usability Questionnaire (MPUQ) [7].
- Heuristic evaluation [8] – your system (it can be a software application, web page, tool,...) is being evaluated by a set of professionals.
- Laboratory testing.
- SUS (System Usability Scale) questionnaire [9] – this method is similar to the “usability questionnaire” method. This questionnaire is given to the users that have already completed the laboratory testing.

Conducting a usability testing requires having resources in people and devices, and for example for the lab testing, an infrastructure that can be used as laboratory. The devices used are Samsung Galaxy S3 and an iPod. Three applications that have multi-touch marking menus and that were used for the testing purposes. **ControlUI** [13], a prototype application for multi-touch screens for a smart home or a vehicle. **Block Blender**, a game for training users to get used to the marking menus. This game has been removed from iTunes. **Text editor**, a prototype mobile application where the marking menus are used to edit a text.

3.1 Testing Setup

Each of the testing methods has its own configuration and methodology.

The first method, the questionnaires have been filled in by twelve users with different professions, age, technical knowledge and experience. Four of the users were female and eight were male. Two of the users were Medical Doctors, one QA engineer, six IT engineers and one pharmacist and two students. The age span was from 21 to 54 years old. These users were filling out google forms, from their homes, for the three groups of questionnaires.

For the heuristic evaluation, five experts in mobile technologies and the user experience design were engaged. Two of the experts are user experience designers, two are front-end programmers and one is a solution architect. The results from the heuristic evaluation from the experts are analyzed separately and after that they are analyzed all together. At the end a list with identified problems is formed. They give their comments for the following heuristics [10]:

- HM1: Visibility – the options and the actions should be clearly marked, intuitive and not ambiguous;
- HM2: Consistency and standards in navigation – the actions have to be consistent and logical;
- HM3: The operations should not be destructive – after an error, the return to the previous state should be easy and transparent;
- HM4: Scalability – The menu usage should be easy for different screen sizes;
- HM5: Reliability – The operations should finished without any undesired events;
- HM6: Help during usage – the marking menus can be used without documentation, there should be a manual or user guide on how to use the marking menus;
- HM7: Response – The marking menus should respond nearly instantaneously after the user action;

For the laboratory testing of the multi-touch marking menus, an unmoderated laboratory testing has been performed, i.e. guerrilla testing [11] with twelve testers. All of the testers are male and IT engineers with different experience. The tests were in duration of 20 minutes. After the testers have passed some arbitrary time with multi-touch marking menus, they started executing the predefined scenarios. The measured parameters during the laboratory testing are:

- The elapsed time for executing the scenario;
- The number of actions for completing the scenario;
- Identifying the problems the testers face during the execution of the scenarios;
- The positive and the negative remarks of the testers.

The scenarios used for the laboratory testing:

- Scenario 1 (marking menus in the ControlUI app): you are entering in your smart home. You can control your home from your phone by using the ControlUI app. By using the app, turn on the music and the air conditioning.

- Scenario 2 (Text editor): You have an unedited document to send to the management. Luckily it is not a big deal and you can do it while taking the train with the Text edit app. Make two paragraphs bold, and cut and paste two paragraphs.
- Scenario 3 (Block blender): There is a game that your friend have mastered and are using it in master mode. You want to catch their level so you need to practice. Play the game with shown moves and after that without.

The SUS questionnaire was filled out by the same users that did the laboratory testing. It is populated just after the laboratory testing, but before the discussion with the user. The users are asked to give an answer to the questionnaire without second thinking [12]. The questionnaire contains ten questions, on a scale from one to ten. If the user cannot decide for an answer, then the middle answer is taken as the chosen answer.

4 Results from the Usability Testing

In this section the results of each of the testing methods are shown.

4.1 Usability Questionnaire

The possible answers in the “preference measurement” questionnaire are: Strongly don’t agree, Don’t agree, Neutral, I agree and “Strongly agree”.

1. The multi-touch marking menus are useful – (16.7% don’t agree, 16.7% neutral, 33.3% agree and 33.3% strongly agree).
2. The multi-touch marking menus meet your expectations – (16.7% don’t agree, 16.7% neutral, 50% agree and 16.6% strongly agree).
3. The features of these menus are corresponding to your need – (8.3% don’t agree, 58.3% neutral, 50% agree and 16.7% strongly agree).
4. The multi-touch marking menus are easy to use – (33.3% neutral, 41.7% agree and 25% strongly agree).
5. The multi-touch marking menus are easy to learn – (25% neutral, 16.7% agree and 58.3% strongly agree).
6. The multi-touch marking menus are easily accessible. (8.3% strongly don’t agree, 8.3% don’t agree, 16.7 neutral, 50% agree and 16.6% strongly agree).
7. I need help to use the multi-touch marking menus – (8.3% strongly don’t agree, 25% don’t agree, 41.7% neutral, 33.3% agree)
8. The multi-touch marking menus are better than the linear menus - (8.3% strongly don’t agree, 8.3% don’t agree, 41.7% neutral, 41.7% agree)
9. I would use the multi-touch marking menus all the time. (8.3% don’t agree, 25% neutral, 41.7% agree and 25% strongly agree).
10. The multi-touch marking menus should be more present in the mobile devices. (16.7% don’t agree, 25% neutral, 25% agree and 33.3% strongly agree)

