Fraunhofer
PORTUGAL

FRAUNHOFER CENTER FOR ASSISTIVE INFORMATION AND COMMUNICATION SOLUTIONS – AICOS

REPORT

Keyboards usability comparison testing GoLivePhone / SmartCompanion

By: Ana Vasconcelos

Date: October 2015

Fraunhofer Center for Assistive Information and Communication Solutions – AICOS
Associação Fraunhofer Portugal Research
Rua Alfredo Allen 455/461, 4200-135 Porto, PORTUGAL

Table of Contents

1 Summary

System and objectives

This experiment was conducted within the context of the GoLivePhone/SmartCompanion Project (https://www.aicos.fraunhofer.pt/en/our_work/projects/SmartCompanion.html), during the development of a custom designed keyboard. During this evaluation we tested three different keyboards, commercially available as the default keyboards on the three main mobile operating systems: Android, iOS and Windows Phone. The goal was to analyze the interaction of senior users with on-screen smartphone keyboards. The results of this test will be used to inform the design of a new on-screen smartphone keyboard targeting senior users.

Method

Three keyboards were tested with 13 users at FhP's living lab. Participants were asked to perform the same text input task in the different keyboards.
To evaluate the usability of the system, we measured effectiveness (via task completion), efficiency (via task completion time) and satisfaction (via an interview).

Results

On average, the Windows Phone keyboard was the most error-prone with seven errors, followed by Android with five and iOS with four. Maybe due to the fact that it was more error-prone, the Windows Phone keyboard was also where users spent most time to write the text (3m24s), followed by Android with 2m54s. iOS was the most efficient keyboard with an average of 2m43s.
The iOS default keyboard was perceived as the easiest by five participants and as easy as the others by four participants, summing up

to a total of nine participants. Two participants found the Android default keyboard to be their favorite and five others placed it among their top, seven in total. Only one participant considered the Windows Phone default keyboard as better than the others, while three more participants considered it to be as good as others, completing a total of four participants.

2 Introduction

2.1 System description

During this evaluation we tested three different keyboards, commercially available as the default keyboards on the three main mobile operating systems: Android, iOS and Windows Phone. The OS versions used were: Android 5.1.1, iOS 9.0.2 and Windows Phone 8.1. These keyboards are intended for any smartphone user and are used for text or number input. Even though this test was developed under the scope of the GoLivePhone/SmartCompanion project, the tested products are not a direct result of this project.

2.2 Test objectives

The goal of this test was to analyze the interaction of senior users with on-screen smartphone keyboards. We were interested in understanding common challenges posed by these keyboards, measure performance of tasks, and get an overall opinion from users.

The results of this test were used to inform the design of a new on-screen smartphone keyboard targeting senior users.

3 Method

3.1 Test facility

Tests took place at FhP's living lab. Each participant had a one-on-one session with one facilitator, in a room that simulates a living room.

Figure 1. Participant and moderator during a test

3.2 Equipment

During each session participants interacted with three different smartphones, with different operating systems and using different note-taking applications, as shown on the following table and figures.

Smartphone	OS	Application
Nexus 5	Android 5.1.1	Google Keep
iPhone 6	iOS 9.0.2	Notes
Nokia Lumia 630	Windows Phone 8.1	OneNote

Table 1. Used equipment

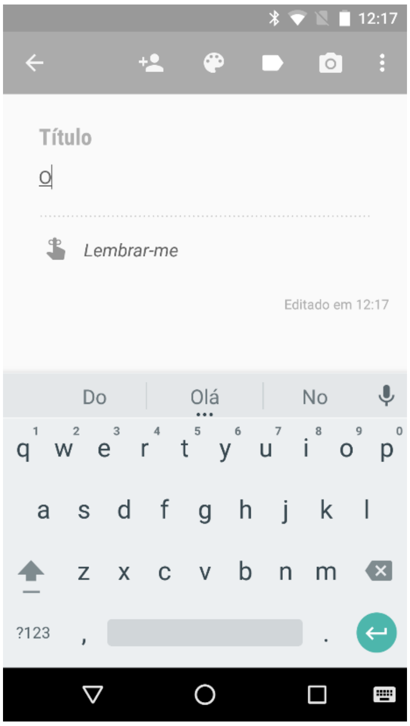

Figure 2. Android keyboard/Google Keep application

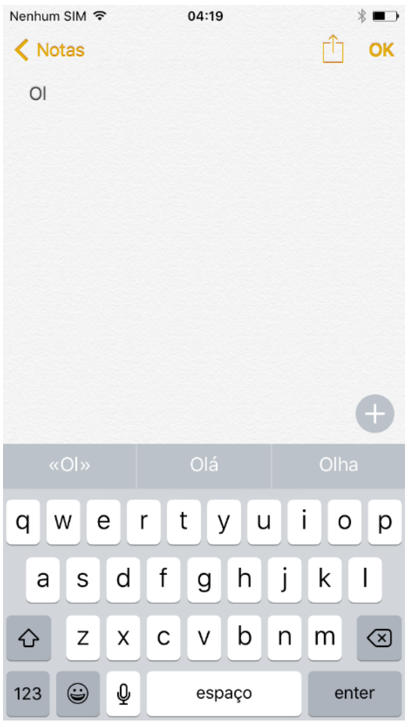

Figure 3. iOS keyboard / Notes application

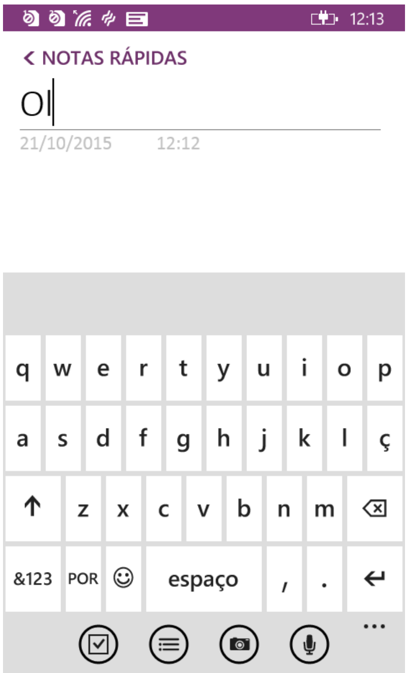

Figure 4. Windows Phone keyboard/ One Note application

Participants were asked to specifically interact with the default keyboard on each smartphone, set for the Portuguese language, with suggestions enabled but disabled autocorrect.

3.3 Procedure

The sequence of events from greeting the participants until their dismissal was the following:

- Participants were greeted by the facilitator.
- They were given an informed consent to sign (available at section 6.2 of the Annex).
- A short background questionnaire (age, smartphone ownership and previous experience with SMS) was administered.
- Information about the test was read aloud from the script (available at section 6.1 of the annex) by the facilitator. Participants were asked to try and complete the task as if the

facilitator was not present, but to ask for help if they felt they were stuck or did not understand the task description. The facilitator also tried to elicit some comments from participants during task execution to understand their though process.

- Participants were given the smartphones sequentially, in a random order.
- Questions and comments about the product were solicited.
- The facilitator thanked and dismissed the participants.

There was no time limit to complete a task and participants were not compensated.

3.4 Tasks

In each session, participants were asked to write a sample test using the three smartphones and their respective on-screen keyboards. The used text was: "*Olá! Hoje faço anos, queres ir lá a casa?".* Participants were given this text on a piece of paper, since there was no need for them to memorize it.

3.5 Usability metrics

3.5.1 Effectiveness

Effectiveness relates the goals of using the product to the accuracy and completeness with which these goals can be achieved. The following metrics were collected:

- **Completion rate**: The proposed task was considered completed when the user finished typing the given text on the smartphone.
- **Errors**: An error was counted every time the participant performed an action that did not contribute to task completion.
- **Assists**: An assist was considered every time the participant requested the assistance of the facilitator in order to perform the

task. If the assistance was required because the task was not well explained it was not considered.

3.5.2 Efficiency

Efficiency relates the level of effectiveness achieved to the quantity of resources expended. It was assessed by the mean time taken to complete the given task.

3.5.3 Satisfaction

Satisfaction describes a user's subjective response when using the product. After the task, participants were informally asked (there was no formal questionnaire made) to give their opinion regarding the evaluation and to compare the several keyboards they had just used.

4 Results

4.1 Participants

For this experiment we collected data from 13 senior participants with ages ranging from 58 to 77 years old (average = 68), six males and seven females. Seven participants used a feature phone while the remaining six already owned a smartphone. Six participants were already familiar with writing SMS's while seven of them had never tried that functionality on their phones.

Participant	Age	Gender	Type of mobile phone	Experience with SMS
P1	58	Male	Feature phone	Yes
P2	68	Female	Smartphone	Yes
P3	65	Male	Smartphone	Yes
P4	69	Female	Smartphone	Yes
P5	69	Female	Feature phone	Yes
P6	70	Male	Feature phone	No
P7	72	Female	Feature phone	No
P8	65	Male	Feature phone	No
P9	69	Female	Smartphone	No
P10	65	Female	Feature phone	No
P11	67	Female	Smartphone	Yes
P12	77	Male	Feature phone	No
P13	72	Male	Smartphone	No

Table 2. Participant's list

4.2 Performance results

All testing sessions were recorded using a video camera. The recorded videos were imported and coded using the Observer XT software. The coding scheme designed for this study had two groups. The Metrics group included Errors (point event), and Assists (point event). The Comments groups included Comment (point event) and Task (state event).

The data were then analyzed in Observer XT, allowing the researchers to quantify the number of errors and assists for each participant, as well as register the task (beginning and end) and comments.

An error was marked every time the participant performed an action that did not contribute to task completion. An assist was considered every time the facilitator intervened (at the participant's request or due to the facilitator's judgment) to help the participant in the completion of the task.

