Deliverable 5.1

Usability Test Report

Project Number	AAL-2012-5-249
Project Name	MYLIFEMYWAY
Duration	March 2016 – Feb 2019 (36 Months)
Coordinator	engie
Document ID	D 2.1
Release Number /Date	V2.0/March 2018
Checked and released by	Daniel Bolliger
Document Type	Project Deliverable
Original Due Date	October 2017
Dissemination Level	Public
Main Editor	Daniel Bolliger (IHL)
Contributing Partners	VIR, IHL, End Users
Reviewed by	Herman Slagman

Abstract

This documents contains feedback for the currently used system in phase I.

1. An internal usability assessment using the scheme of Sniderman.
2. A collection of user feedback on different topics
3. An end user rating of improvement proposals, presented in form of mockups

For rollout phase II and III the document will be updated

What's new in this deliverable since MTR on November 2017?

In November 2017 There was v1.1. available, containing the usability assessment of iHomeLab, dated March 2017 (chapter 2).

For Version 2.0 this document is extended by a general usability feedback V1.2 (Chapter 3) and the usability improvement proposal survey V1.4 (Chapter4). Additionally the structure of the document is adapted, with adding executive summary (chapter 1) and conclusions (chapter 5).

The MyLifeMyWay project is Co-funded by the European AAL Joint Programme

List of Authors

Daniel Bolliger (IHL)
Martin Wüthrich (IHL)
Florian Graf (IHL)
Herman Slagman (VIR)
Dennis Bussink (VIR)

Version	*Date*	*Changes*	*Name*	*Organisation*
1.0	2018-03-27	Initial Creation	Daniel Bolliger	iHomeLab
1.1	2018-03-27	Review and released	Martin Wüthrich Florian Graf	iHomeLab
1.2	2017-11-15	General Usability feedbacks	Herman Slagman	Virtask
1.3	2018-01-24	Rework Grammar and Spelling	Daniel Bolliger	iHomeLab
1.4	2018-02-15	Usability Assessment Virtask	Dennis Bussink	Virtask
2.0	**2018-03-25**	**Update to Version for MTR2**	**Daniel Bolliger**	**iHomeLab**

Content:

1 Executive Summary

The deliverable describes the usability feedback based on the existing SW version for phase I. This software version was rolled out to the first end users in spring 2017 and was tested there in different locations.

Deliverables related to this documents are:

- *Usability Requirements Specifications D3.2* giving the system requirements in the field of usability. Now we see here the end usability feedback that validates the requirements and enlightens shortcomings of the current system version.
- *Feedback Integration Plan D 4.1* which describes the process flow how the feedback of the end users is utilized, in order to sharpen the product. The usability testing feedback is a very important source for further implementation steps in this project.

The findings in D5.1 give advice for improvements for the system of phase II and III in with a user centered design approach to the *real* end user needs.

This document comprises (until March 2018) the following three usability testing studies

- *Usability testing by iHomeLab team*. This testing follows the theoretical usability assessment technique of Shneiderman. It identifies the weakest points of the implementation in phase I and gives some proposals for improvements of the system.
- *All user feedback from the field test phase I* is compiled in this study. All points are commented by the project team. The inputs are then included for the improvements of the next versions of the system
- Virtask performed a *user group survey, to rate some design variants*. Main focus was the improvement of the usability the currently available features in January 2018. Some of this feedback is already taken into account for the phase II system version.

This document will be updated after the field tests of phase II and serves as feedback source for improvements in the last project phase

2 Usability Assessment (January 2017)

2.1 Usability Assessment

In Human Computer Interface (HCI) community there is a lot of knowledge available, how to design effective, easy to use, intuitive and pleasant interfaces for interaction between humans and computers.

In the project MyLifeMiWay we address HCI by means of a touchable user interface and relying mostly on speech recognition and speech synthesis for natural interaction between Anne and the end user.

Principles for evaluation of a specific HCI implementation or design can be evaluated in a standardized way either by Jakob Nielsen[1] or Ben Shneidermans[2] rules and principles.

In the next section 2.1.1 the principles Shneiderman are stated and in the following section 2.1.2 a rating of current implementation of Anne V.5.31 (System SW for the first field test in phase I) is given. This chapter is summarized in section 2.1.3

2.1.1 Shneiderman's 8 Golden Rules of Interface Design

Principles and Questions to consider
1. Strive for consistency *Is the style of this element maintained across your site/app? Is this content placed in the correct location according to the site hierarchy? Does this follow the conventions for your chosen platform? How can you make your designs more consistent?*
2. Enable frequent users to use shortcuts *Are there shortcuts available for your more experienced users? Whom this product is designed for? Will there be a need to consider experienced users? How can you make it easier and quicker for experienced users?*
3. Offer informative feedback *Does the user know where they are at in the process? Does the user know what they have done after performing this action? How are you communicating this feedback to your user?*