The answers of the questions for each of the areas are summarized and divided with the number of the questions for the same area. If the user did not answer certain ques-

tion, it gets the value of a neutral answer. According to this questionnaire the certain area is a “pass” if the total points are above four for that area. The questionnaire is a “pass” if the average value of the points is above four.

The results show that the multi-touch marking menus in the ControlUI application are: usable in the ELU, EAMC and EC areas and unusable for the AOPS and the CMML areas. The total average value is 4.33, which means that the marking menus have a “pass” value in terms of usability.

The results show that the multi-touch marking menus in the TextEdit application are: usable in the ELU, EAMC and EC areas and unusable for the AOPS and the CMML areas. The total average value is 3.89, which means that the marking menus have a “no pass” value in terms of usability.

The results show that the multi-touch marking menus in the Block Blender game application are: usable in the ELU, EAMC, CMML and EC areas and unusable for the AOPS area. The total average value is 4.21, which means that the marking menus have a “pass” value in terms of usability.

Results from the pre-test questionnaire:

- How old are you? - Most of the users are between twenty and thirty five years;
- Sex - Eight of the users are male and four are female;
- For how long do you use a mobile device? - The chosen users have experience with mobile devices;
- Do you own a smart phone with a touch screen? - Most of the users have smart phones (83.3%);
- How much time do you spend on your phone per day? - Half of the users use their mobile device between one and three hours;
- On which utility do you spend most of the time on your phone? - The mobile devices are mostly used for the applications and the multimedia functions;
- What kind of apps do you usually use on your phone? - The business apps are used from 41.6% of the users and 25% of the users use social apps;
- Do you know what are marking menus? - Most of the users didn’t know.

4.2 Results from the Heuristic Evaluation

The results from the heuristic evaluation are shown in **Table 1**, **Table 2** and **Table 3**. In each of the tables the rows are the evaluators and the columns are the heuristics as shown in **3.1**. Each of the fields is the grade given by the evaluator for the specific heuristic. Grading levels:

1. There is a serious problem which makes the multi-touch marking menus unusable;
2. There is a serious issue that has to be removed;
3. There is an issue which doesn’t interfere with the user’s workflow, but removing this issue will improve the usability;
4. No issue found.

The results from the heuristic evaluation of the multi-touch marking menus in the application ControlUI are shown in **Table 1**. The lowest grade is given to the heuristic HM6. Some of the evaluators consider that the marking menus are unusable for the heuristic HM1. Poor grades were given also for the heuristic HM2. For heuristic HM3 in certain situations the number of fingers is not well detected.

Table 1. Heuristic evaluation of the multi-touch marking menu in the application ControlUI

	<i>HM1</i>	<i>HM2</i>	<i>HM3</i>	<i>HM4</i>	<i>HM5</i>	<i>HM6</i>	<i>HM7</i>
1	1	2	4	4	2	1	4
2	2	2	3	2	3	1	4
3	2	1	3	4	2	1	4
4	2	2	4	3	2	1	4
5	1	1	3	4	3	1	4

The results from the Heuristic evaluation of the multi-touch marking menu in the application TextEdit are shown in **Table 2**. The lowest grade was given for the heuristic HM6; also low grades were given for the heuristic HM1. While the menus are in landscape mode of the screen, the menu items are not well visible. Regarding HM2, the menu items are difficult to select and also hard to remember. For HM3 the remark is that some menu items like “undo” and “redo” should be easily accessible. No remarks for the heuristics HM4 and HM7. For the heuristic HM5, the menus should be more accurate in detection and more fault tolerant.

Table 2. Heuristic evaluation of the multi-touch marking menu in the application TextEdit

	<i>HM1</i>	<i>HM2</i>	<i>HM3</i>	<i>HM4</i>	<i>HM5</i>	<i>HM6</i>	<i>HM7</i>
1	1	3	4	4	3	1	4
2	1	3	3	4	4	1	4
3	3	3	3	4	3	1	4
4	2	2	3	4	3	1	4
5	3	3	4	4	4	1	4

The results from the Heuristic evaluation of the multi-touch marking menu in the application Block Blender are shown in **Table 3**. There were no remarks from the evaluators regarding the heuristics HM3 and HM7. The evaluators gave lowest grades for the heuristic HM6.