4.2.1 Android overall results

Participant	Unassisted Task Effectiveness [(%)Complete]	Assisted Task Effectiveness [(%)Complete]	Task time (min)	Errors	Assists
P1	0%	100%	05:17	14	2
P2	0%	100%	01:51	6	1
P3	100%	100%	01:24	2	0
P4	100%	100%	01:11	1	0
P5	100%	100%	02:45	4	0
P6	0%	100%	04:13	11	2
P7	0%	100%	04:20	1	2
P8	100%	100%	01:34	3	0
P9	0%	100%	04:00	5	3
P10	0%	100%	05:14	12	1

P11	100%	100%	01:34	2	0
P12	100%	100%	02:42	3	0
P13	100%	100%	01:41	0	0
Mean	57%	100%	02:54	5	1
St. Deviation	0,51	0	01:31	4,57	1,07
Min	0%	100%	01:11	0	0
Max	100%	100%	05:17	14	3

Table 3. Android Results

4.2.2 iOS overall results

Participant	Unassisted Task Effectiveness [(%)Complete]	Assisted Task Effectiveness [(%)Complete]	Task time (min)	Errors	Assists
P1	100%	100%	04:13	13	0
P2	100%	100%	01:12	0	0
P3	0%	100%	01:31	1	1
P4	100%	100%	01:34	1	0
P5	100%	100%	01:58	0	0
P6	0%	100%	03:00	5	3
P7	0%	100%	03:12	4	2
P8	0%	100%	03:10	4	2
P9	100%	100%	02:45	4	0
P10	0%	100%	04:16	9	3
P11	0%	100%	01:43	1	2
P12	0%	100%	04:21	2	2
P13	0%	100%	02:29	4	1
Mean	36%	100%	02:43	4	1
St. Deviation	0,50	0	01:06	3,75	1,17

Min	0%	100%	01:12	0	0
Max	100%	100%	04:21	13	3

Table 4. iOS Results

4.2.3 Windows Phone Results

Participant	Unassisted Task Effectiveness [(%)Complete]	Assisted Task Effectiveness [(%)Complete]	Task time (min)	Errors	Assists
P1	0%	100%	06:28	26	3
P2	0%	100%	02:54	7	1
P3	100%	100%	01:52	2	0
P4	100%	100%	01:28	1	0
P5	100%	100%	02:52	3	0
P6	0%	100%	03:59	14	2
P7	0%	100%	04:06	6	1
P8	100%	100%	02:44	4	0
P9	100%	100%	04:26	10	0
P10	100%	100%	03:43	9	0
P11	100%	100%	01:27	2	0
P12	0%	100%	05:01	10	3
P13	0%	100%	03:09	3	1
Mean	50%	100%	03:24	7	1
St. Deviation	0,52	0	01:27	6,81	1,14
Min	0%	100%	01:27	1	0
Max	100%	100%	06:28	26	3

Table 5. Windows Phone Results

4.2.4 Comparison results

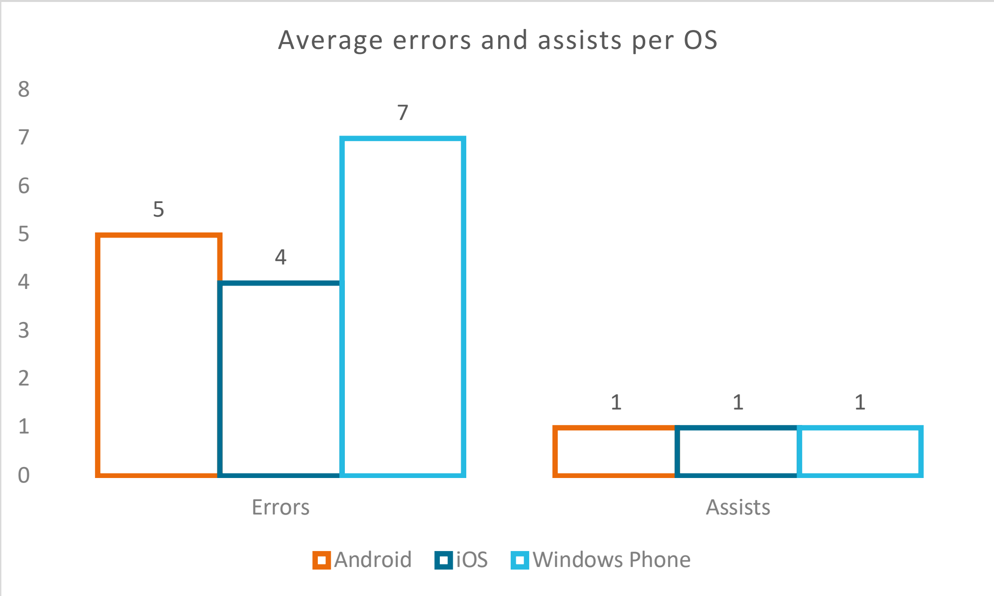

Figure 5. Average errors and assists per OS

On average, the Windows Phone keyboard was the most error-prone with seven errors, followed by Android with five and iOS with four. Regarding assists, there was no difference among operating systems with each user, requiring on average one assist per OS.

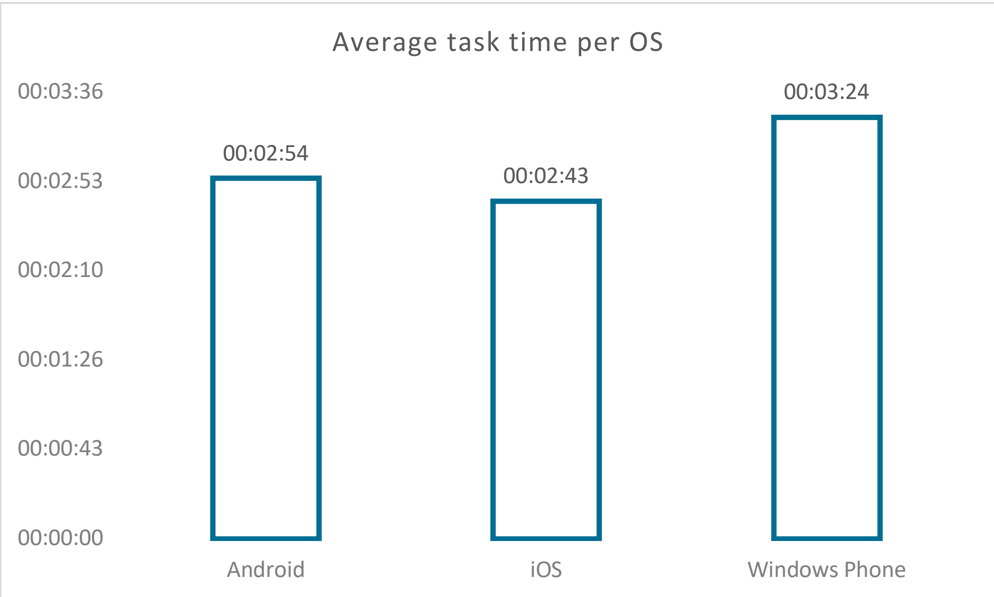

Figure 6. Average task time per OS

Maybe due to the fact that it is was more error-prone, the Windows Phone keyboard was also where users spent most time to write the text (3m24s), followed by Android with 2m54s. iOS was the most efficient keyboard with an average of 2m43s.

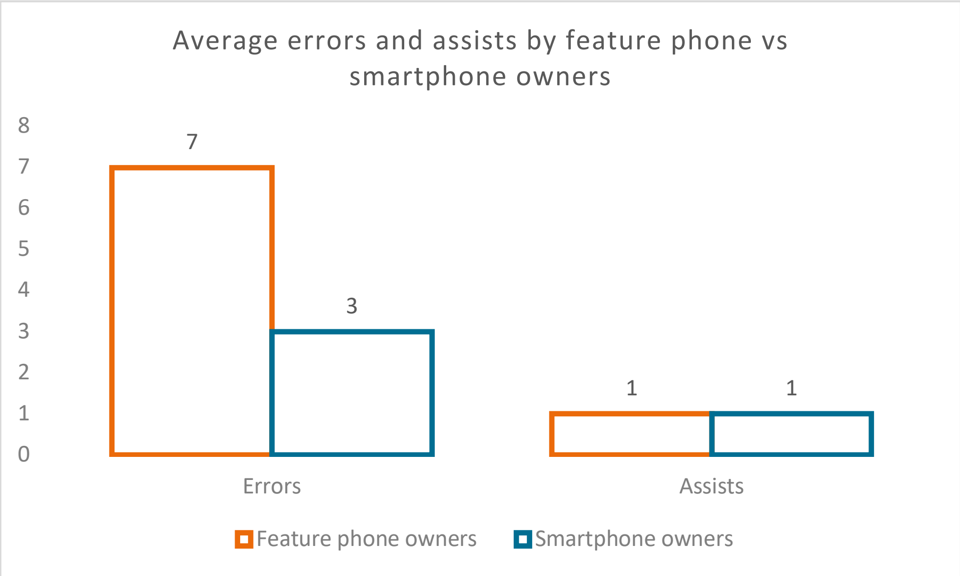

Figure 7. Average errors and assists by feature phone vs smartphone owners

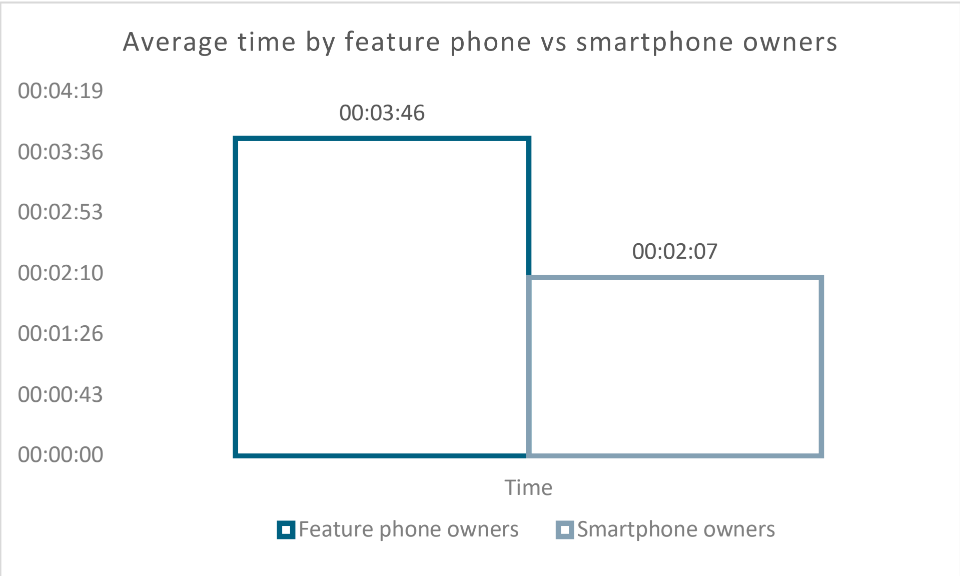

Figure 8. Average time by feature phone vs smartphone owners

In order to understand the impact of having previous experience with a smartphone in the performance of this evaluation, we analyzed the difference between average errors, assists and time by segmenting users into two groups: feature phone owners and smartphone owners. Results show that on average feature phone owners commit more errors but needed the same number of assists. Smartphone owners were an average of 1m39s faster than feature phone users.