[1] https://www.nngroup.com/articles/ten-usability-heuristics/
[2] SHNEIDERMAN, B.; PLAISANT, C. Designing the user interface: strategies for effective human-computer interaction: Pearson Education. 2004.
https://public-media.interaction-design.org/pdf/Shneiderman.s.Eight.Golden.Rules.Worksheet.pdf

4. Design dialogue to yield closure *Does the user have to do any guessing here? Is it clear and obvious enough for your intended audience? Are there any next steps for the user? How are you communicating the system status with the user?*
5. Offer simple error handling *Have you done everything imaginable to prevent this error from happening on your end? Is this error avoidable in the first place? If the user does make an error, how easy is it for them to fix it?*
6. Permit easy reversal of actions *How many steps does the user have to take to reverse their actions? Will the user quickly realize they need to reverse the action in the first place? How can you make your users detect the possibility of reversal?*
7. Support internal locus of control *Will the user feel in control at this specific touch point in your app? Will they be surprised in an unpleasant manner? Does the site feel easily navigable? Does the user feel safe and in control? How can you make the user feel more safe and in control?*
8. Reduce short-term memory load *Are there enough visual cues here for the user to find the functionality or item? Do they have to remember things to understand what is going on? How can you help the user recall?*

Figure 1: Golden Rules Shneiderman

2.1.2 Rating Anne with Shneiderman's 8 Golden Rules

The Shneiderman's rules are applied to the system, rolled out in phase I (SW version 0.5.31) to the end users. Thes system has the following features and characteristics as shown schematically in Figure 2

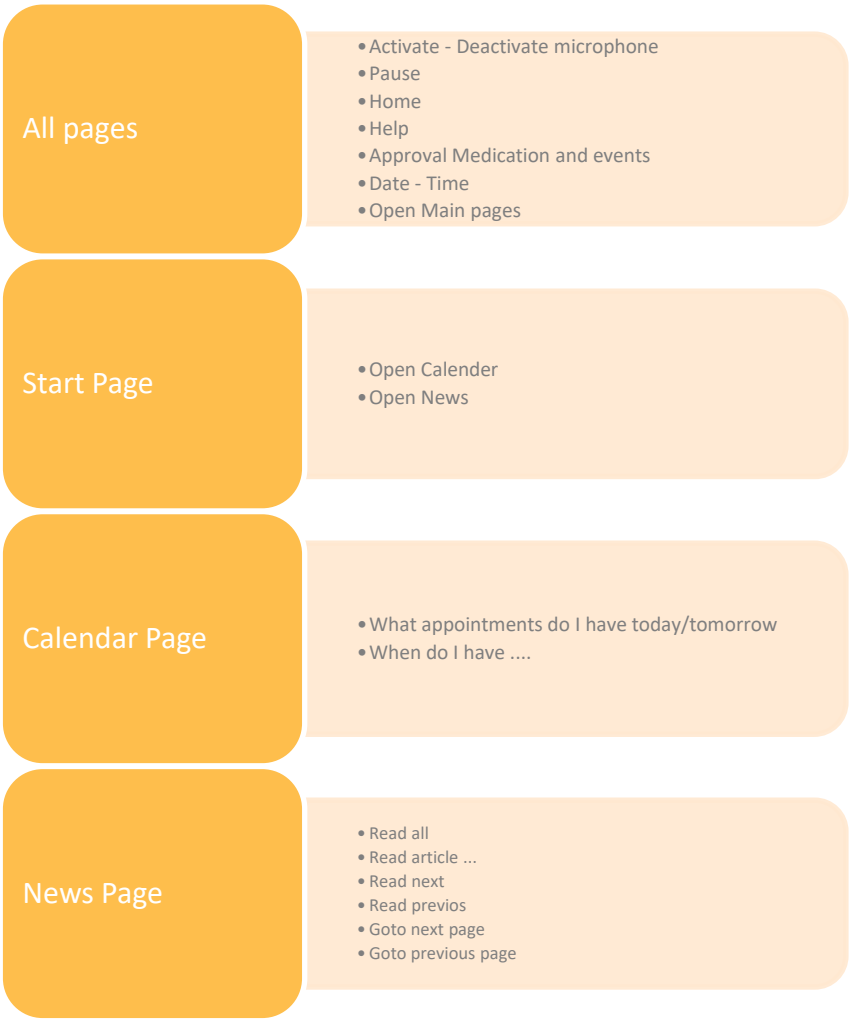


Figure 2: Features Available in V.5.31 of Anne

In the following section all Shneiderman's questions are discussed for the current software version. Most rules are followed quite well. The points that can be improved are indicated for each point in bold.