Table 3. Heuristic evaluation of the multi-touch marking menu in Block Blender

	<i>HM1</i>	<i>HM2</i>	<i>HM3</i>	<i>HM4</i>	<i>HM5</i>	<i>HM6</i>	<i>HM7</i>
1	3	3	4	4	4	1	4
2	3	3	4	3	3	1	4
3	2	4	4	3	3	1	4
4	4	3	4	3	2	1	4
5	3	2	4	4	4	1	4

4.3 Results from the Laboratory Testing

The problems found during the execution of the laboratory testing of the Scenario 1:

- The user expects the menus to have indication that they can be called.
- The user expects the items of the menu to be unambiguous.
- The user expects to be able to display all the menu items.
- The user expects accurate finger detection.
- The user expects the menu action then one finger is used.

The problems found during the execution of the laboratory testing of the Scenario 2:

- The user expects the menus to be equally accessible when the screen is vertical (portrait mode) and when the screen is horizontal (landscape mode);
- The user expects to have enough space to view the text in landscape mode;
- The user expects the frequent menu items to be easily accessible.
- The user expects to have more visibility when the menus are in the “menu” mode.

The problems found during the execution of the laboratory testing of the Scenario 3:

- The users expect user guide for the marks in the game
- The users expect accurate finger detection.

The measured parameters during the laboratory testing are shown at section 3.1. Each of the scenarios is executed three times. The measurements regarding multi-touch marking menus in the application ControlUI while executing Scenario 1 (section 3.1) are shown in **Table 4**. From the table we can observe that the time required to finish the scenario is shrinking.

Table 4. Results from the measurements when executing Scenario 1

	Try 1		Try 2		Try 3	
	Competition time (seconds)	Stroke number	Competition time (seconds)	Stroke number	Competition time (seconds)	Stroke number
U1	22	7	10	4	4	3
U2	20	6	8	3	5	3
U3	18	4	7	3	4	2
U4	17	5	8	3	3	2
U5	24	7	7	3	2	2
U6	25	8	7	3	2	2
U7	19	6	8	3	3	2
U8	10	3	6	3	3	2
U9	14	4	5	2	2	2
U10	16	5	4	2	2	2
U11	13	4	5	2	2	2
U12	7	2	2	2	2	2
Average	15,6	5,1	6,4	2,8	2,8	2,2

In **Table 5** Error! Reference source not found. Error! Reference source not found. are shown the results from the measurement during the execution of Scenario 2 (section 3.1). The scenario has been executed in two modes, the “menu” mode and the “mark” mode. The scenario is expected to finish in sixty seconds. It can be seen that during the first try the users spend a lot of time to complete. As the tries are repeating, the elapsed time is shrinking. After some tries the users are becoming experts, but some still need more tries to be able to use the menus in “mark” mode.

Table 5. Results from the measurements when executing Scenario 2

“menu”	Try 1		Try 2		Try 3	
	Competition time (seconds)	Stroke number	Competition time (seconds)	Stroke number	Competition time (seconds)	Stroke number
U1	30	15	20	10	12	8
U2	35	18	17	9	9	6
U3	29	15	11	6	6	6
U4	37	20	21	10	10	6
U5	25	14	15	9	8	6
U6	28	14	15	8	7	7
U7	31	13	14	8	8	7
U8	30	13	13	7	8	6
U9	35	17	20	11	12	9
U10	29	11	15	9	11	9
U11	24	12	11	6	9	6
U12	26	12	16	9	8	6
Average	29,9	14,5	15,7	8,5	9	6,8
“mark”	Try 1		Try 2		Try 3	
U1	Not finished	35	30	15	10	8
U2	Not finished	55	36	18	15	10
U3	Not finished	51	29	15	13	8
U4	Not finished	62	37	20	10	9
U5	45	22	15	14	12	8
U6	48	28	15	14	11	10
U7	Not finished	40	19	13	10	10
U8	Not finished	39	30	14	14	11
U9	40	20	35	18	10	9
U10	Not finished	65	29	11	12	11
U11	35	20	24	12	10	9
U12	38	25	26	13	10	7
Average		38,5	27,1	14,5	11,4	9,2

In **Table 6** are shown the results from the measurement during the execution of Scenario 3 (section 3.1). The measurements are made when the user is a novice and when the user becomes an expert. The expected time of completion of the scenario is two min. In the game the blocks are falling faster in each try. In expert mode all users

failed to complete the scenario because they could not follow easily the needed strokes (game over).

