Large Scale Collaborative Project

7th Framework Programme

INFSO-ICT 224067

Report on Usability Benchmarking

Deliverable n.	D4.8.4	Report on Usability Benchmarking	
Sub Project	SP n.4	Evaluation and Assessment	
Workpackage	WP n.8	Bench-marking from the consumers point of view	
Task n.	T n. 4	Usability Benchmarking	
Author(s)	Pontus Wallgren, Nadja Lejon	File name	D4.8.4_Report on_Usability_benchmarking.docx
Status	Final		
Distribution	Public (PU)		
Issue date	...	Creation date	2012/11/13
Project start and duration	1st of June, 2008 – 54 months		

Project co-funded by the European Commission

DG-Information Society and Media

in the 7th Framework Programme

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	DESCRIPTION
ESoP	the European Statement of Principles
TMC	Traffic Message Channel
P	Participant
CV	Cognitive Walkthrough
HTA	Hierarchical Task Analyses

REVISION CHART AND HISTORY LOG

REV	DATE	AUTHOR	REASON
1	2013-01-15	Pontus Wallgren	Changes suggested by the internal reviewers

EXECUTIVE SUMMARY

It can be assumed that the usability of the devices tested in the TeleFOT FOTs will have an impact on if and how the participants use the devices and therefore on which impact the devices will have on the different impact areas of TeleFOT. Low usability may both lead to dissatisfaction and consequently lower use, and therefore lower impact. Low usability may also infer on safety in that the user have to spend more time and cognitive effort on the device instead of on the primary task of driving. Two principal strategies can be used when analysing usability, heuristic evaluation or usability tests. Four of the devices used in TeleFOT have been evaluated using an heuristic approach, in this case a checklist, and one of the four devices have also been evaluated with usability testing. The conclusion is that the devices used in the FOTs are fairly typical for the product category and that no really severe usability problems were found. All devices, however, have some issues that might affect their usability. Furthermore, it can be concluded that the way one evaluate usability will have great impact on what result one gets.

INTRODUCTION

It could be argued that in order for any nomadic and aftermarket device to have any impact on any of the impact areas TeleFOT is studying, it has to be used. This is discussed in detail in the deliverables on user uptake. Furthermore, one pre-requisite for user uptake is the usability of the device. ISO defines usability as "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." Analysing the definition it becomes clear that usability isn't a property of the device alone, but is created in the use context. If one is to do benchmarking on usability one therefore has to consider not only the device on its own, but in its intended context of use.

Two different methodological approaches have been used in order to evaluate the usability of the tested devices; heuristic methods and usability testing.

There are many different heuristic methods that can be used, the most well known are probably Cognitive walkthrough (CV) and Hierarchical Task Analyses (HTA). Since the analyses by necessity had to be performed by the different test sites, rather than by the Usability experts in TeleFOT together as would have been preferred, the project opted to use a checklist; the "In-Vehicle Information Systems Electronic Checklist v.1.01". The checklist also has the added benefits of ensuring that the context of use would be properly addressed.

This checklist have been used to evaluate the device used in the Spanish LFOT1, the Italian LFOT and the UK LFOT, as well as the device used in the Finnish LFOT and the one used in the Swedish LFOT2. Furthermore it has been used to evaluate the device used in the UK DFOT1. The device used in the Swedish LFOT 2 has also been evaluated in usability tests in order to compare the quality and quantity of usability related issues that can be extracted using the two approaches.

HEURISTIC USABILITY EVALUATION

TeleFOT has used the "In-Vehicle Information Systems Electronic Checklist v.1.01" developed by TRL (available for download at www.trl.co.uk). The checklist covers Installation, Information presentation, Interaction with displays and controls, System Behaviour, and Information about the system. It provides guidelines for design referring to the European Statement of Principles (ESoP), International Standards, Statement of Principles and Regulations. The checklist is made up by 57 question areas with one or several sub-questions to be answered, followed by an assessment of severity and a comments box. The checklist then provides a summary of all issues that the analyst has found in some way problematic. The design of the assessment tool is that the checklist should be filled in after the analyst(s) have tested the device in its intended context of use. The focus of the checklist is clearly on safety and usability issues, mostly on an interaction level. The inclusion of user and the use situation are very much up to the analyst.

As stated earlier four devices have been evaluated using the checklist, the Blom navigator used in the Italian LFOT, Spanish LFOT1, and UK LFOT; the FootLITE device used in one of the UK DFOTs; the LATIS mobile application used in the Finnish LFOT; and the Garmin navigator used in the Swedish LFOT2. The results from the heuristic usability evaluation are presented below.

2.1 LATIS

LATIS™ (provided by Logica), is a Location aware traffic information solution for drivers. It is based on Logica's Enterprise Mobility framework. The system provides both traffic & road weather information as well as speed limit information and speed alerts. Mediamobile Nordic provides the TRAFFIC and road weather information.

LATIS™ utilizes a built-in speech synthesizer to read aloud announcements of nearby incidents or other relevant info. On-line map service is used to display the user's position and the exact location of the incident. Current speed and speed limit are also displayed for users equipped with GPS (as were all phones included in the test). The current speed is read aloud, if it exceeds the speed limit. As the information exchange in LATIS™ works both ways by nature, all users produce also advanced FCD information. Manual "one

button" reporting of traffic incidents enables even a limited number of users to effectively provide traffic information. LATIS™ mobile phone application works also on the background, enabling a simultaneous use of navigator software.