1. Strive for consistency ☺

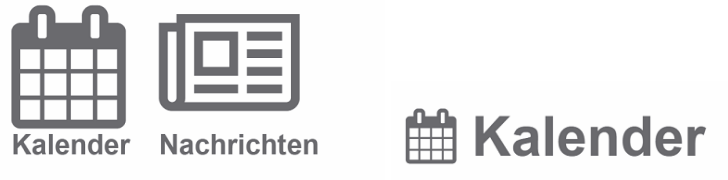
The graphical layout of the user interface is consistent over all implemented pages. Anne is represented on the right lower part of each screen. Because Anne is on the top layer, sometimes the underlying text is masked by the Anne outline

In the lower left part of the screen the following icons are visible always:

 . The red embossed icons are touchable.

In the right upper part of the screen, reminders and date-time information is visible on all screens

on top either the actual page location or the tiles for main pages navigation are visible:



Consistent and clear pictograms labeled on top pages, enable the users easyly to understand what he/she can do. It is clearly labeled on which site he/she is at the moment.

Additionally positive can be statet that on each page direct access to the home button is available. The location of the icons is consistent through all pages.

Missing on the design, is the differentiation between clickable/executable (Kalender, Nachrichten) screen elements information only elements (WLAN, Battery…) and static content (Article)

2. Enable frequent users to use shortcuts ☺

The combination between touchable icons and speech commando allows the user to navigate directly to the place in the software he/she wants to go. No hierarchical constructs are required to reach the desired function . This really helps the user to navigate very quickly to the desired place in the software.

Positive is, that on the lower icon ribbon, common main functionality is available in a consistent way.

The list of available speech commands is not available on the application itself. Either the end user knows the required command exactly by heart, or the user has to consult the user documentation (quick reference, command lists…).

3. Offer informative feedback ☹

Yes – the user always knows, where she/he is in the interface (see also point 1).

In the 'happy-day scenario' the HCI reacts prompt to the speech command or the touch on an icon.

Assuming, that the user wants to initiate an action by speech – and the action fails – the user does not get any qualified feedback from the system that something went wrong. The system stays in the state as it was before the action attempt. **This is really confusing for the user and reduces strongly the user experience! Here major improvement has to be implemented**
4. Design dialogue to yield closure 😐 This is only applicable in multi-step actions (e.g. fill in a form in several steps). With the currently implemented feature set, navigation to the main page is feasible with one click. From the main page all main functions are also reachable with one interaction step. By extending the application with additional features, this currently available interaction concept has to be followed further.
5. Offer simple error handling 😐 The system is designed in a way that it is robust against any input of the user. It should not crash or be confused after any user input. If the system reacts not in that way, as the user expects, a second and third attempt is possible without any limitations. A plus is, that the icon to reach the entry point (house symbol) is always available. It makes also sense to have as many icons operable by speech command as well as by touching them. **Referring to point 3 – the reaction of the system to non-successful speech commands is missing or very poor. There improvements have to be implemented.**
6. Permit easy reversal of actions ☺ Because there are no nested procedures to initiate an action, the reverse action is quite simple: Go to the start page and re-initiate the desired action again. **Only for actions that provoke a longer feedback of the system (as in the news section performing the command read all), the microphone is muted during speech synthesis of the system. Therefore there is no speech command available to stop the current action. The user has to wait until speech output is finished.**
7. Support internal locus of control ☺☹ In the 'happy day scenario' the user has everything under control, feels secure and gets the intended reaction of the system → thumb up

Operation of the system by touching on the screen is straight forward and very intuitive → thumb up **If something goes wrong (especially speech detection) the system behaves not predictable. Either there is no reaction or a reaction that is not expected. Due to the fact, that there is no feedback of the system in such a case, the user is almost lost. This effect is already described in the rules 3 and 5. → thumb down and major actions to be taken.**
8. Reduce short-term memory load ☺☹ Visual representation of the system state and icons available for touch interaction are well placed and also available in the right amount (not too few, not overcrowded). The user can easy operate the system → thumbs up **For speech commands the system relies on commands (fixed patterns – no free speech analysis), that have to be initiated:** • **with exactly the expected wording** • **with almost written language** **This requires from the user, that he/she exactly knows by heart, what are the valid commands in this context. This is on one hand a matter of training and documentation. On the other hand this is also depending on the ability of the user to learn all commands by heart, and to know where which command is available. Until now, the system does not support the user interactively.**

Figure 3: Shneiderman Anne HCI Assenment

2.1.3 Conclusion

The *majority of the software implementation is working well and follows the usability rules* of section 2.1.1. Albeit a lot of work is already done the following points reduces the usability and user experience in the current version of MLMW:

- Graphical differentiation between executable icons and only info icons / sections on the GUI.
- Missing context sensitive presentation of available speech commandos
- Missing feedback of the system on speech recognition (quality, what is understood, what action is performed).

2.2 Usability Blockers

There are two main areas that limit the usability of Anne:

- There is almost no feedback of Anne on a performed user interaction
    - In the 'happy-day' scenario the reaction of Anne is as expected and everything is fine. Feedback of Anne is not needed, because the reaction of Anne indicates to the user that everything went fine.

