D4.1 User testing evaluation report

Project Number	2018-5-46-CP
WP Number	4
WP Title	Evaluation & Impact Assessment
Type	Report
Dissemination Level	Public
Version	v3.0
Delivery Date	22.February 2022
Author	Michael Gstöttenbauer, LFTL
Internal Reviewer	Bernhard Matzner, YOU

Version History

Version	Date	Author	Partner	Description
v0.1	02.04.2020	Michael Gstöttenbauer	LFTL	1st draft, ToC
v0.2	12.5.2020	Michael Gstöttenbauer	LFTL	1st round SALSA Fun Wireframes, Heuristics
v0.3	13.5.2020	Cheyenne Schuit	NFE	1st round SALSA Fun in NL
v0.4	15.9.2020	Michael Gstöttenbauer	LFTL	1st round SALSA Fun, AT, SUI
v0.5	16.9.2020	Michael Gstöttenbauer	LFTL	1st round SALSA Health
v0.6	13.4.2021	Michael Gstöttenbauer	LFTL	2nd round SALSA Fun in NL
v1.0	1.9.2021	Michael Gstöttenbauer	LFTL	Finalisation for internal review
v3.0	22.2.2022	Michael Gstöttenbauer	LFTL	Finalisation for external review

Table of Contents

1   Introduction

This deliverable gives an overview of the methods and results of the pre-prototype test activities during the design and early implementation phase of the SALSA app until January 2021.

Wireframes, mockups and pre-prototypes were used to evaluate the early design before and during the actual real technical development. The goal of the pre-prototype testing is to gather as much feedback as possible, which helps to identify any (design) issues before the development gets to the expensive part of the process.

These prototypes were tested with different users from an early stage on and iteratively. They generally included a basic navigation structure, some interactive functionality and later on the design. They were tested with experts and real end-users from the target groups in evaluation sessions.

An evaluation manual was created to help the pilot partners perform the evaluation sessions. This manual contains basic information on the evaluation methods as well as the tasks, that the end users should perform within the application.

All feedback was collected and presented using the online tool Trello (https://trello.com/). The findings were discussed, evaluated and assessed in iterative workshops and finally led to changes/redesigns in order to define the MVP (minimum viable product), the final version, that would be used in the pilot phase.

Following the recommendations of the Lean Startup Academy SALSA was designed as two separate solutions, SALSA Fun and SALSA Health: SALSA Fun strongly influenced by Walking Sports groups in the Netherlands and targeting the needs of its trainers, SALSA Health targeting physiotherapists and their patients.

After the mid-term review meeting and in order to follow the recommendations given a number of corrective measures were implemented regarding SALSA Health. Instead of pursuing the initial approach of building a tele-rehab solution from scratch and including parts of the MIRA system for exergames, the consortium revised the technological architecture and decided to build more strongly on the readily available MIRA platform and supplement it with new functionalities, especially for home use.

2   Methods/Procedure

The evaluation ideally consists of two parts:
1) **Heuristic evaluations** done by evaluators from partner organizations
2) **Thinking aloud sessions** with target users observed by researchers from the pilot partners

The interactive prototype evaluation should be conducted with 3-5 end users per target group (people 55+, trainers, …) per pilot site. The heuristic evaluation is to be carried out by 2-3 experts of each partner organization. Theory says that in an early phase of a project where the focus is more towards usability it provides no additional value to include a bigger amount of users to gain feedback.[1]

The identified issues from the evaluations are documented in Trello[2], where boards for each user group and experts per partner are prepared.

- Trello board for heuristic evaluation: http://bit.ly/34eqYKa
- Trello board for evaluation with people in 3rd age (55+): http://bit.ly/3bfZCXV
- Trello board for evaluation with trainers: http://bit.ly/2vKsuau

All feedback was collected and presented to the technical partners using the online tool trello. The findings were discussed, evaluated and assessed in iterative workshops and finally led to changes/redesigns in order to define the MVP (minimum viable product), the final version, that will be used in the pilot phase.