The nomadic device, on which the user interface of the applications is implemented, is the users' own mobile phones. The applications support Nokia Symbian phones (N and E series phones, as well as Nokia 6210 Navigator.

2.1.1 Installation

Due to the system being installed on a mobile phone, all questions regarding installation of the physical device are dependent on the participants' phones and therefore not possible to evaluate. The installation of the software on the phone on the other hand, was quite difficult and it was considered that this was not a task you could do without special skills. Much for this reason, a help desk was installed. Reports from the help desk indicated that with such a variety of mobile phones, the help desk people sometimes had troubles supporting the installation.

2.1.2 Information Presentation

A serious concern was that Traffic info messages are shown both as symbol and texts. On the map view, symbols sometimes overlap each other rendering them unreadable. Another serious concern was that general icons are used for the messages. These icons may not always cover the specific content.

A minor concern was that the mobile phone brightness on most of the phones in the test does not change in function of the surrounding light. There was also no special "Night mode in the software. Another minor concern was that traffic info icons are superposed on a map display, obscuring part of the map.

Another minor concern was the audio messages. These are based on text messages that are read through speech synthesis. Text messages related to traffic info are long texts and therefore sometimes difficult to grasp when presented as speech. Audio messages related to speed limit, however, are short and clear. The service was during the test only available in Finnish, but this was not considered a problem as all participants in the FOT spoke Finnish. Another part of the checklist concerned the loudness of the auditory information. This is, of course, dependent on the phone used, but it was noted that there

generally was no connection between the cars infotainment system and the phone so e.g. the radio wasn't muted when traffic info were read by the device.

During the FOT it was noted that messages, e.g. on road works may be out-dated, when the end of an event has not been registered by the traffic info server. Also, the correctness of speed limits depends on the accuracy at the server, e.g. if speed limits are changed at the correct time between summer/winter.

2.1.3 Interaction with Displays and Controls

There is very little interaction with LATIS. The only menu interaction is for controlling the settings, which should not be done while driving. It was noted that since the software was running on standard Symbian phones, all interaction had to be done through the numeric keypad. This was found to sometimes be not easy or intuitive.

2.1.4 System Behaviour

Normally the driver does not have to interact with the service during driving. The only input possible is traffic events reported by the user.

2.1.5 Information About the System

Installation instructions and first use are rather complex. Apart from that, use is quite simple, except for the entering of traffic info events that is not self-explanatory. This is however not something that the average user is expected to do that often. The guidelines state that program should not be used when driving.

2.1.6 Overall assessment and recommendations

Overall the LATIS service was found to be quite easy to use. The small screen of the devices tested in some cases rendered information difficult to read as icons and text obscured the map. There were also concerns that data input, although rare, was quite difficult.

2.2 BLOM Ndrive Touch XL SE 'TeleFOT Release'

The NDrive G800 is a personal navigation solution based on GPS technology. It provides navigation through visual and voice instructions, which includes names of roads and locations, door-to-door navigation and detailed information about points of interest. Furthermore, is provides speed limit information and speed alert.

2.2.1 Installation

The device is mounted on the windscreen with the help of a suction cup. Depending on where on the windscreen you mount the device and the design of the vehicle it might be difficult to reach the device. In the test vehicle used for the evaluation, the analysts found that the device was difficult to reach when the driver had their seatbelt on. This restricted their ability to interact with the device for necessary tasks such as adjusting the volume. As the device is mounted on the windscreen, the device impaired some of the 'swept windscreen area', reducing the driver's view of the road.

2.2.2 Information Presentation

The 'main menu' button's function was not obvious from its appearance and was, therefore, easy to overlook. The icons at the top right of the map screen were also not intuitively obvious as to what their functions were and were quite cluttered and close together. Furthermore, the interface when giving directions didn't inform you which road you were heading onto e.g. 'take the A6'.

To adjust the volume from the map screen required clicking on the 'battery and GPS' logo which was not at all intuitive, followed by requiring a further button press to make the slider controls appear to allow adjustment of the volume. This would likely be a common interaction with the device and was seen as being unintuitive and overly complex.

It was suggested that presenting auditory information, on which road to enter, would add clarification and allow the driver to match audio instructions to the road signs observed.

2.2.3 Interaction with Displays and Controls

A small number of the turning instructions were observed to be a little late in their presentation when coming off a dual carriageway and onto a roundabout. There was also no easy way to ask the device to repeat a missed audio instruction, leading to a greater reliance on information displayed on the screen and more eyes off road time than necessary. Furthermore, the icons on the map display were quite small and therefore it was difficult not to inadvertently activate another control. It was also not possible to locate the controls non-visually due to the interface being a touch screen, meaning the buttons weren't differentiable by texture or relief.

2.2.4 System Behaviour

A serious concern was raised as the device didn't inform the user that the driver shouldn't interact with the device whilst the vehicle was in motion. It also doesn't inform them that they take responsibility for their own actions when interacting with the device. The user manual also lacks a statement to inform the driver that they retain responsibility for complying with traffic regulations.

Although the instruction manual informed users not to attempt to use the system whilst in motion the device was still fully operational whilst the vehicle was moving, perhaps offering a temptation to interact with the device.

2.2.5 Information about the system

The devise was delivered with a manual in the relevant languages.

2.2.6 Overall assessment and Recommendations

Overall the device didn't perform as well as many of the class leaders but was generally usable and navigated the researchers from point A to B with only very minimal confusion caused. It should be acknowledged that the device was still in prototype form and it is likely that the only serious concerns raised by the assessment, that the device itself did not inform users not to interact with it whilst in motion, would likely have been addressed in the finished product.