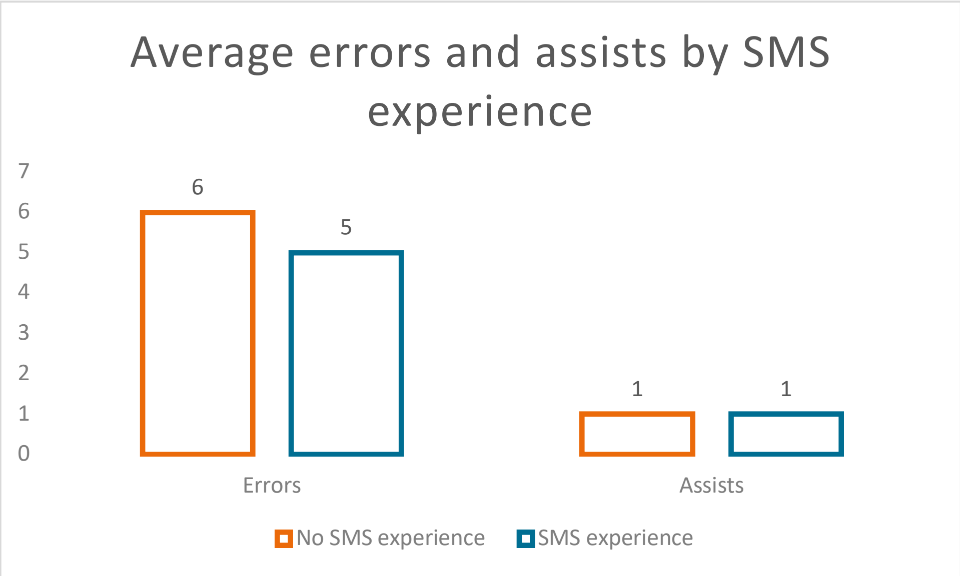

Figure 9. Average errors and assists by SMS experience

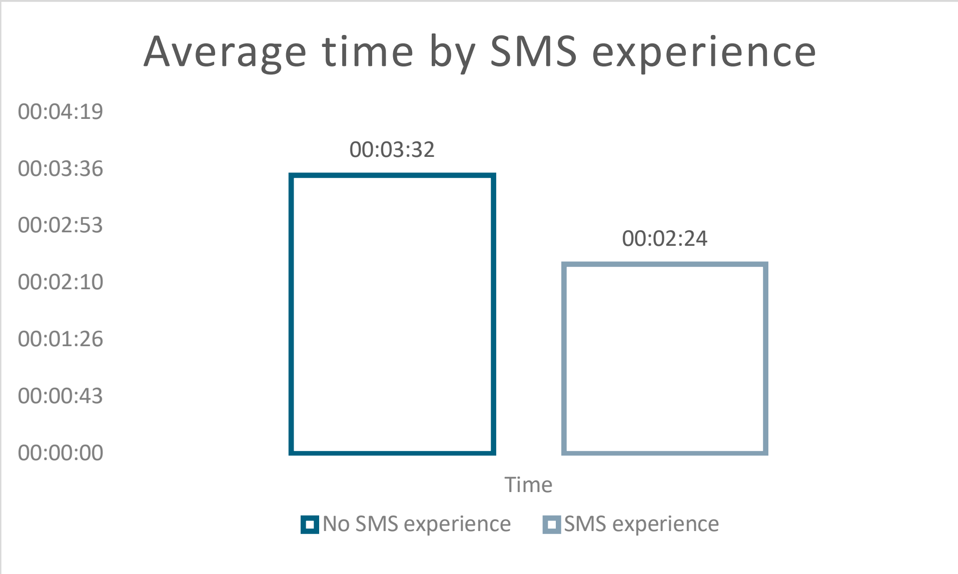

Figure 10. Average time by SMS experience

Next, we analyzed the impact of having previous experience with SMS by segmenting users into two groups. Results show that differences in errors made and assists needed by users with and without SMS experience are not significant. On average, users with previous experience were 1m8s faster to write the text.

During the tests it was also noted that most participants improved at least one performance metric (errors, assists or time) overtime, i.e., even though the keyboards were different they were able to use the knowledge acquired during the test to improve their performance. Detailed results can be found in Section 6.3 of the Annex of this report.

4.3 Satisfaction results

After the evaluation session participants were asked to indicate their preferred(s) keyboard. When a participant indicated two keyboards as similar, we considered both to be his favorites. Results are in the table below:

Participant	Android	iOS	Windows Phone
P1			X
P2	X	X	
P3	X		X
P4	X	X	
P5		X	
P6		X	
P7	X		
P8		X	
P9		X	
P10		X	
P11	X	X	X
P12	X		
P13	X	X	X
Total	7	9	4

Table 6. Participants' preferred keyboard

The iOS default keyboard was perceived as the easiest by five participants and as easy as other by four participants, summing up to a total of nine participants. Two participants found the Android default keyboard to be their favorite and five others placed it among their top, seven in total. Only one participant considered the Windows Phone default keyboard as better than the others, while three more participants considered it to be as good as others, completing a total of four participants.

4.3.1 Observations and participants' comments

Participant	Observations and comments
P1	Confused the Enter key with the Delete key; Difficulties in pressing the 'a' button on Windows Phone – probably because it was too close to the screen border; Long pressed the 'a' key (it has been previously used to introduce 'á') to search for 'ç' Found the ',' on the main keyboard; Commented that he felt some challenges during the first task but acknowledge that it was normal since it was his first time interacting with a smartphone. We also indicated that after the first task he felt more confident and already knew how to introduce punctuation; He felt that the font on the Android keyboard was smaller than the others; Acknowledged that the iPhone keyboard made a small sound when pressing a key.
P2	Used the keyboard suggestions to introduce punctuated letters or complete words; Commented that the WP keyboard seemed more complete because it included the 'ç' in the main keyboard.
P3	Already knew how to introduce punctuation and even used a long press on the '.' key to introduce the '!'.
P4	Did not know how to introduce punctuation or the 'ç' but learned with just one explanation and applied the knowledge to the following keyboards; Found the ',' on the main keyboard; Commented that the WP keyboard was harder because it lacked space between keys; Was able to switch to capital letters alone.
P5	Already knew how to introduce capital or delete; Used the keyboard suggestions to introduce punctuated letters or complete words; Already knew how to introduce punctuation; After choosing a suggestion introduced a space after the word (which had already been done automatically); Difficulties in pressing the 'ç' button on Windows Phone – probably because it was too close to the screen border.
P6	Searched for 'ç' on the punctuation keyboard; Faced challenges when first trying to introduce punctuated words (drag gesture); Did not know how to introduce punctuation but learned with just one explanation and applied the knowledge to the following keyboards; Searched for 'á' on the punctuation keyboard; Used the iPhone's home button while trying to introduce a space; Tried long pressing the punctuation keyboard key - was confused about how to introduce 'á' or other punctuation; After learning how to delete, used that knowledge for other keyboards.
P7	Used the ',' on the main keyboard; Needed two assists to learn how to delete.
P8	Needed one assist to learn how to introduce 'á', punctuation and delete; Used the iPhone's home button while trying to introduce a space;

	Tried long pressing and dragging on the punctuation keyboard key; Confused the Enter key with the Delete key; Searched for 'ç' on the punctuation keyboard; Commented that the punctuation keyboard had smaller fonts.
P9	Consistently and purposely introduced two spaces between words – sometimes this activated the '.' Shortcut; Commented that she felt more confident after completing each task; It was difficult for her to maintain the long press due to hand tremors.
P10	After an assist to learn how to introduce the 'á' was able to use that knowledge to find the 'ç'; Pressed the keys with more pressure and for longer than necessary – on Android that meant that certain letters were switched to numbers; Started using the suggestions once she figured out what they were; Pressed the delete key for longer than necessary which meant that more letters than intended would be deleted.
P11	After an assist to learn how to introduce the 'á' was able to use that knowledge to find the 'ç'; Did not know how to introduce punctuation but learned with just one explanation and applied the knowledge to the following keyboards; Used the keyboard suggestions to introduce punctuated letters or complete words; After choosing a suggestion, introduced a space after the word (which had already been done automatically).
P12	Was able to immediately identify if the keyboard was set to caps and how to change it; Difficulties in pressing the 'a' button on Windows Phone – probably because it was too close to the screen border; Needed one assist to learn how to introduce 'á', punctuation and delete; Faced challenges when first trying to introduce punctuated words (drag gesture); Noticed the zoom of the letters on iPhone; Commented that the Android keyboard seemed more sensitive to touch and therefore easier to use.
P13	Used the keyboard suggestions to introduce punctuated letters or complete words; Confused the ' ' ' with the ','; After choosing a suggestion introduced a space after the word (which had already been done automatically).

Table 7. Observations and participants' comments

5 Recommendations

Based on the results from the tests, the facilitators' notes and observations during the sessions, the user comments and the analysis of the videos coded using Observer XT, the following recommendations were derived.

Base the design of a custom keyboard on the iOS default keyboard. Results show that this keyboard was less error-prone, participants were faster to perform, and, when asked, they also referred to this keyboard as their favorite, or among their favorites.

Make the difference between the enter button and the delete button clear.
On the Android and on the WP keyboards, both buttons used an arrow pointing to the left, which caused some doubts to the users. A solution like the one used by iOS – writing the word Enter on the key – would be more suitable.

Disable autocorrect but maintain the suggestions bar.
Autocorrect can sometimes introduce errors while typing a text. The suggestion bar however is widely used by more experienced users, either to complete words or to introduce punctuated letters.

Let users know if the keyboard is set to caps.
It is easier for a user to know if he is about to write a capitalized letter if the keyboard reflects that. Therefore, if the caps key is activated the keyboard should show the letters in capitals.

Remove the drag option when selecting punctuated letters.
Many users faced challenges with the dragging gesture to select a punctuated letter. A solution as the one proposed on WP – open sticky menu and let users choose the letter – is easier and less error-prone.