2.1   Heuristic Evaluation

Developed by Nielsen and Molich, 1990 [3]

In General

Heuristic evaluation is a method for finding the usability problems in a user interface design so that they can be attended to as part of an iterative design process. Heuristic evaluation involves having a small set of evaluators examine the interface and judge its compliance with recognized usability principles (the "heuristics").

The recommended number of evaluators is three to five (since one does not gain that much additional information by using larger numbers).

Heuristic evaluation is performed by having each individual evaluator inspect the interface alone. The results of the evaluation will be recorded as written reports from each evaluator and aggregated afterwards.

Typically, a heuristic evaluation session for an individual evaluator lasts one or two hours. Longer evaluation sessions might be necessary for larger or very complicated interfaces with a substantial number of dialogue elements, but it would be better to split up the evaluation into several smaller sessions, each concentrating on a part of the interface.

---

[1] https://goo.gl/X271D3
[2] https://trello.com/
[3] https://goo.gl/hqxqGM

During the evaluation session, the evaluator goes through the interface several times and inspects the various dialogue elements and compares them with a list of recognized principles (the heuristics). These heuristics are general rules that seem to describe common properties of usable interfaces. In addition to the checklist of general heuristics to be considered for all dialogue elements, the evaluator obviously is also allowed to consider any additional usability principles or results that come to mind that may be relevant for any specific dialogue element.

In principle, the evaluators decide on their own how they want to proceed with evaluating the interface. A general recommendation would be that they go through the interface at least twice, however. The first pass would be intended to get a feel for the flow of the interaction and the general scope of the system. The second pass then allows the evaluator to focus on specific interface elements while knowing how they fit into the larger whole.

The output from using the heuristic evaluation method is a list of usability problems in the interface with references to those usability principles that were violated by the design in each case in the opinion of the evaluator. The evaluators should try to be as specific as possible and should list each usability problem separately.

Set of Heuristics

1) **Visibility of system status**
   The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2) **Match between system and the real world**
   The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3) **User control and freedom**
   Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support *undo* and *redo*.

4) **Consistency and standards**
   Users should not have to wonder whether different words, situations, or actions mean the same thing.

5) **Error prevention**
   Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6) **Recognition rather than recall**
   Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7) **Flexibility and efficiency of use**

Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8) **Aesthetic and minimalist design**
Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9) **Help users recognize, diagnose, and recover from errors**
Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10) **Help and documentation**
Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

2.2   Thinking Aloud

Developed by Judy Ramey, University of Washington, with additions by Usability Analysis & Design, Xerox Corporation[4].

Think-aloud user tests involve participants thinking aloud as they are performing a set of specified tasks. Participants are asked to say whatever comes into their mind as they complete the task. This might include what they are looking at, thinking, doing, and feeling. This gives the observers insight into the participant's cognitive processes (rather than only the tested prototype). Think-aloud sessions are best done with two observers, one is guiding the test, explaining things, asking questions, ... the other observer is primarily taking notes of what participants say and do, without attempting to interpret their actions and words, and especially noting places where they encounter difficulty.

The idea is to look if the participants think the application is logical and if they have any ideas about the design. It is important to see what the participants like, dislike, understand and (most important) don't understand. It might be interesting to ask the participant some additional questions about the buttons. These questions would be about the place of the buttons ("Does it make sense where the buttons are?"), if there are too many or too few buttons ("Were there too many functions or maybe did you miss a function?") and if the buttons were clear to them ("Are the buttons/texts/icons clear to you?").

So in each task comment on the following:

- Like/dislike and why?
- Understand/don't understand and why?
- Does it make sense to you where the buttons are?
- Are there too many functions?
- Did you miss a function?

---

[4] http://bit.ly/397tFiO

- Are the buttons clear enough?

In the manual (see Appendix) we provided guidelines and recommended techniques for the observers.