2.3 FootLite

FootLite is a system that captures, processes and displays information about journeys and driving style through a smart phone application. The information presented ranges from Green Driving Support to lane keeping assistance and headway keeping assistance. With web access, drivers are able to see post-journey feedback for a comprehensive review of driving behaviour and also to share data with the FootLite on-line community.

2.3.1 Installation

The installation is dependent on the smart phone used and have not been analysed. The smartphone connects wireless to the camera (for lane keeping and headway assistance) and a special FootLite interface to the car.

2.3.2 Information Presentation

There was a slight concern of the use of green and red colours together, however, issues with this were mostly overcome by the graphics in question being distinguished by positioning as well as colour. The graphic informing where the gearshift information is displayed currently uses the word 'gearchange' this was found to be difficult to read due to the font being vertical. Under very exceptional situations the display can become a little overloaded, such as when overtaking on single carriageway, where lane departure, acceleration and headway warnings were all shown simultaneously.

2.3.3 Interaction with Displays and Controls

The system requires no interaction once it has been launched.

2.3.4 System Behaviour

It was suggested by the analysts that adjusting any of the system's settings, apart from volume, should be made impossible once the vehicle is in motion. Furthermore, they concluded that if the volume is switched off or turned down low then this information should be made displayed to the user. Also if the system crashes or stops working then it should be made obvious to the user that the system is no longer functioning.

2.3.5 Information About the System

The FootLite system, while base on mature technology, is still in the prototype phase. There still don’t exist much printed material on the system.

Overall assessment and Recommendations

Overall there were only a small number of minor concerns with the design. The interface design was thought to show clear, and intuitive graphics. Overall the design was thought to be intuitive, believed to cause minimal distraction and effectively communicated greener driving as well as improved driver safety practices.

There were a few design recommendations. The first has already been mentioned in the report above, namely to make it impossible to interact with the device's settings while the vehicle is in motion as well as displaying if the volume is turned down low or turned off. The graphic informing where the gearshift information is displayed is suggested to be replaced with a simple image of a gearbox to alleviate the problem of it currently being vertical. Prioritizing and reducing the information displayed in circumstances such as overtaking would improve the design.

2.4 Garmin Nüvi 205WT

The Garmin Nüvi 205WT is a fairly standard GPS Navigator. In addition to navigation it provides traffic information through TMC, and a simple green driving support function called Garmin EcoRoute. It is mounted in the windscreen with a suction cup.

2.4.1 Installation

The device is mounted on the windscreen with a suction cup, power is fed from the cigarette lighter socket. No problems regarding installation could be conceived.

2.4.2 Information Presentation

One concern regarding information presentation was that many common settings are hidden in the menu structure and it is not obvious how to access e.g. Traffic information.

Another concern was that "take the third exit to the right" is used as a way to say "turn left in the roundabout". This may take some time to get used to. During the test it was apparent that Traffic info is often late and only covers some traffic info, or even worse is irrelevant. Furthermore, when traffic info is available the traffic info sign changes colour, which isn't very obvious. Green driving advice is presented as a number from 1-100 and although quite easy to see and interpret was seldom considered relevant as an indication of green driving.

2.4.3 Interaction with Displays and Controls

Some things, such as traffic info requires lots of button pressing in order to see what type of problem there is and exactly where it is located.

2.4.4 System Behaviour

Although the instruction manual informed users not to attempt to use the system whilst in motion the device was still fully operational whilst the vehicle was moving, some functions (TI) requires interaction while driving. The display of regulatory messages on start-up infers on the usability criteria "efficiency", and may lead to people starting driving and then entering information as the car is moving.

2.4.5 Information About the System

The manual did not include information on “ECOROUTE” the simple green driving support function included in the device. These instructions were in a supplement only available in English

2.4.6 Overall assessment and Recommendations

The device is generally quite well thought out. Navigation is clearly the primary function and is given priority over the other functions. Traffic information and Green driving support have more usability related issues.

USABILITY TESTING

One of the devices tested in TeleFOT, the Garmin Nüvi 205WT, was in addition to the heuristic evaluation, subjected to a Usability test. In a usability test, realistic tasks are created. The subjects then have to try to solve these tasks using the tested device. The work starts with deciding on relevant usability criteria, such as efficiency, learnability, gueassability, etc. (Jordan 1998)[1]. These criteria are then matched to measures such as time to complete a task/scenario, number of interactions, number of erroneous interaction, stated certainty, etc. The result is a good picture of how the usability of a device is in a certain context. Much effort has to be spent in order to assure that usability as it is measured is in line with usability as it will be experienced by the real users in a real use situation.

3.1 Method

3.1.1 Scenarios

Scenarios were composed to cover the most common functions, as well as some scenarios that comprised the set up of the device. The scenarios were selected in such a way that they represented situations similar to real circumstances using the device. The scenarios were created to mainly evaluate the usability criteria gueassability and learnability, as these were thought to be the most relevant for this type of device. Through the comments and by watching the participants navigate through the system, information about Garmin Nüvi's consistency, affordance, visibility, feedback and compatibility was also gained.