Ignore consecutive spaces.

Two situations in particular led users to introduce more than one space between two words: after selecting a suggested word which already introduced a space and hand tremors that led to the space key being pressed twice. Since there are very few situations where more than one space is needed, ignoring the second one (within a time frame) could be a solution.

Disable shortcuts.

As a consequence of the situations referred above of users pressing the space key two times in a row, some keyboards had a shortcut that introduced a '.', which led to texts written as: "Olá! Hoje. Faço. Anos (…)". In addition, on the Android keyboard, a slightly long press on the top row of letters switched to a number. To novice users who press for longer than needed, this introduces errors. Therefore, this sort of shortcuts should be avoided.

6 ANNEX

6.1 Instructions to participants

6.1.1 Introduction

First of all, I would also like to ask you to read and please sign this informed consent form.
Our goal now is to evaluate smartphone on-screen keyboards. In order to achieve that I will ask you to write a specific text on several smartphones. You can ask me about anything that you don't understand and then you can try to accomplish the task. Try to do it as if I was not here but if you feel that you are stuck you can ask me for assistance. You can also voice your opinions. Remember that we are testing the application and not the user, and that there is no right or wrong way to perform a task. Also, we are looking for both good or bad feedbacks so don't refrain from expressing a bad opinion or point out any errors that you may encounter. They are expected and we appreciate it if you let us know.

6.1.2 After the test

Do you have any questions or comments? Thank you very much for your participation in this test, your opinion is very valuable to us.

6.2 Informed consent (Portuguese)

A *Associação Fraunhofer Portugal Research* faz trabalho de investigação destinado a encontrar soluções que promovam o bem-estar da população.

No âmbito do projecto GoLivePhone pretendemos avaliar a usabilidade de diferentes teclados tácteis de forma a desenvolver uma solução adaptada às necessidades dos utilizadores finais. Para o estudo, iremos proceder à recolha de dados relativos ao processo de usabilidade de diferentes teclados, através de gravação vídeo da interacção dos participantes com diferentes smartphones.

Gostaríamos de contar com a sua participação nesta fase da nossa investigação. A participação não envolve qualquer prejuízo ou dano material e não haverá lugar a qualquer pagamento. Os dados recolhidos são confidenciais. A *Associação Fraunhofer Portugal Research* tomará todas as medidas necessárias à salvaguarda e protecção dos dados recolhidos por forma a evitar que venham a ser acedidos por terceiros não autorizados.

A sua participação é voluntária, podendo em qualquer altura cessá-la sem qualquer tipo de consequência.

Agradecemos muito o seu contributo, fundamental para a nossa investigação!

O participante:

Declaro ter lido e compreendido este documento, bem como as informações verbais fornecidas e aceito participar nesta investigação. Permito a utilização dos dados que forneço de forma voluntária, confiando em que apenas serão utilizados para investigação e com as garantias de confidencialidade e anonimato que me são dadas pelo investigador. Autorizo a comunicação de dados de forma anónima a outras entidades que estabeleçam parceria com a Associação Fraunhofer Portugal Research para fins académicos e de investigação científica.

Nome: ____

Assinatura: ____ Data ___ / ___ / ______

Investigador responsável pelo projecto "GoLivePhone":

Nome:

Telefone:

E-mail:

6.3 Detailed results by participant

Participant 1

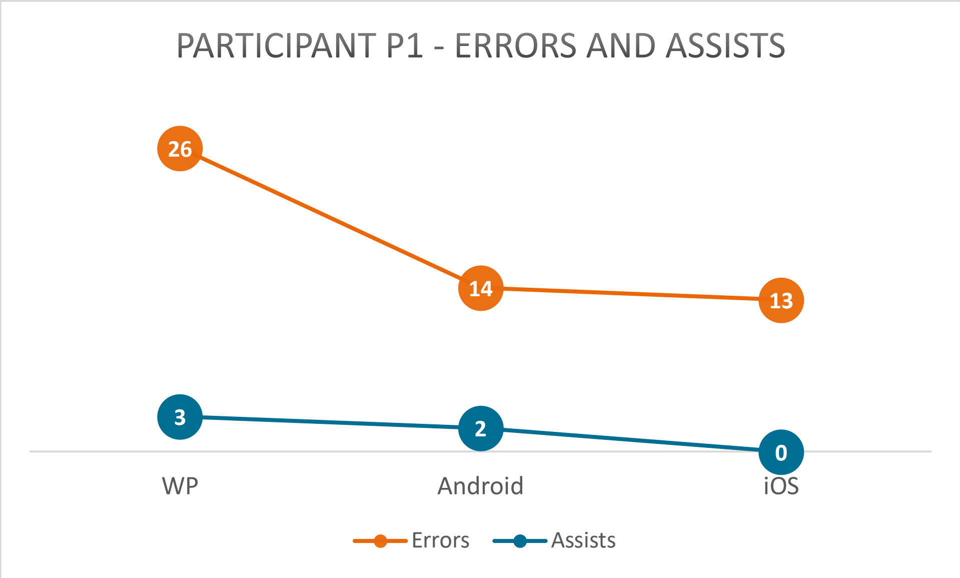

Figure 11. P1 - Errors and assists

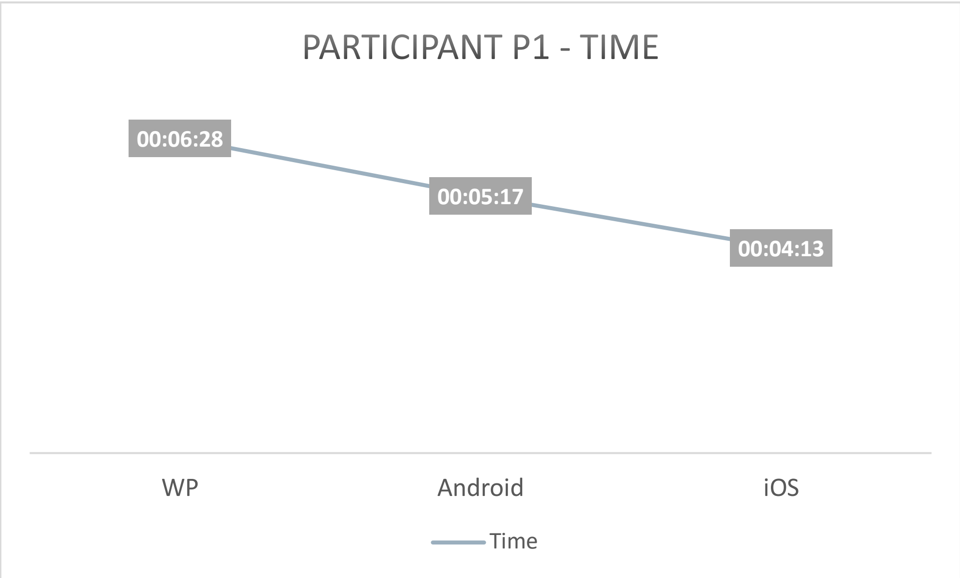

Figure 12. P1 - Task time

Participant 2

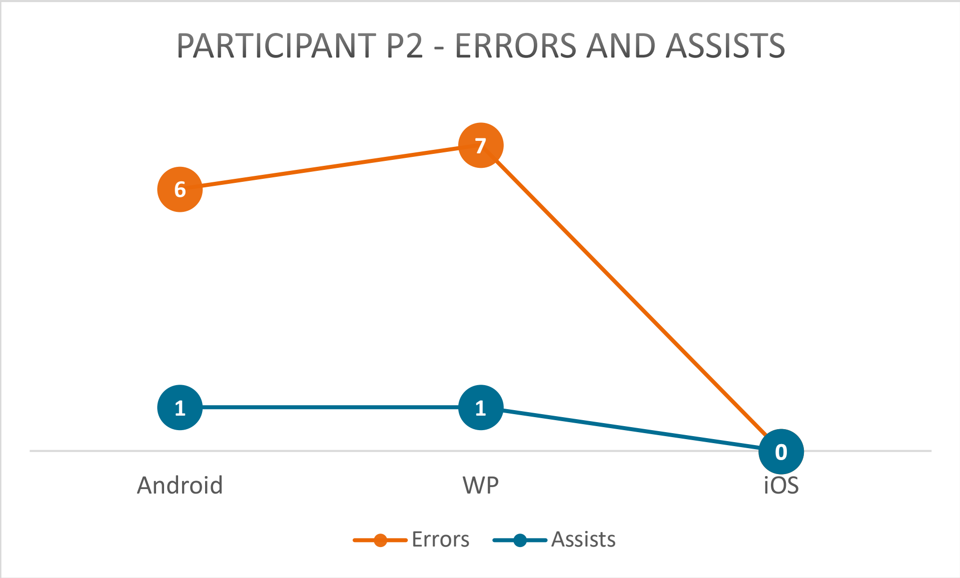

Figure 13. P1 - Errors and assists

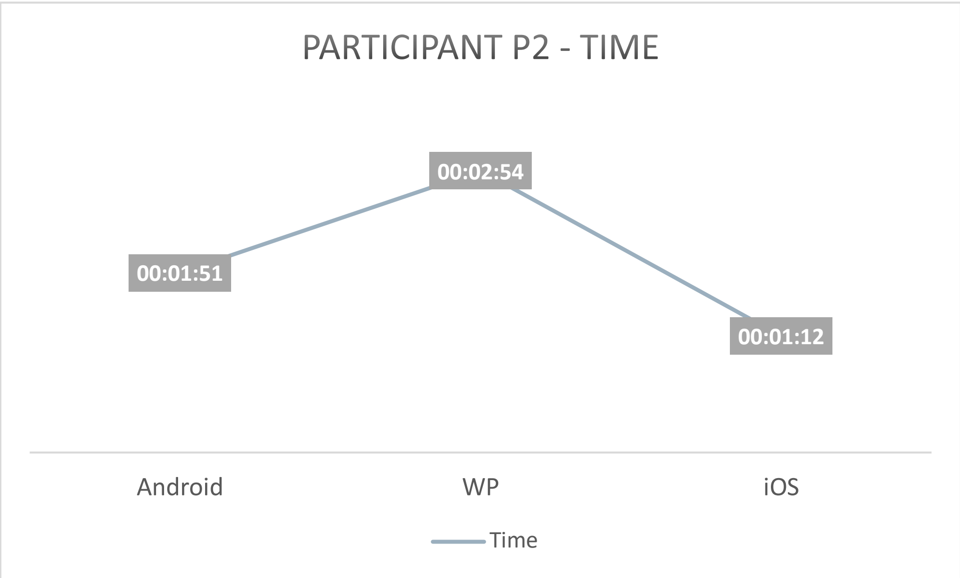

Figure 14. P2 - Task time

Participant 3

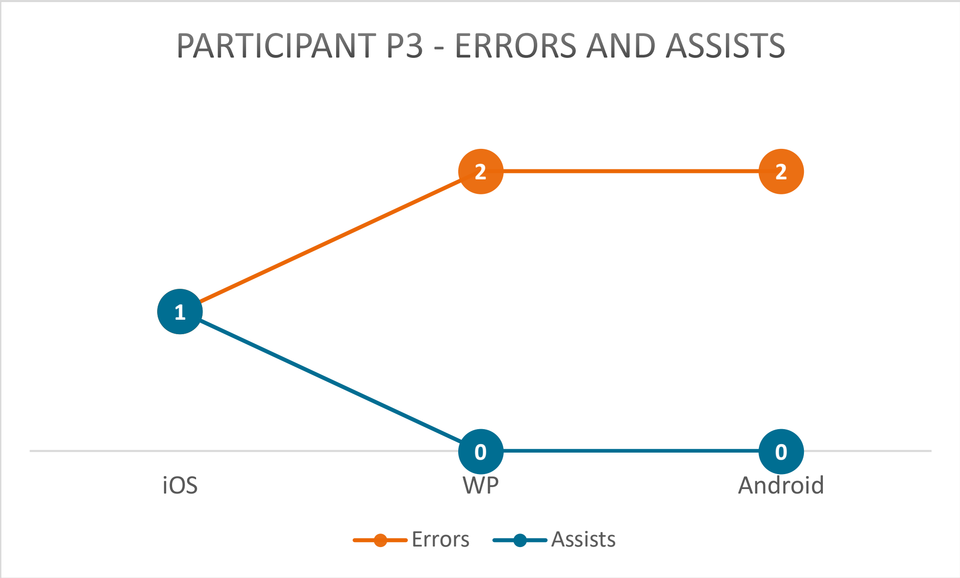

Figure 15. P3 - Errors and assists

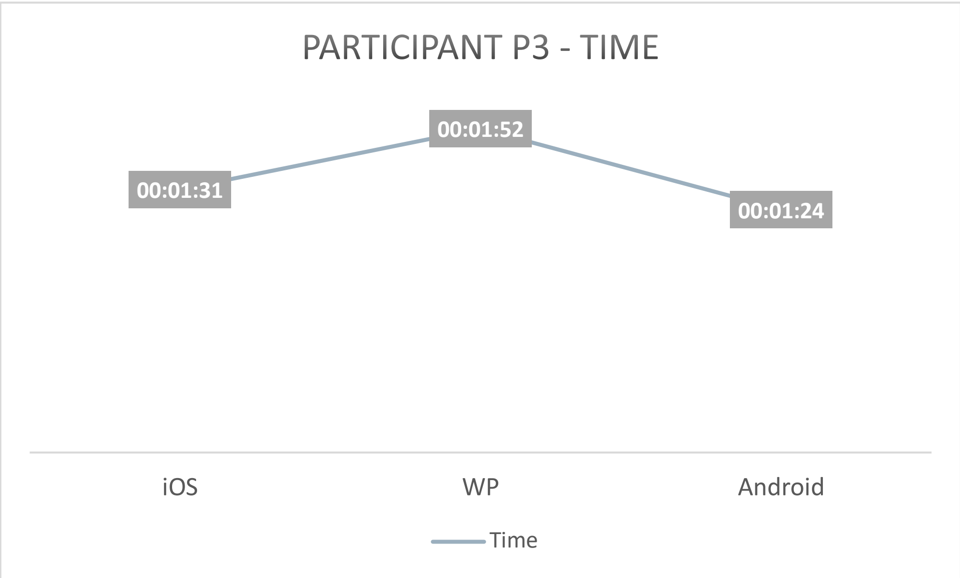

Figure 16. P3 - Task time

Participant 4