- In 'failing-scenario 1' nothing happens on the performed user interaction. The user does not know if Anne has understood him; if there was the wrong/unknown command; if something else is not working.
- In 'failing-scenario 2' an unexpected action as reaction on a performed user interaction takes place. The user does not expect the current reaction of the system. It is not understandable for the user why Anne reacts like this.

- The commands utilized must match exactly the dictionary (lookup table for the grammar) in the current context. The user knows well what he/she wants to do, but does not exactly remember what are the required command(s) for this action.

2.2.1 Commands

Anne is reacting on command based key phrases. They have to be used exactly as stored in the system. If the user provides a synonym or acronym to Anne, she will not understand and therefore not react. For users it is difficult to remember the exact command – it is much easier to remember the context or the meaning behind the commands.

2.2.2 Unexpected Actions

- Without any wanted interaction with Anne, she suddenly opens a help page or shows the date and time. This is scary.
- By sending one clear command, she understands something else and behaves according to her interpretation and gives no feedback to the user what she understood exactly.
- Sometimes she refuses doing anything and the user does not know why.

2.2.3 Feedback

As described in Section 2 Anne has **no explicit feedback** system integrated. In normal human interaction, the feedback of a communication partner is essential for a satisfying communication. Active listening is a measure for functioning and agreeable communication. With active listening, the statement of the partner is summarized, in order to guarantee that the message from the partner is well interpreted. Further, the sender understands the receiver's reaction better.

2.2.4 Emotions

Anne shows emotions. Figure 2 is a common emotion, that is shown by Anne.

This emotion often displays with non-understandable root cause.

In our interpretation, it shows disagreement of Anne towards the user. This is scary because it is not clear why Anne shows this mimic.

In future, the user needs more assistance or feedback from Anne for better understanding her.

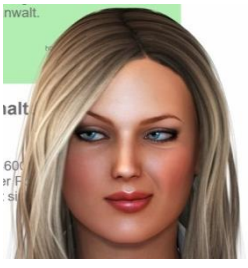


Figure 4: weird Anne

2.2.5 Availability Microphone

- We observed often difficulties to activate the microphone after a longer idle time. The command 'hallo Anne' has to be repeated several times until Anne turns active.
- Often the switching time between saying 'hallo Anne' and the active symbol of the microphone takes 1-2 seconds to be switched. This is alienating the user. He/she does not know if the command is well understood by Anne. As a consequence the command is repeated by the user louder and eventually already impatient or angry.
- While Anne is talking, the microphone is muted. Therefore it is impossible to interrupt her in an ongoing action. This sometimes sometime is annoying.
- After executing actions, the microphone is muted for another 1-2 seconds at least. Thus while restart talking to Anne immediately, Anne is not yet listening to the user.

2.2.6 Dialects and Pronunciation

- Anne is available in Dutch, Flamish and German. The multilingual speech commands and displayed contents can be switched quite easily, in the application configuration settings.
- Pronunciation in 'standard' language is required, because text to speech is utilized by Anne for recognition of the key words.
- The overall pronunciation of Anne is sometimes bad and thus therefore difficult to understand.

2.3 Improvement Proposal

The ideas / requests below propose a possibility to implement the requested enhancements. Graphical layout and specific implementation can be discussed more precisely during the realization phase. The samples and ideas for layout serve as base for discussion, the real implementation has to be defined by the whole consortium.

2.3.1 Clickable vs. non Clickable

- Make microphone state clickable, in addition to the speech commands 'hallo Anne' and 'danke'. With this additional option the system can be waked up in any case.
- Distinguish graphically what icons and regions are clickable and which are not. Underlay the clickable area with a light colour shade, that can be configured in the dashboard

sample for clickable icons	sample for non clickable icons
 	

Figure 5 Clickable and non-Clickable Regions

2.3.2 Context-Sensitive Speech Commands Display

Give the user a dynamical list of available commands that is context-sensitive. This additional list of available speech commands should be displayed on request. The speech command for this action could be the key-word ‘commands‘ and the clickable icon on the upper right side of each site as displayed in the following figure:

Figure 6 Commands Context-Sensitive Activation Icon

If the key-word ‘commands’ is initiated – or the icon is clicked, a dynamic list with available speech commands is displayed. The most probable command is highlighted. All items are clickable.

Figure 7 Clickable Speech Commands - Popup

After clicking one list-item or using the corresponding speech-command, the list is automatically collapsed. The list is also collapsed by touching the commands icon or after some seconds idle time.

2.3.3 System Information – Current Status

Missing feedback of the system is one of the most alienating issue of the current system version. Moreover, it really lowers the usability experience very much.