3.1.2 Subjective analyses

[1] Jordan, P., (1998), *An Introduction to Usability*

Between each scenario the participants got to answer to which extent they were certain that they had done the task correctly and to which degree they found the task difficult on a seven degree scale. In the scale 1= ”not at all sure”, 4= ”neither sure or unsure” and 7= ”completely sure” respectively 1= ”very difficult”, 4= ”neither difficult or easy” and 7= ”very easy”. After the test they were asked about their general impression of the navigator and what they thought of finding different functions.

3.1.3 Objective analyses

During the test the participants were filmed so that the number of clicks used to execute a task could be counted. They were also timed. If the participants asked for help, this was noted together with what information he/she got.

3.2 Test conduction

The usability test of Garmin Nüvi was set out in two different tests. One part was set out in a car and the other in a lab environment. The test participants varied in age and gender as well as their experience of using navigators. The tests aimed to include all major functions as well as some more seldom used ones.

3.2.1 Participants

The test participants were chosen to form a varied test group as the device is intended to be used by most people of legal age to drive. The age varied between 23 to 69 years, and the aim was to have an equal distribution between men and women with 4 women of 10 in the first test and 5 women of 8 in the second test. Differences between the two groups are mainly the previous experience of GPS and navigators where the participants in the car test had more experience than the group set in lab.

Table 1: Test participants for test in car

Test participant	*Age*	*Gender*	*Took drivers licence*	*Experience of GPS*
a	69	Female	1961	No
b	24	Male	2006	Yes
c	61	Female	1971	Yes
d	23	Female	2012	Yes
e	53	Male	1977	Yes
f	38	Male	1992	Yes
g	24	Male	2006	Yes
h	67	Female	1967	Yes
i	51	Male	1978	Yes
j	28	Male	2007	Yes

Table 2: Test participants for test in lab

Test participant	*Age*	*Gender*	*Took drivers licence*	*Experience of GPS*
1	37	Male	1994	Yes
2	25	Male	2007	No
3	46	Female	1984	No
4	26	Female	-	Yes
5	57	Female	1973	Yes
6	27	Female	2003	No
7	48	Male	1982	Yes
8	67	Female	-	No

3.2.2 Test in car

The Garmin Nüvi was tested in a car with 10 test participants. All manipulations of the device were done with the car standing still. The scenarios used took the participants for a shorter trip in the Gothenburg area with two stops on the way.

3.2.3 Test in lab

The Garmin Nüvi was tested in a lab environment with 8 test participants. The same scenarios as in the test conducted in car were used although, for obvious reasons, there was no actual driving taking place. Since there was no car driving these tests could be performed in a shorter timespan, which opened up for the possibility to better evaluate the device's learnability. In order to do so, the test was doubled with the same tasks being performed twice but with different locations and addresses.

3.2.4 Test composition

The functions that were meant to be used were:

- ecoRouteTM - Vehicle Profile and Fuel Report and Driving Challenge
- Set a location – Either: Go Home and use my current location or Where Am I and Save Location
- Set a destination – Where to? And Address
- Set a via point – Where to? And search for extra stop and touch Add as Via Point
- Set avoidances of the route – Tools, settings, navigation, avoidances, enabled.
- Set a favourite or home location to a destination

3.2.5 Scenarios

The scenarios below are the ones that the test participants used in order to solve the tasks. The scenarios were read out aloud by the test leader and repeated if needed. If there was a problem of spelling street names the test participants got to see the name in writing form.

The scenarios used in the car trials and in the first group of lab scenarios were:

1. You are at a car rental firm and have just rented a car for the weekend, as you will help a friend move this weekend. You are an environmentalist and want to be aware of how much fuel has been consumed when you have finished moving and also get indications of whether you drive ecofriendly. The rental firm told you that the car needs 12 litres/100 km in the city and 7 litres/100 km on the highway.
2. The car rental firm is located in an area where you do not find your way so well and you want to be sure to find your way back there quickly and easily when you return the car. You therefore install the device so you can find your way back.
3. Your friend lives on Framnäsgatan 35. You cannot find your way there from the car rental company, and therefore set the destination with the navigator.
4. Since you are going to move some large pieces of furniture you have rented a trailer at Shell on Helmutsrogatan. You want to go there and get the trailer on the way to your friend.

5. Since you will drive an unbraked trailer you want to be sure that you will not go on the highway, or having to make u-turns. Set your device and start driving.
6. At last, the move is finished and it is time to return the car at the rental company. Drive back as instructed by the navigator.
7. In order to climate compensate you want to see how much gasoline has been used during your travels.

Scenarios used in part 2 of the lab test:

1. You borrow a car of a friend to go on a weekend trip. You are still an environmentalist and want to be aware of how much fuel has been consumed when you have finished moving and also get indications of whether you drive ecofriendly. The friend told you that the car needs 13 litres/100 km in the city and 8 litres/100 km on the highway.
2. Your friend lives in an area where you do not find your way so well and you want to be sure to find your way back there quickly and easily when you return the car. You therefore install the device so you can find your way back.
3. You are going to a spa in Lundsbrunns Kurort and on their web page you have found information on how to find your way there. Use the information you need and set the destination with the navigator.

   Lundsbrunn is situated between Skara and Götene In the heart of West Sweden – about a km away from the E20 (three different exits). The routes to the spa is clearly marked with road signs, both form E20 and route 44 between Lidköping and Götene. From Gothenburg the journey is about one hour and a half and from Stockholm approximately four hoursr.

   Address:□Lundsbrunns Kurort Brunnsvägen 32□533 72 LUNDSBRUNN Latitud: 58°28'26.39"N

   Longitud: 13°26'49.79"O

4. It is a long way to go and you have planned to stay and eat on your way. You have been recommended restaurant Sushi Yaki in Skövde and want to go there on your way to the spa. The adress is Kyrkogatan 20, 54130 Skövde.