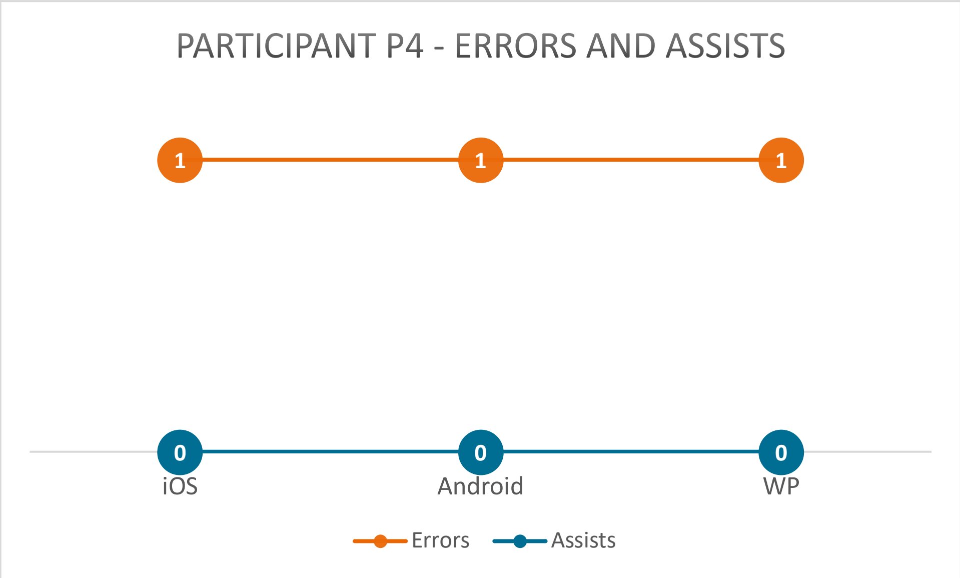

Figure 17. P4 - Errors and assists

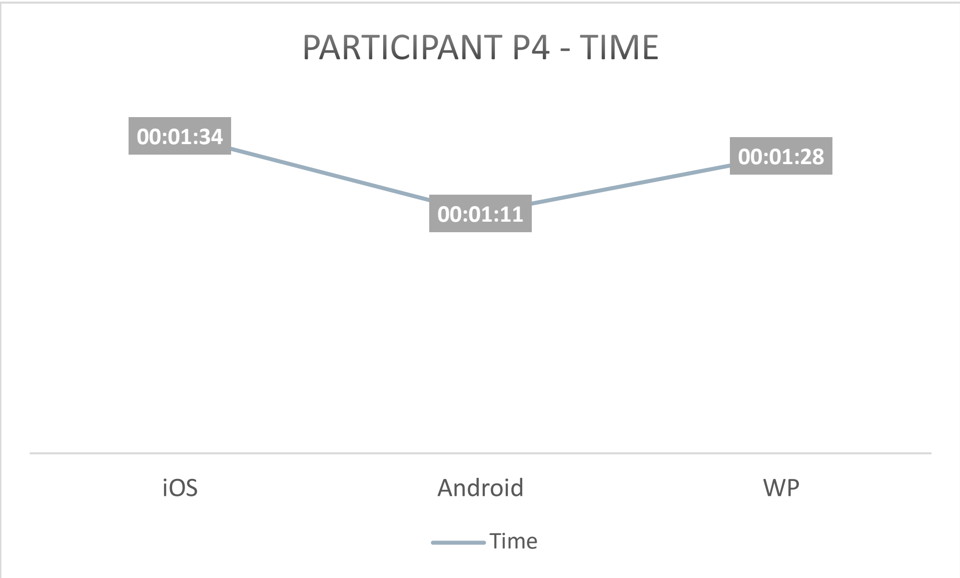

Figure 18. P4 - Task time

Participant 5

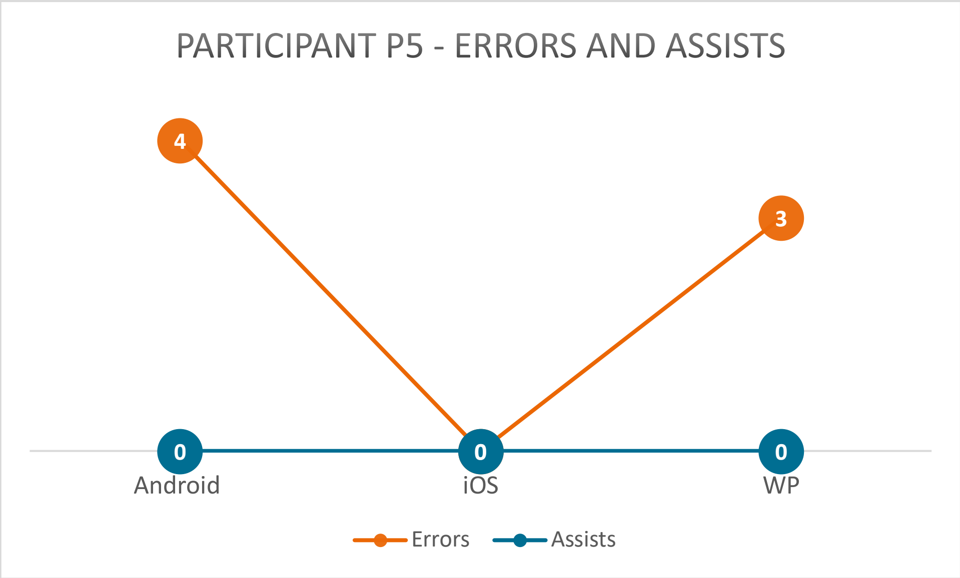

Figure 19. P5 - Errors and assists

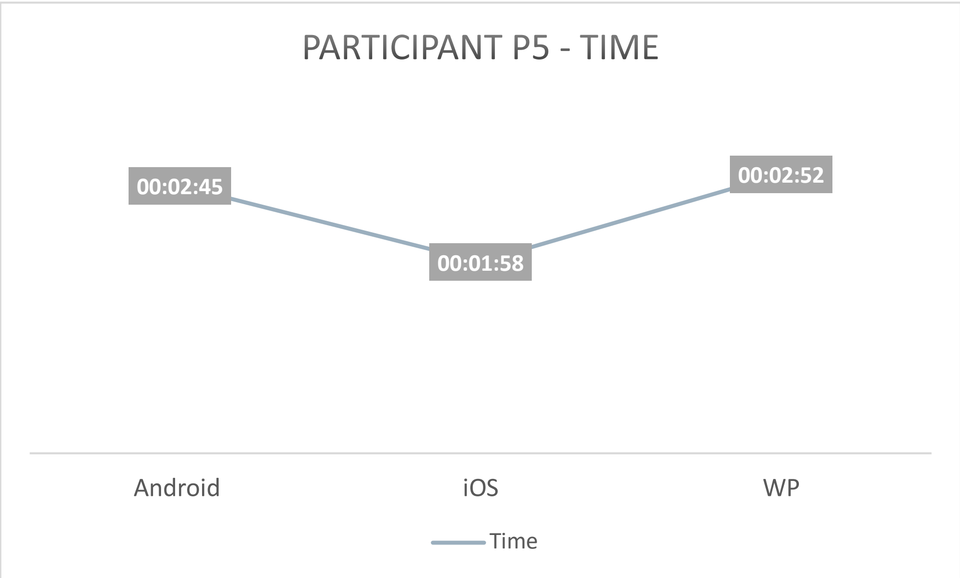

Figure 20. P5 - Task time

Participant 6

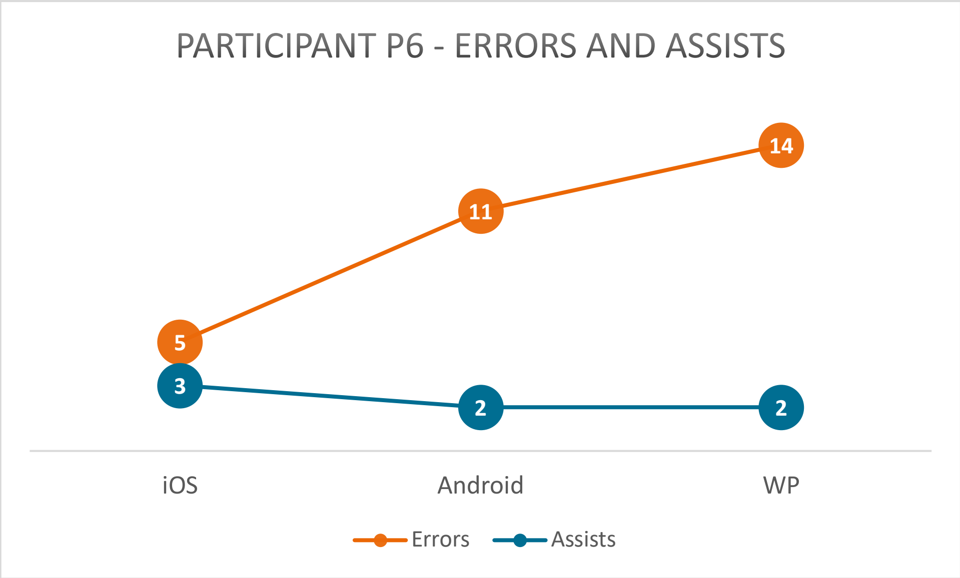

Figure 21. P6- Errors and assists

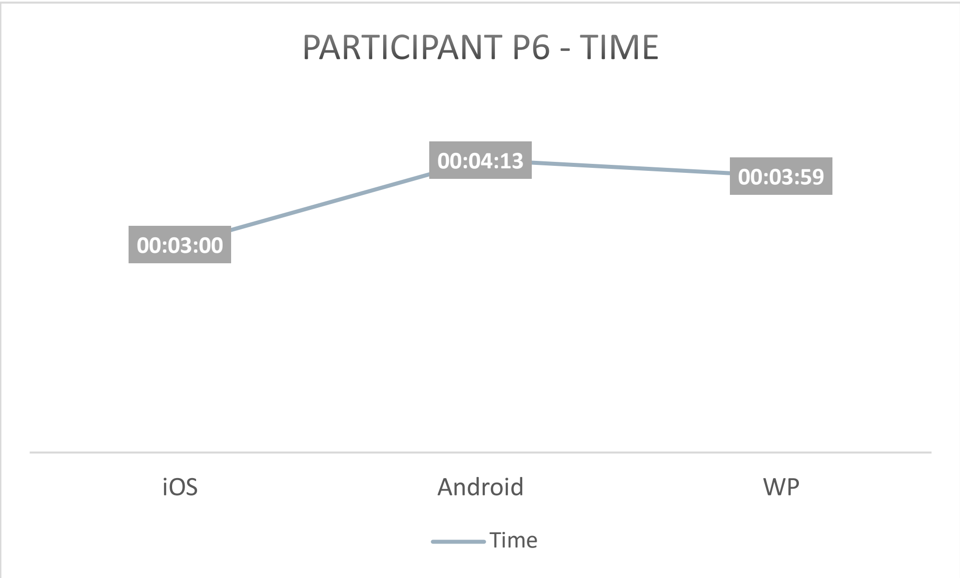

Figure 22. P6 - Task time

Participant 7

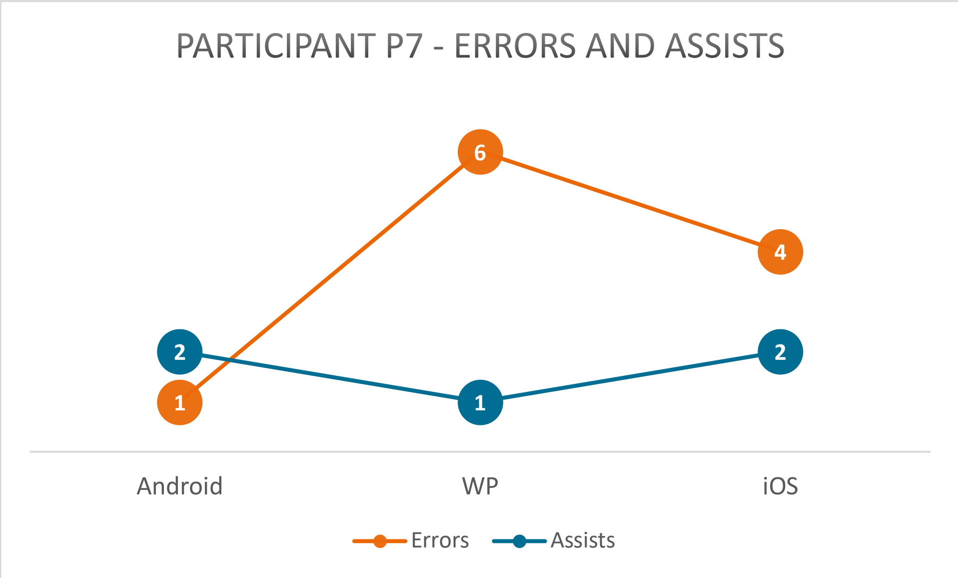