Our proposal is to have an additional option that can show the current action and speech recognition of the system in an easy way for the user

- Have an additional 'ear' icon on the upper right. By clicking (clicking only), a running text line is shown on the lower part of Anne, providing actual system information

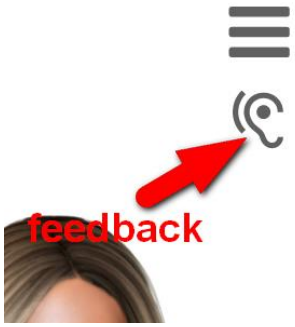


Figure 8 Clickable Feedback Button

- Recognized speech inputs are displayed with a colored smiley (red – to low acceptance, or/and not understood, yellow critical acceptance and text, green with recognized command. Additionally log messages can be displayed in this line
- Make the output graphically colored, depending on the operation status – some kind of chat between end user and Anne.

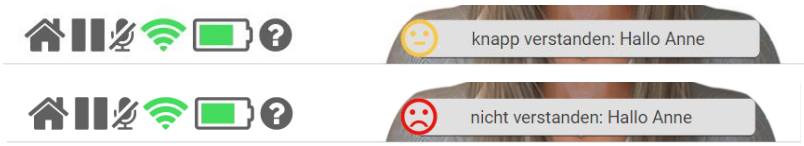


Figure 9 System Feedback Qualified with Emoticons and Colors

3 General Usability Feedback (Fall 2017)

Software system with version 0.5.31 was rolled out to our end users from spring 2017 onwards. Virtask and iHomeLab are for all sites the technical hotline. In this function a lot of direct contact with the end users took place. Virtask collected all these inputs, complaints and questions from them. The resulting grouped list can be found just below. In the right column comments of the project team is given.

User Input	Comment
Speech recognition	
1. Full speech recognition is already there why not in Anne? 2. Anne is listening poorly sometimes she hears stuff and does things that i haven't asked off her 3. If the surroundings are busy, Anne will just do stuff 4. It is better to just let Anne listen by pushing a button and stop listening by pushing the same button 5. Hybrid version will be better. first You have to push a button to activate Anne, and to deactivate again push a button	*Speech recognition is not stable enough yet. It is done at the moment on command base, and therefore due to the limited comparison possibilities in the grammar often misinterpretation takes place. There are two measures to compete with this weakness:* *Combination of touch and ASR will be done in critical modules as intermediate solution.* *On mid term range a new ASR method is developed and implemented that's based on free text input, and promise to be more robust*
Speech adaptability	
1. Same sound/volume. For example when a conversation partner isn't loud enough 2. Can the user change the speech speed, volume, etc. etc. We are getting more and more capable end users of course.	*1 is an interesting idea, we will follow up. 2 can be done at the moment in the dashboard easily. By intension there is no possibility to change parameters on the device itself. We think about simplifying the dashboard*
Agenda and Medication	
1. Agenda is the least clear overview. Maybe comparing this with other existing agendas? Or integrate a good other agenda? 2. Agenda an medication separated is handy if I will still get my warnings in connection with actions 3. Where does the medication notification go to? Can i also get an overview of a week or more than a week myself?	*1,2 There is planned for the rollout of phase II to separate the agenda (google agenda) from the reminders and to make the events much better readable. This can be seen also in section 4.3 what end users like most.* *3 medication reminder feedback is available in the dashboard for administrators of the end users already* *4 For elderly users it is too complex to fill in an agenda entry. And with the lack of a*

User Input	Comment
4. Filling the agenda in the Anne program itself (there are at least 5 - 10 users who would like that already)	*keyboard, only with speech recognition the process is quite difficult. With the use of the Google agenda in the background there is the possibility of using the API functionality to build an easy to operate add-on as standalone app or to be launched with Anne*
Individualization	
1. Making text size adjustable in either Anne or the dashboard 2. Deciding the size of Anne yourself 3. Changing the background yourself 4. Could the sleep screen also be a photo screen on which i could play my own pictures (Dia-show) 5. Is the avatar adjustable, can i choose a different avatar already 6. I would like Anne to be on screen all the time; she is becoming a main focus point for me. If she is there sometimes and sometimes not it will give confusion.	*1-6 Especially the active end users are able to handle individualization of their devices. In our opinion this is an important identification booster for the whole system.* *But not all users are capable of adjusting everything right.* *Therefore we propose to extend the dashboard in order to individualize and configure Anne.*
Additional Features	
6. Can something flicker if I have to do something? For example the medication notification as a kind of attention 7. Can i chat with Anne, for example: asking Anne to open a webpage? Or asking Anne to show the weather? 8. I would like to be able to receive and answer e-mails with Anne. 9. There hast to be a night screen, because Anne is much too bright and disturbs me much, when I have her in my bedroom next to me	*6 very interesting idea of making notifications or reminders more prominent. It is thinkable to make visual and sound more prominent in the notifications. Alternatively it could be possible to attach an external blue-tooth linked alarm device to Anne.* *7 Weather is under consideration. General purpose search requests, seems to be very delicate, especially a competition analysis showed that the robustness of such systems is still very poor even at Google, Alexa etc.* *8 Mails receiving should not be a problem but answering with ASR only is not stable enough at the moment* *9 This also a very valuable input from our end users. The development takes place on daytime in an office environment. So*

User Input	Comment
	the developers did not think about the typical end user situation at home and in the night. In the next section 4.1 some night screen proposals are demonstrated to the end users

We take the input of the end users collected here together with other information sources in consideration for the next feature releases.