5. As you are in no rush, you would rather drive fuel efficient than fast and avoid traffic as much as possible. Set the navigator to that.
6. At last it is time to return the car to the friend. Drive back as instructed by the navigator.
7. In order to climate compensate you want to see how much gasoline has been used during your travels.

While the two scenarios number 5 looks quite different, they actually are very similar requiring the participant to go to the same settings page, but tick another box.

3.3 Results

What can be seen in the tests is that the time and number of clicks are similar between the test participants in car and the result from part 1 in lab environment and that both time and number of clicks are reduced in part 2 compared to part 1. This implies that for a first time user it takes more time and energy to look for the right functions and their place in the menu than if you have worked a bit with the device. This means that the system has a high degree of learnability and in some cases low gueassability. In the test set in lab environment, people found the tasks easier to do in part 2 than in part 1, which implies that they learn where to find certain functions and feel more confident. The most common mistakes done by first time users are listed below.

Table 3: Common mistakes and their roots

Common mistakes when first using Garmin Nüvi	Possible reasons for mistake
Try to use the arrows as delete button instead of the eraser symbol	Comparability, as delete buttons on computers and phones are displayed as a left pointing arrow
Write the wrong letters repeatedly	Visual clarity, the buttons are too small for people to aim right.
Try to scroll on the page by stroking the finger over the screen instead of using arrows	Comparability, as most people have experience from smartphones where this way of scrolling is used
Try to set Via Point by pressing Detour firstly and then Extra	Gueassability, as most people expect to be able to write an address after pressing detour. The next place they expect to write an address is after pressing extra.
Try to set the fuel data via Fuel Report instead of Vehicle Profile	Comparability, people expect to both find the information and change it in the same place
Writing the street name instead of city when first writing the destination.	Affordance, in Swedish the word for City (ort) can be interpreted to destination, which could be a cause for this confusion

The average time and number of clicks to conclude the tasks of the scenarios are shown in the table below. A division is done between part 1 and part 2 of the test set in lab to show the difference more clearly.

Table 4: Average time and number of clicks per Scenario

	Test in car		Test in lab			
			Part 1		Part 2	
	Time	*Clicks*	*Time*	*Clicks*	*Time*	*Clicks*
Scenario 1	02:41	33	03:13	37	00:50	16
Scenario 2	01:07	21	01:27	20	01:07	25
Scenario 3	01:15	32	01:54	50	01:11	26
Scenario 4	03:03	49	03:00	56	01:56	44
Scenario 5	02:24	29	02:15	31	01:49	27
Scenario 6	00:42	9	00:59	12	00:19	7
Scenario 7	00:46	5	00:28	6	00:08	4

Table 5: Average rating per scenario

	Sure of doing right			Level of difficulty			Failed task		
	Car	Part 1	Part 2	Car	Part 1	Part 2	Car	Part 1	Part 2
Scenario 1	4,4	4,75	6,9	4,3	3,5	5,3	f,h	2,5,7,8	1,2,7,8
Scenario 2	5,4	5,75	6,5	4,7	4,5	5,1		8	3,7,8
Scenario 3	6,9	6,1	6,9	5,2	4,9	6			8
Scenario 4	4,2	4,4	5,9	3,3	2,6	4,1	b,c,d,e,f,g,h,i,j,	1,5,6	1,7,8
Scenario 5	6,5	5,9	6	3,6	4	3,7		7,8	5,7,8
Scenario 6	6,5	6,1	6,8	5,7	5,3	5,5	a,d	5,7,8	3,7,8
Scenario 7	6,8	6,1	7	5,9	5,1	6,8		8	3,8

After each task the test participants answered how sure they were that they had done right and the level of difficulty. In the scale 1= ”not at all sure”, 4= ”neither sure or unsure” and 7= ”completely sure” respectively 1= ”very difficult”, 4= ”neither difficult or easy” and 7= ”very easy”. Of the total number of scenarios (182)[2], 28 were not correctly executed (not including interrupted tests). In 11 of these 28 cases (40%), the participants were certain that they had executed the task in the correct way. That 40% of the people that did not complete a task thought they completed the task correctly

[2] (10 participants x 7 scenarios + 8 participants x 2 x 7 scenarios) = 182

might mean that the feedback is not sufficient or that they interpreted the task differently than what the test leader intended.

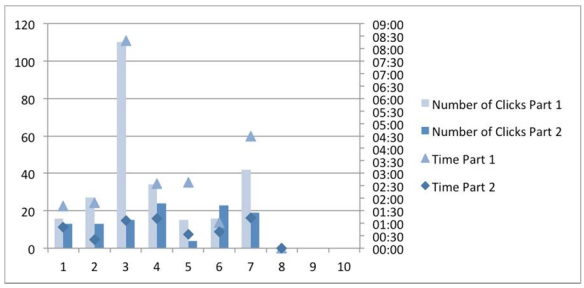
3.3.1 Scenarios

The results from the different scenarios will be shown in the paragraphs below, scenario by scenario, explaining the similarities and differences of time and number of clicks needed to fulfil the tasks.