Figure 23. P7- Errors and assists

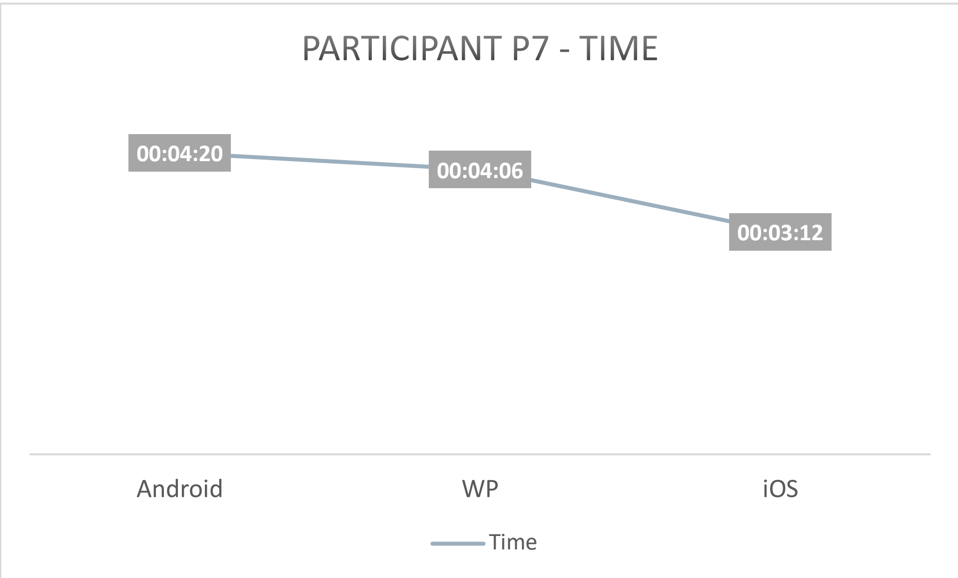

Figure 24. P7 - Task time

Participant 8

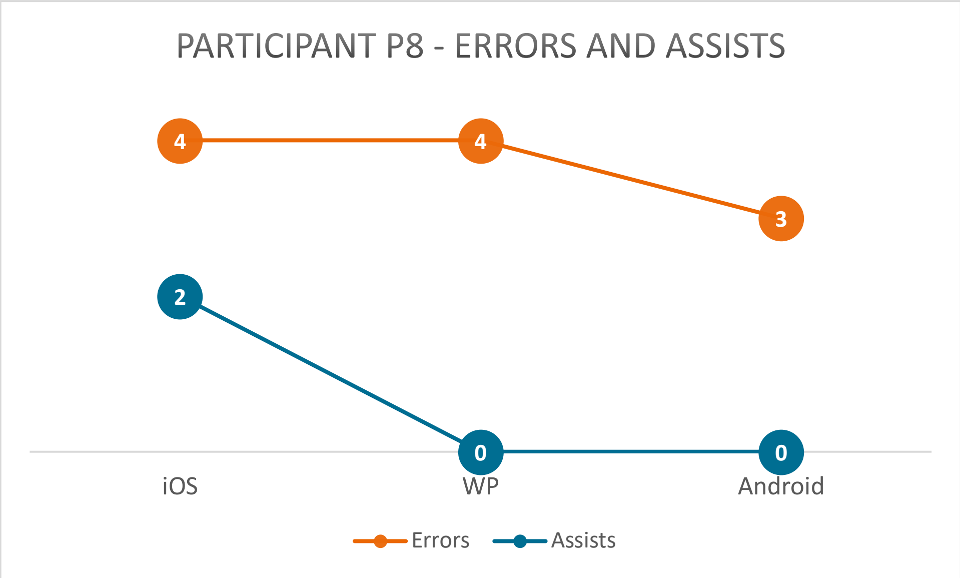

Figure 25. P8 - Errors and assists

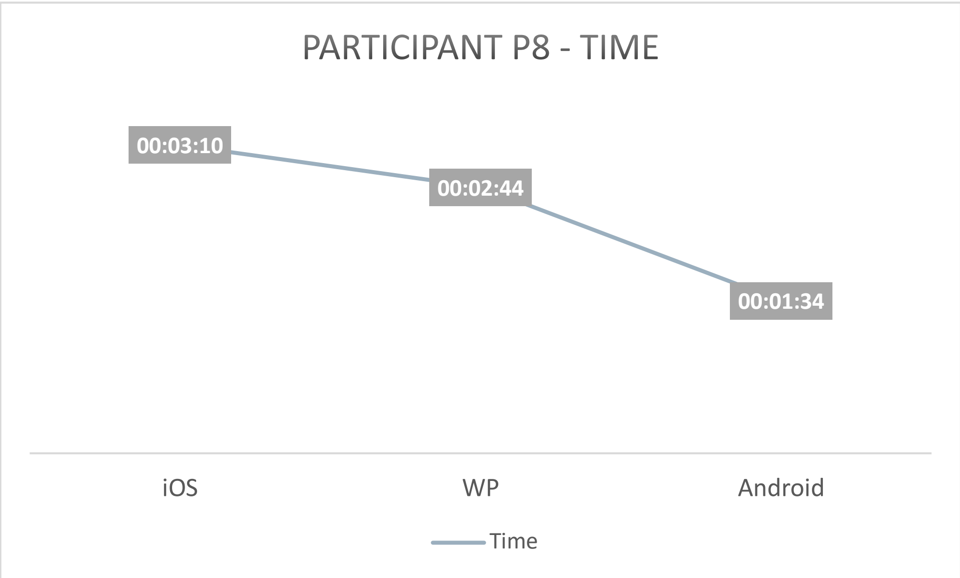

Figure 26. P8 - Task time

Participant 9

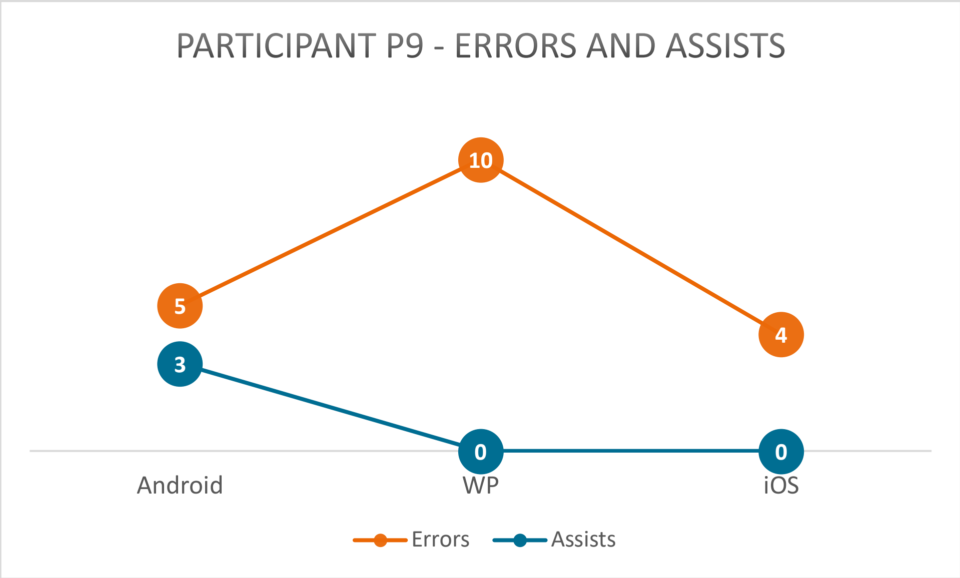

Figure 27. P9 - Errors and assists

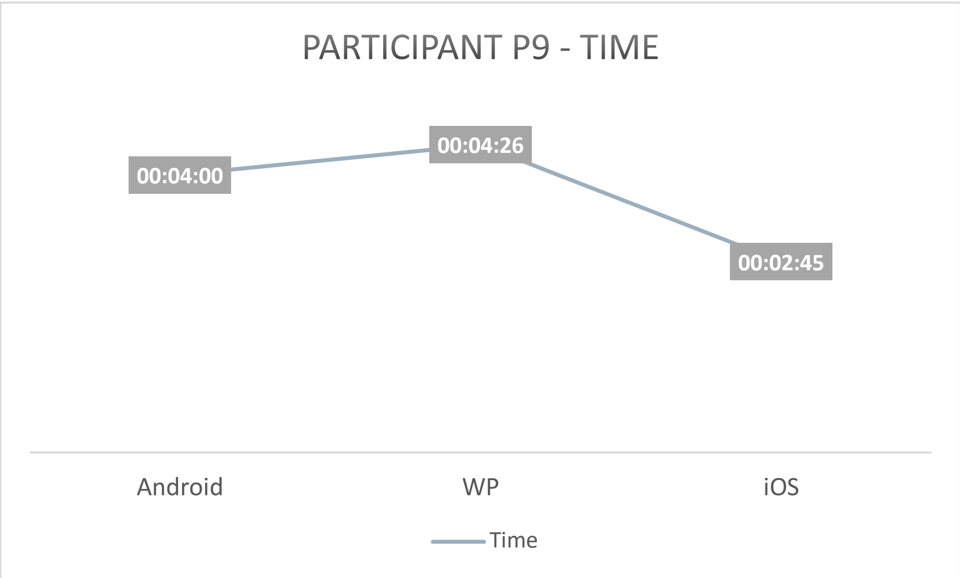

Figure 28. P9 - Task time

Participant 10

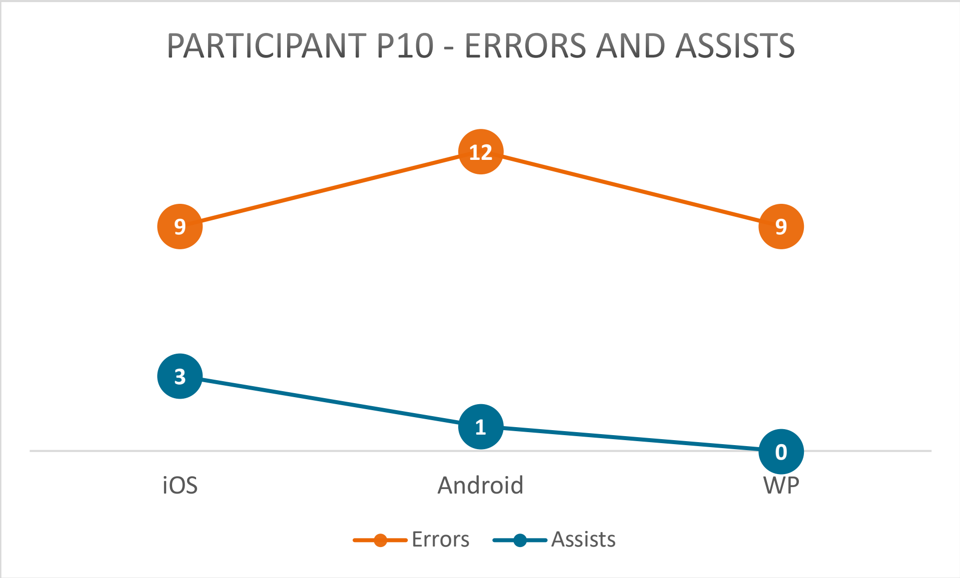

Figure 29. P10 - Errors and assists

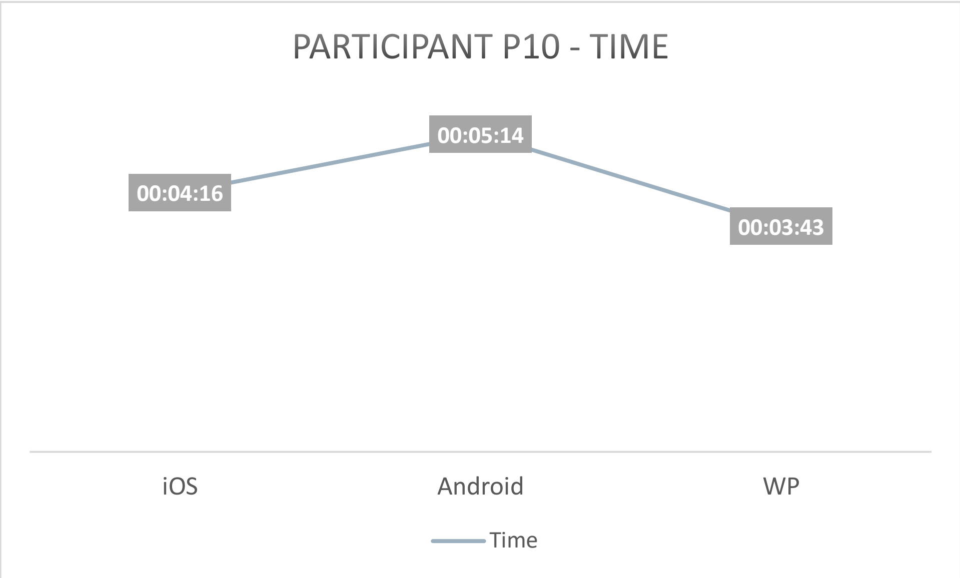