Table 6. Results from the measurements when executing Scenario 3

Novice	Try 1		Try 2		Try 3	
	Competition time (seconds)	Stroke number	Competition time (seconds)	Stroke number	Competition time (seconds)	Stroke number
U1	120	244	120	245	89	185
U2	120	245	120	244	65	135
U3	120	248	120	240	120	240
U4	120	240	120	240	78	160
U5	120	249	120	249	79	161
U6	120	250	120	197	55	115
U7	120	244	120	254	110	223
U8	120	255	89	250	99	200
U9	120	258	120	242	89	184
U10	120	245	120	248	120	240
U11	120	260	120	247	120	240
U12	120	258	105	215	95	194
Average	120	249,7	114	239,3	93,3	189,8
Expert	Try 1		Try 2		Try 3	
U1	20	30	18	23	18	20
U2	15	22	15	21	22	25
U3	22	30	20	26	14	15
U4	25	36	22	29	12	13
U5	35	50	15	22	11	13
U6	40	57	30	39	28	29
U7	29	42	18	22	14	16
U8	27	39	20	23	10	11
U9	24	35	26	31	18	20
U10	20	28	29	34	14	17
U11	30	44	33	38	25	28
U12	33	50	31	37	20	23
Average	26,7	38,6	23,1	28,8	17,2	19,2

For the results of the SUS questionnaire: The total result of 68,125 makes the multi-touch marking menus in the ControlUI application usable. The total result of 66.3 makes the multi-touch marking menus in the TextEdit application unusable. The result of 70.8 makes the multi-touch marking menus in the Block Blender game usable.

5 Conclusion

In the beginning of the paper there were four questions on which the usability testing will give answer.

1. If the radial menu is present, then the multi-touch marking menus are easy to be learnt. The transition from novice to expert is faster and more efficient. If the menu has a lot of submenus then the transition is slower, but for a frequently used elements from the menu, the transition takes around ten tries.
2. The completion of the task depends of the user level and if there is a “menu” mode present. For a beginner, the task completion can be confusing.
3. The marking menus are not the usual type of menus and the users need help when they are using them for the first time. In a mobile app there should be a way to guide the users while using the app. The best way to achieve fast and ad-hoc help is through the so called “overlay help” screens.
4. The first impression is positive, but the enthusiasm drops when faced with unfamiliar interface. The daily interaction with the linear menus has a huge impact over the perception on how menus are expected to work. The games are an exception where the users expect a period of learning. A drawback is the small screen size.

The recommended areas of usage are games and in applications that do not require high precision and user concentration (similar to the ControlUI app). The design should be fault tolerant, but on the other hand to be precise and unambiguous.

References

1. Kurtenbach, G., & Buxton, W.: Issues in combining marking and direct manipulation techniques. Proceedings of the Fourth ACM SIGGRAPH Symposium on User Interface Technology (UIST), 137-144 (1991)
2. Kurtenbach, G.: The design and evaluation of marking menus. Ph.D. dissertation, University of Toronto. (1993)
3. Zhao, S., and Balakrishnan, R.: Simple vs. compound mark hierarchical marking menus. UIST '04 Proceedings of the 17th annual ACM symposium on User interface software and technology. 33-44. (2004)
4. Kin, K., Hartmann, B., and Agrawala, M.: Two-handed marking menus for multitouch devices. ACM Trans. Comput.-Hum. Interact. 18, 3, Article 16 (July 2011), 23 pages (2011)
5. Nielsen Norman Group, <http://www.nngroup.com/articles/usability-101-introduction-to-usability/> (last visited May 2015)
6. Nielsen Norman Group, <http://www.nngroup.com/articles/which-ux-research-methods/>
7. Y., S., Ryu.: Development of usability questionnaires for electronic mobile products and decision making methods. Dissertation Submitted to the Faculty of Virginia Polytechnic Institute and State University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Industrial and Systems Engineering. (2005)
8. <http://www.usability.gov/how-to-and-tools/methods/heuristic-evaluation.html> (May 2015)
9. <http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>

10. Shirlina, Po1, S., H.: Heuristic Evaluation and Mobile Usability: Bridging the Realism Gap. Interaction Design Group, Department of Information Systems, The University of Melbourne. Australia (2004)
11. <http://www.jisc.ac.uk/guides/usability-and-user-experience> (last visited May 2015)
12. Brooke, J.: SUS: A "quick and dirty" usability scale. Usability evaluation in industry, London: Taylor and Francis pp. 189-194 (1996)
13. <http://matthaeuskrenn.com/new-car-ui/> (last visited May 2015)