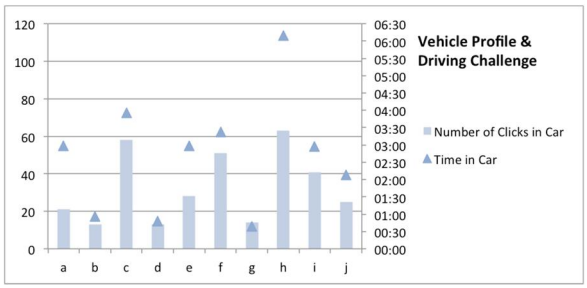
Scenario 1

In scenario 1 the task is to set the fuel economy for the car as well as turn on the driving challenge. Not everyone turned on the driver challenge and this can likely be explained by it being the first question and that there was a lot of information to take in. The scenario does not specifically ask for the driver challenge but instead that they want information of how well they drive while they drive. As seen in the charts below the time and number of clicks needed varies a lot for the first time users (light blue) whereas in the second try (dark blue) both time and number of clicks needed to complete the task are largely reduced to a third of the time implying learnability.

Here recognisability is seen in the system because people see the ecoRoute symbol and associate it with ecology and environment friendliness.



P 5,7,and 8 did not manage the scenario the first time (P=Participant)



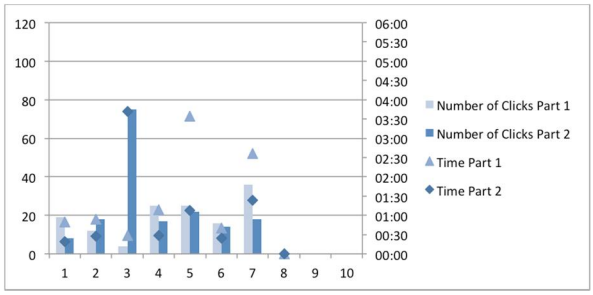
P f and h did not manage the scenario

Figure 1: Time to completion and number of clicks for the first Scenario

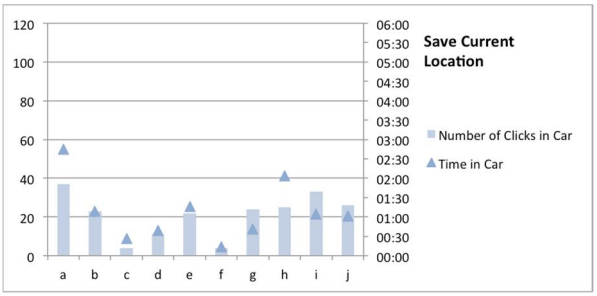
Scenario 2

In scenario 2 saving the current location could be done either by pressing *where am I* and save location or save the location by making it the *home* location. Both alternatives were considered to be the right way of solving the task, however the ones choosing to save the location as *home* struggled in part 2 as changing to another location there is not

done in the same place in the menu. This explains why participant number 3 managed the task in just 30 seconds in part 1 but tried nearly 4 minutes in part 2 without solving the task. Participant 7 was sure that he had completed the task even though he did not. That the home location is changed in another location in the menu implies that there is bad consistency.



P 8 did not manage the scenario the first or second time, P 3 and 7 did not manage it the second time.



All managed the scenario

Figure 2: Time to completion and number of clicks for the second scenario

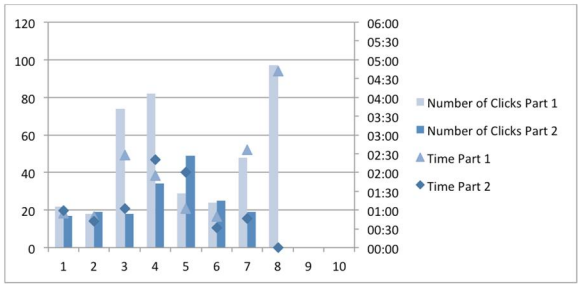
Scenario 3

In Scenario 3 setting a destination was something everyone managed to complete. Everyone hesitated at the screen where one can choose to spell city, spell postal code or search all. Four out of 18 wrote the street name where the city is going the first time using the function. The reason for this could be that the word for city in Swedish (ort) can also mean destination. If another word for city had been used, ”stad” for example, would probably the confusion be reduced. Another mistake that was done mainly by the older participants is that they would rather go back after writing the street name than pressing the square with the street name that pops up. This behaviour causes irritation with the technique because they feel that they cannot go forward with the task. The older generation asks for a confirmation button to clarify that they are on the right track. The function where the navigator suggests addresses based on the letters that have been written caused a lot of frustration for some, especially when they wrote the street name instead of the city.

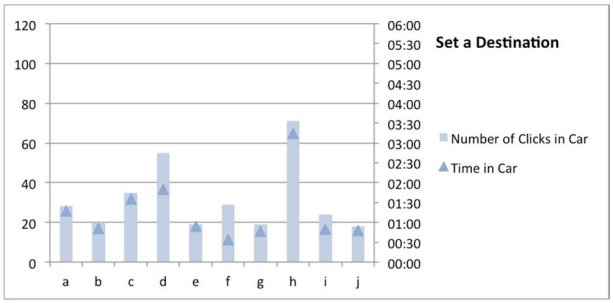
"If I try to write the street name and it gives me another suggestion, then I don't know what to do"

- Test participant H

The time consumption in part 1 and part 2 are not very differentiated which implies that the set destination function is generally self-explanatory with good gueassability.