4 Usability Improvement Survey January 2018

In this document section, you can find the mock-ups, slides with different look and feel of Anne. These different look and feel of Anne (mock-ups) were shown to around 25 end-users. The mock-ups were shown them and their judging and feedback was noted. That gave meaningful input, used for improvements of the representation of Anne on the screen.

The results will be implemented step by step in the next releases.

The sections below are organized according to the occurrence in the application. Each section is organized with different versions of the screen layout and the reaction of the end users to the versions with the two categories:

- The "votes" for each slide.
- Additional comments for each slide.

4.1 Sleep screen

In an inactive state Anne is still very bright and disturbs the users. Additonally it is not clear visible wether Anne is active or in a 'sleep mode'.

Slide 1:	People who like it	14
	Additional comments:	- weather incl. temperature in degrees Celsius
Slide2:	People who like it	7
	Additional comments:	-
Slide 3:	People who like it	5

	Additional comments:	-
Slide 4:	People who like it	1
	Additional comments:	-

4.2 Homepage

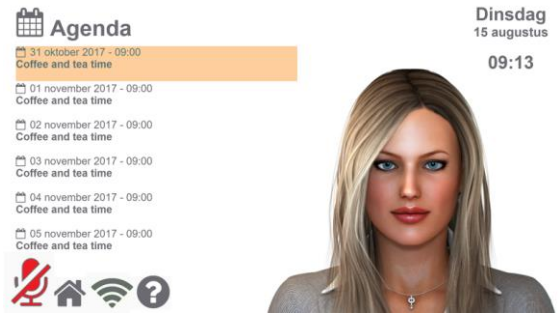
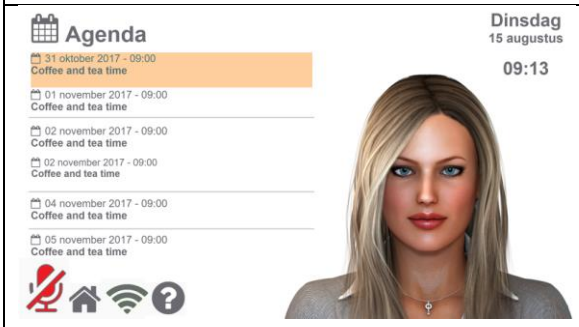
Is there a need for reorganizing the home page? – Sizes, accents, colors …

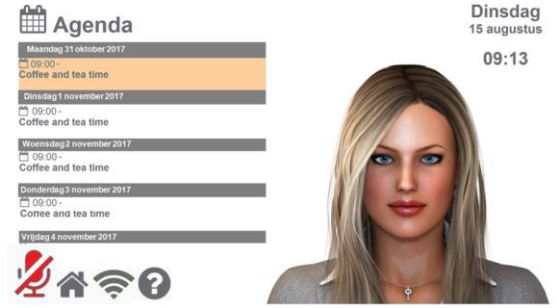
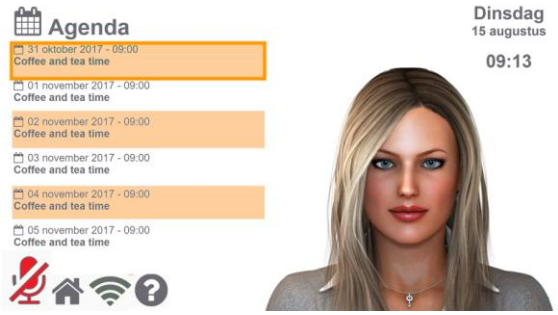
Slide 1:	People who like it	18
	Additional comments:	- can it also be modified to my wishes? Colour or picture background?? - Making it choosable for the user what the 3 main functionality's are. Give those a big Icon on the start screen, make the other icons small to make it less busy.

Slide2:	People who like it	9
	Additional comments:	- Could the year also be on screen?

4.3 Agenda

The agenda of version 1 is unstructured, only a list. A lot of users complained about and wished to have a better structure and visual differentiation and grouping of the entries.