Figure 30. P10 - Task time

Participant 11

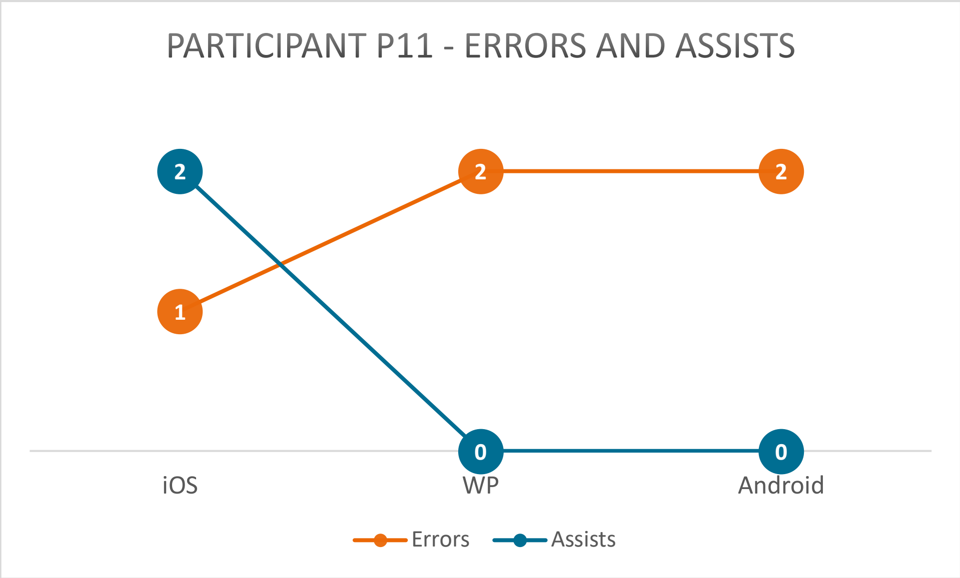

Figure 31. P11 - Errors and assists

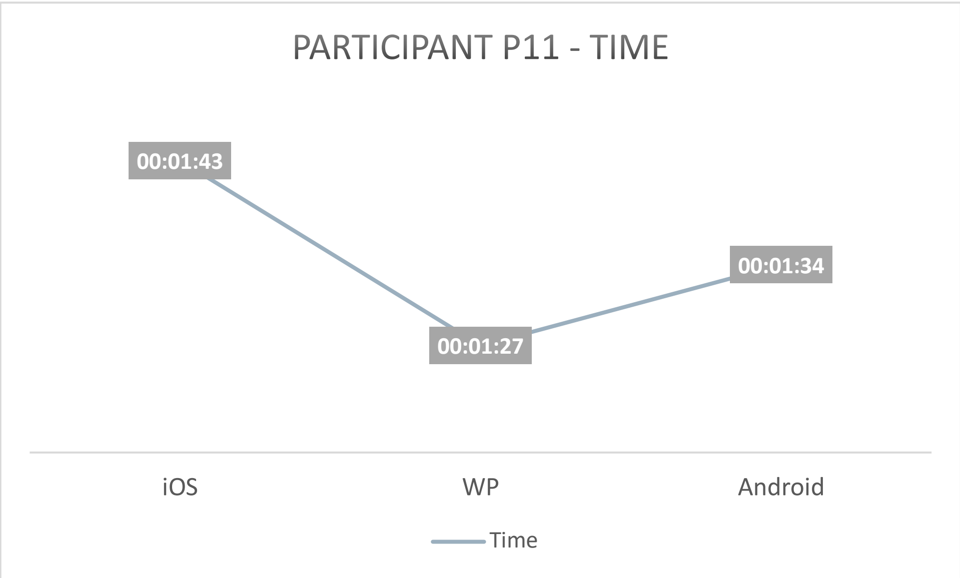

Figure 32. P11 - Task time

Participant 12

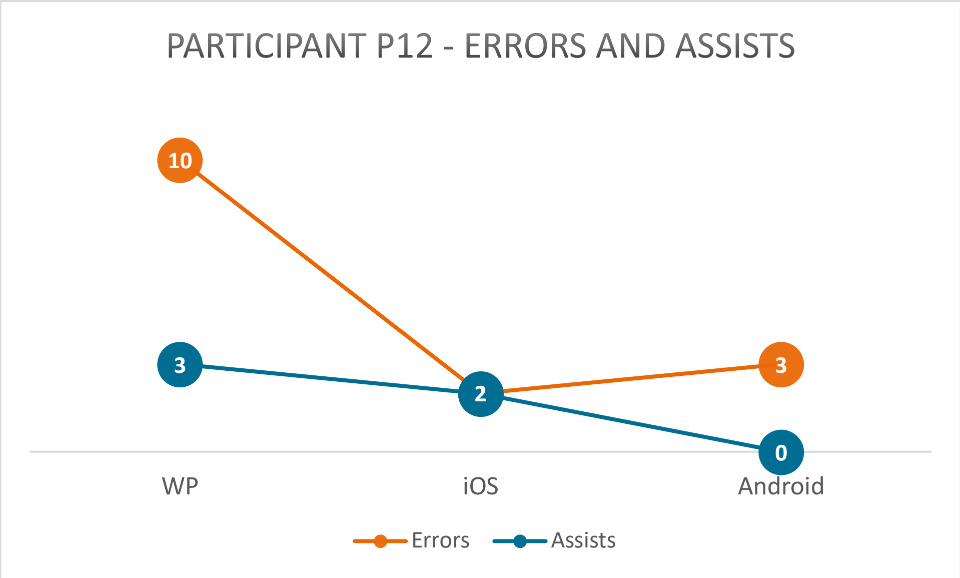

Figure 33. P12 - Errors and assists

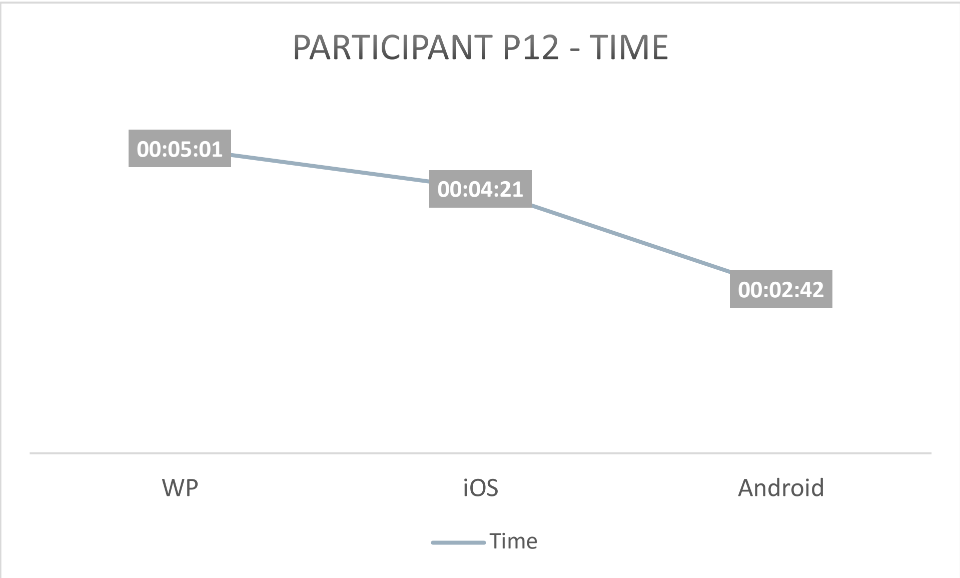

Figure 34. P12 - Task time

Participant 13

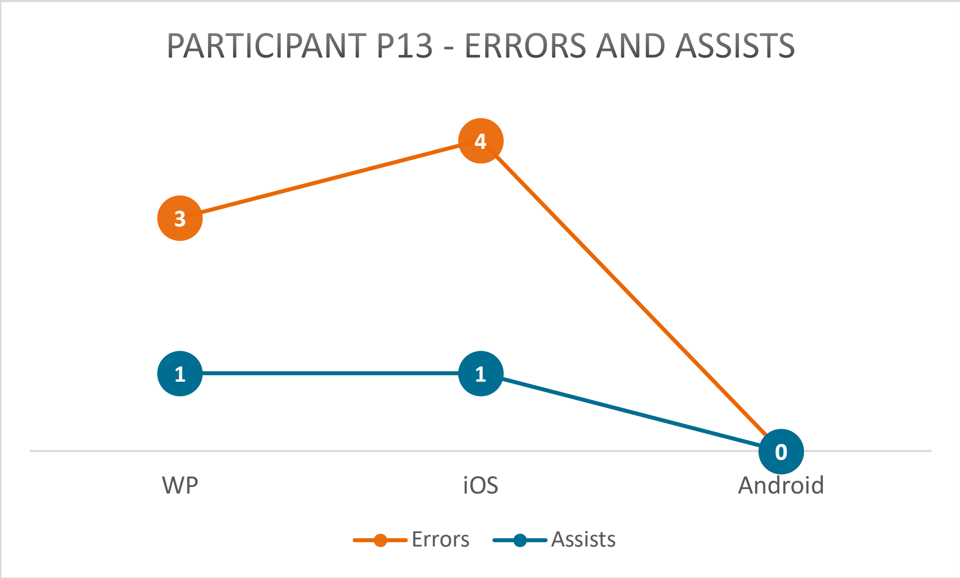

Figure 35. P13 - Errors and assists

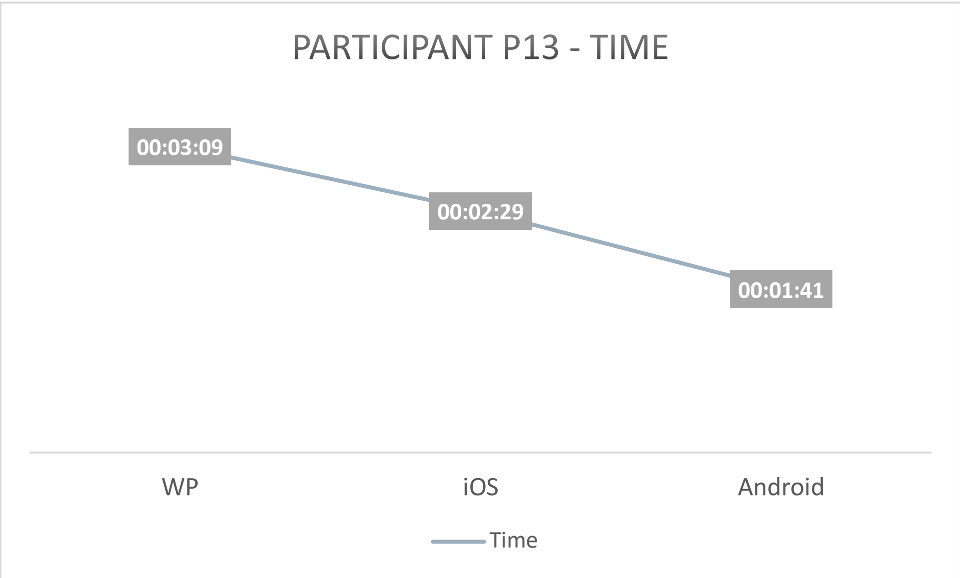

Figure 36. P13 - Task time