All managed the scenario



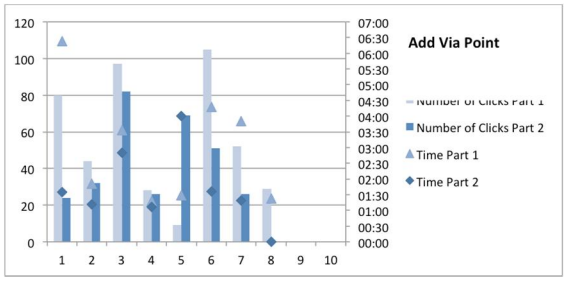
All managed the scenario

Figure 3: Time to completion and number of clicks for the third scenario

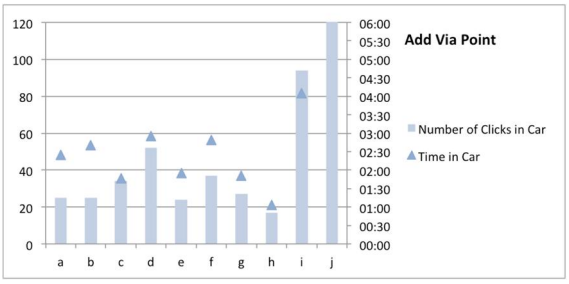
Scenario 4

In scenario 4, adding a "via" point with the Garmin Nüvi is not an easy task. In the test in car only 1 out of 10 managed to do this (a). In the test set in lab more people managed to complete the test. The differences between the tests were that the one in lab was made without GPS signal, which might have inflicted with the reactions of the navigator. The test participants in car all eventually wrote the address in the search field but only one got the choice to set it as a via point, whereas the lab participants all got the alternative to set the new destination as a via point. Common comments for the participants were that the level of feedback is too small. There is no way of seeing which destination is your first or second. Most people's first choice for setting the via point was to go through the Detour button where they expected to have the choice of writing a new address. The next choice was in most cases to go through the menus to find something that would allow to set an extra destination, leading to pressing the Extra button, which in this case means extra functions that this navigator did not have. Eventually people reasoned that they could set a new destination and add the first destination after they

had reached their via point. The choice of setting a via point come too late and the feedback is too week, resulting in a system with poor affordance and feedback.



P 1,5, and 6 did not manage the scenario the first time, 1 and 7 did not manage it the second time

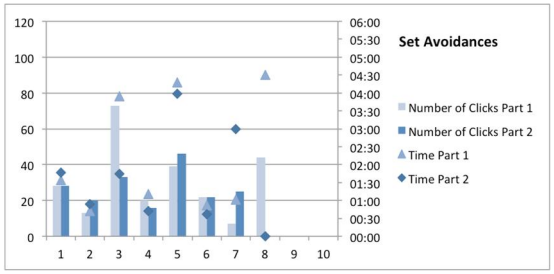


Only P a managed the scenario

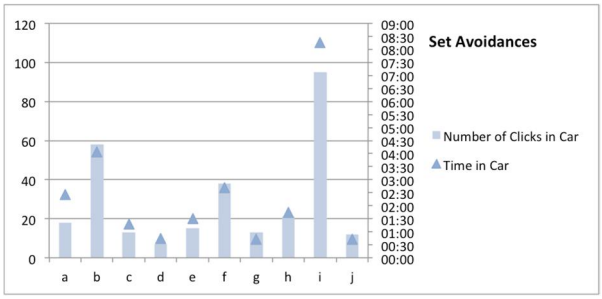
Figure 4: Time to completion and number of clicks for the fourth scenario

Scenario 5

In scenario 5, setting avoidances was a difficult task and people put low scores on the difficulty level even though most people were fairly sure that they had done right eventually. As seen below to the right the result shows that the time and clicks for part 1 and part 2 are very similar except for participant 3 and 7, which drastically decreased and increased both time and clicks respectively. Five of 18 tried to go via the traffic symbol on the map screen first, implying poor gueassability.



P 7 did not manage the scenario the first time, 5 and 7 did not manage it the second time

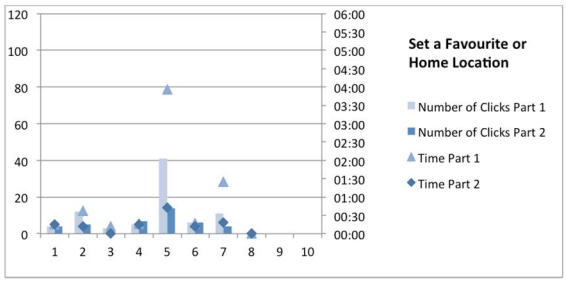


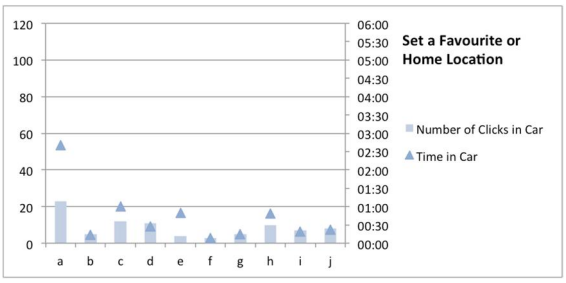
All succeeded

Figure 5: Time to completion and number of clicks for the fifth scenario

Scenario 6

In scenario 6, setting a saved location to a destination was something most people managed to do both fast and with a few number of clicks. Comments that arose was that there might have been easier to find if Favourites, displayed with a heart, would have been called Saved Locations and symbolised with a file ordering system. As the task was to save the location of the car rental firm there were a few that did not consider it to be a favourite location. Having the file system symbol would also provide compatibility towards computers and mobile phones as this is the symbol for saved files.