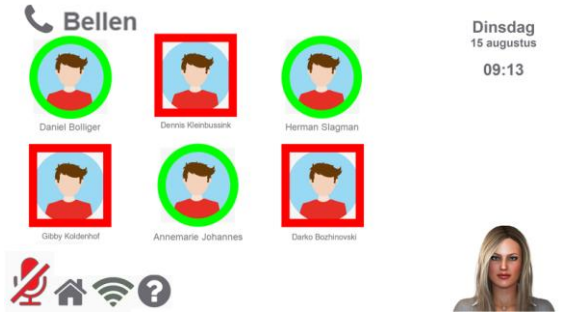
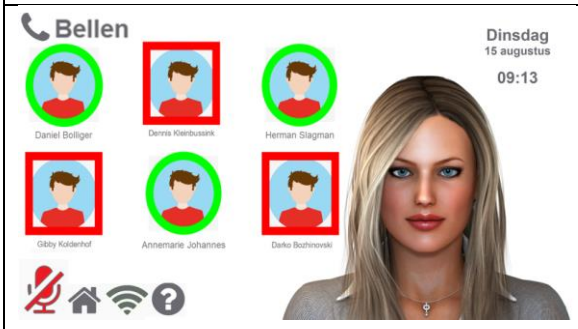
Slide 1:	People who like it	2
	Additional comments:	-
Slide2:	People who like it	9
	Additional comments:	-

Slide 3:	People who like it	15
	Additional comments:	-
Slide 4:	People who like it	0
	Additional comments:	-
Slide 5:	People who like it	1
	Additional comments:	-

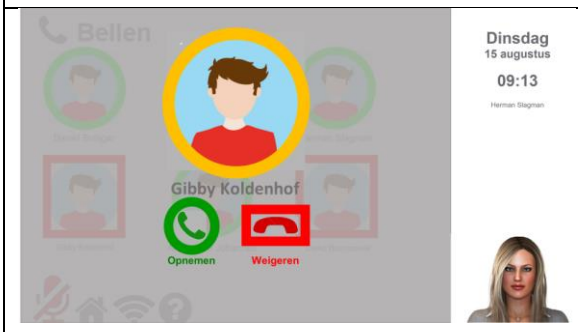
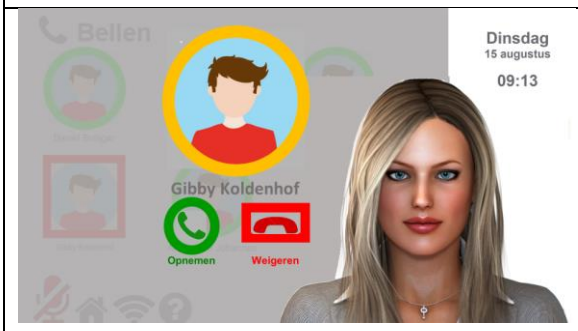
4.4 Videocall – Calling

Especially the graphical representation of the video calling functionality was of interest. Shape vs. color, as well as size of Anne.

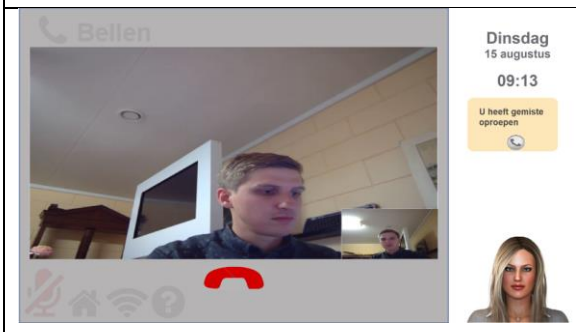
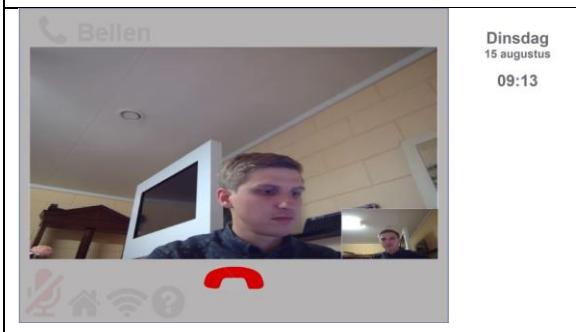
Slide 1:	People who like it	14

	Additional comments:	- When a contact is offline blur the contact and lose the red square. So just the green for online contacts and blurred offline contacts. This makes it less busy.
Slide2:	People who like it	11
	Additional comments:	- When a contact is offline blur the contact and lose the red square. So just the green for online contacts and blurred offline contacts. This makes it less busy.

4.5 Videocall – Incoming Call

Slide 1:	People who like it	20
	Additional comments:	- When a contact is offline blur the contact and lose the red square. So just the green for online contacts and blurred offline contacts. This makes it less busy.
Slide2:	People who like it	4
	Additional comments:	- When a contact is offline blur the contact and lose the red square. So just the green for online contacts and blurred offline contacts. This makes it less busy.