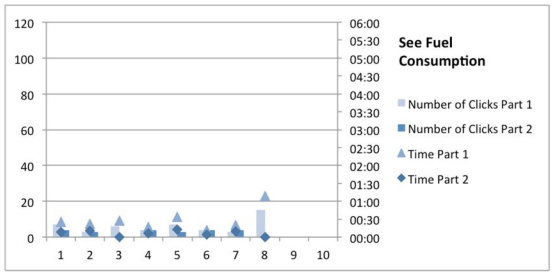
P 5 and 7 did not manage the scenario the first time, 7 had difficulties even the second

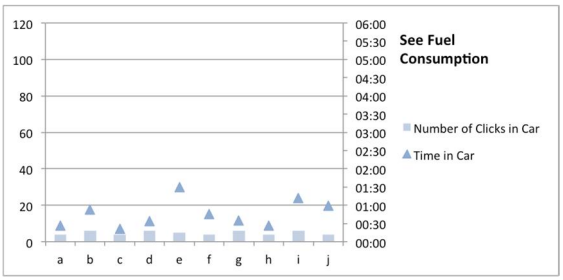
All succeeded

Figure 6: Time to completion and number of clicks for the sixth secenario

Scenario 7

In scenario 7, finding the fuel consumption was an easy task, maybe partly because people had been in that part of the menu system before when programming the vehicle profile. As seen in the chart to the right, both time and number of clicks have been reduced in part 2.





All succeeded

All succeeded

Figure 7: Time to completion and number of clicks for the seveth scenario

3.3.2 Drive Challenge

The device has, as mentioned, a simple green driving support function indicating how fuel efficient one is driving trough a leaf symbol that changes colour and an index ranging from 0-100%. Since the usability tests were done in city traffic (where it works quite badly) and in a lab (where it doesn't work at all), the use of this functionality wasn't

tested in the scenarios. To get some indication to what extent people understand the symbol, the test participants were asked about what they thought the drive challenge symbol meant. The result was varied. Many people related to the symbol of the leaf to have something to do with the environment. Some comments were that it could be something with the weather, the temperature and risk of slippery roads due to the red colour. The ones who used the driver challenge while driving did not trust the information given as it seemed that one got higher points when driving uphill and using the gas pedal more intense.

"Why should the GPS have this function when cars already have them, knowing which gear I'm using and all other information that the car knows that the GPS doesn't"

- Test participant 2

3.3.3 General Views of Garmin Nüvi

The general opinions stated after using Garmin Nüvi is that it is unnecessarily difficult to find functions in the menus and that too many clicks are required to fulfil a task. The via point function was too difficult and there was a few requests that there should be a "Via" button located on the start page. The instruction manual was too difficult to understand and the people using the manual (c, 4, 7 & 8) still found it difficult or impossible to fulfil some tasks.

"It was easier the second time. The standard functions were easy but the special functions more difficult. The navigator works fine, I find my way quite fast. Some functions have strange names though, like Favourites."

- Test participant D

CONCLUSIONS

The conclusions of the Usability benchmarking are that the devices tested in the TeleFOT FOTs could be described as fairly typical for the category of products and services that they represent. There were some usability issues noted, but nothing that would render the devices unusable. This finding is in line with the results reported in D4.7.2 Implications for User Uptake. The Garmin navigator was the only device that was put through a Usability test and not surprisingly a much greater number of usability related issues were found with this device.

Comparing the results from the Checklist and the Usability test for the Garmin Nüvi, the results are in accordance with each other. The tasks that requires the user to enter the menus are much more difficult than the basic “take me to this address”-task, which could be argued is the right choice in terms of interface design as these tasks are much more common in daily use.

The two devices tested that were most similar and therefore relevant to compare were the Blom and the Garmin Navigator. They both are Navigators with some extra functionality. Of the two the Garmin seems to be the device with the least usability related concerns, with more thought out symbols and messages. They both, however, share the problems related to using a touch display as the only input device, such as having to click into multiple menu steps in order to make common adjustments such as sound level and the need for the user to take the eyes off road as no tactile feedback is given.

One conclusion is that Usability testing is much more reliable than checklists, and that you are much more likely to find actual usability problems if you use this method. While many heuristic usability evaluation methods such as Cognitive Walkthrough and Predictive Human Error Analyses have been criticised for over reporting usability problems, the check list used in TeleFOT clearly under reports instead. One can argue that this necessarily isn’t the case for all checklists, but clearly for the “In-Vehicle Information Systems Electronic Checklist v.1.01”. One reason for the checklist to under report on usability related problems is that it focuses on only some aspects of usability (interaction), and that it has a strong focus on safety related issues, rather than on the

issue whether the user can actually use the device in a real use context. The lesson is that you have to be very careful when you choose which checklist to use.

Another finding is that the one filling in the checklist affects the result to a large extent. For instance, the analysts that filled in the check list concerning the Blom device thought that the most severe issue with the device was that the device didn't clearly warn the user not to interact with the device while driving, while the analyst filling in the check list for the Garmin Nüvi commented that the warning that is shown every time you start the device lowers the efficiency and can lead to the driver starting the trip and then interacting with the device. One can imagine that the ones doing the Blom device analyses have high concerns about safety, while the one doing the Garmin device analyses focuses more on use.

The recommendation here is that the same experts should have completed the heuristic evaluations of all systems. This had at least ensured that the different devices could be compared to some extent. Time constraints and the fact that the devices only worked in specific countries unfortunately made this impossible.