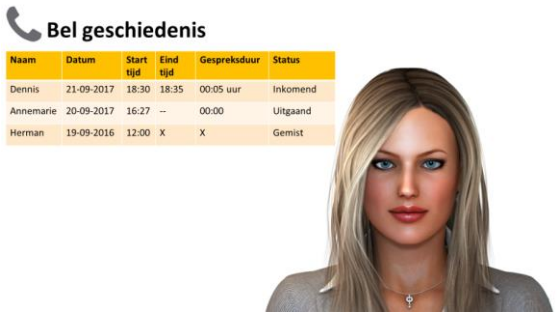
4.6 Videocall – During a Call

Slide 1:	People who like it	22
	Additional comments:	-
Slide2:	People who like it	2
	Additional comments:	-

4.7 Videocall – History

This is a new idea, if we should provide in a second step the video calling history to our users

Slide 1:	People who like it	2
	Additional comments:	- This is clearer, but i still want Anne to be visible
Slide2:	People who like it	11

	Additional comments:	- with pictures of the caller - Making it possible to directly call back a person from this page.

4.8 Home Automation

Slide 1:	People who like it	8
	Additional comments:	-

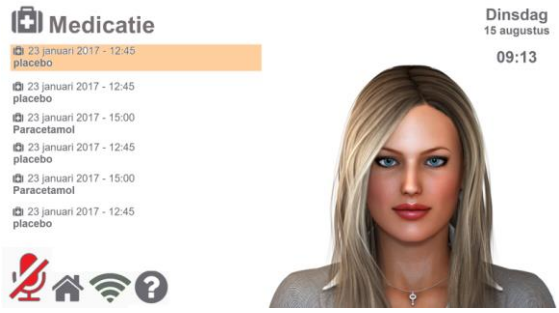
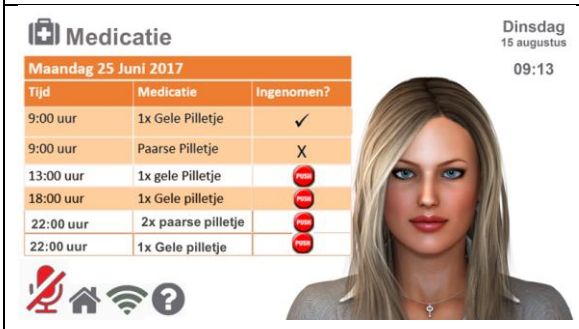
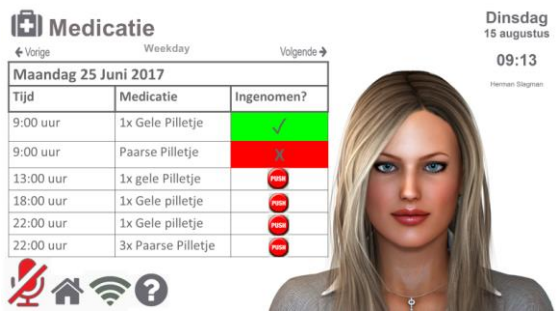
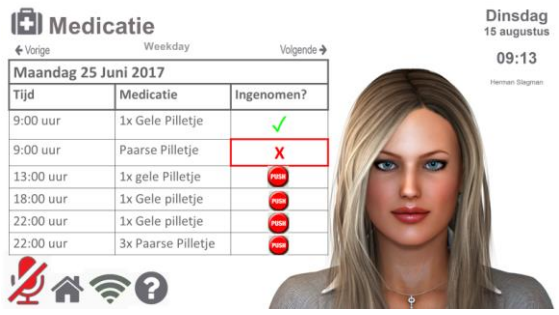
4.9 Entertainment

Slide 1:	People who like it	8
	Additional comments:	- Maybe add a radio and music option?

4.10 Medication

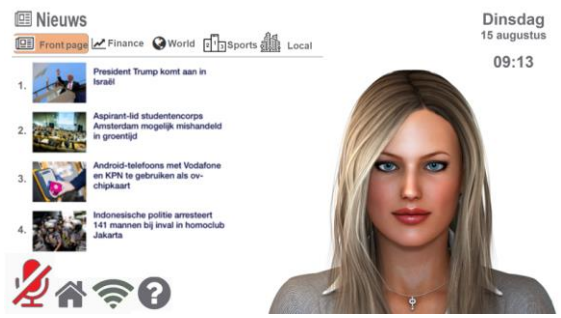
Is there a need to have an overview over the history of taken/accepted reminders?

Slide 1:	People who like it	1

	Additional comments:	-
Slide2:	People who like it	3
	Additional comments:	-
Slide 3:	People who like it	6
	Additional comments:	-
Slide 4:	People who like it	17
	Additional comments:	- For me it is, but not for my partner with dementia. - Just let the pushbuttons appear at the times you should take the medication. So not for the whole day or further. - Background colour for the medication that you have to take at that moment or in the timespan.

4.11 News

Rss Feeds: is there a better way to present them to our users?

Slide 1:	People who like it	2
	Additional comments:	-
Slide2:	People who like it	1
	Additional comments:	-
Slide 3 and 4:	People who like it	23
 	Additional comments:	- Maybe add a read list so that you can scroll all titles first and let Anne remember what you want to hear. So she can later read all the articles you selected in a row.