2.3   UEQ

For testing the user experience, we make use of the validated User Experience Questionnaire by Laugwitz, Held and Schrepp [5], which allows a quick and simple assessment and comes with a tool for evaluating the results. Besides it is available in all pilot site languages. For obtaining feedback on specific problems and suggestions for improvements, additional open questions are necessary.

The format of the questionnaire supports users to immediately express feelings, impressions, and attitudes that arise when they use a product. The scales of the questionnaire cover a comprehensive impression of classical usability aspects and user experience aspects:
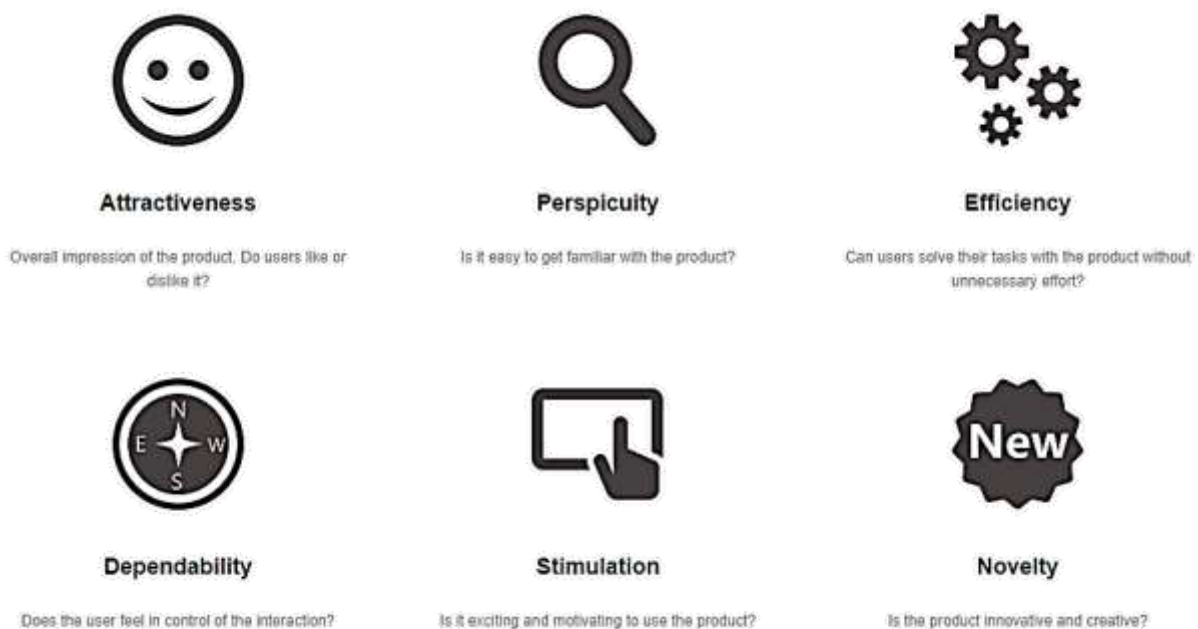
Attractiveness
Overall impression of the product. Do users like or dislike it?

Perspicuity
Is it easy to get familiar with the product?

Efficiency
Can users solve their tasks with the product without unnecessary effort?

Dependability
Does the user feel in control of the interaction?

Stimulation
Is it exciting and motivating to use the product?

Novelty
Is the product innovative and creative?

Figure 1: Scales of the UEQ

The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. The agreement can be expressed with the attributes by ticking the circle that most closely reflects the impression about the product. The user should decide as spontaneously as possible in order to make sure that the original impression is conveyed. In order to analyze the results of the questionnaire an MS Excel tool is available where you can fill in the data.

Example: This response would mean that you rate the application as more attractive than unattractive.

| attractive | ○ | ⊗ | ○ | ○ | ○ | ○ | ○ | unattractive |

---

[5] https://goo.gl/a5jVtC

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Figure 2: The 26 items of the UEQ

3   Wireframes Testing SALSA Fun

In February 2020 the first round of pre-prototype testing started, focussing on the wireframes of the SALSA Fun webpage.
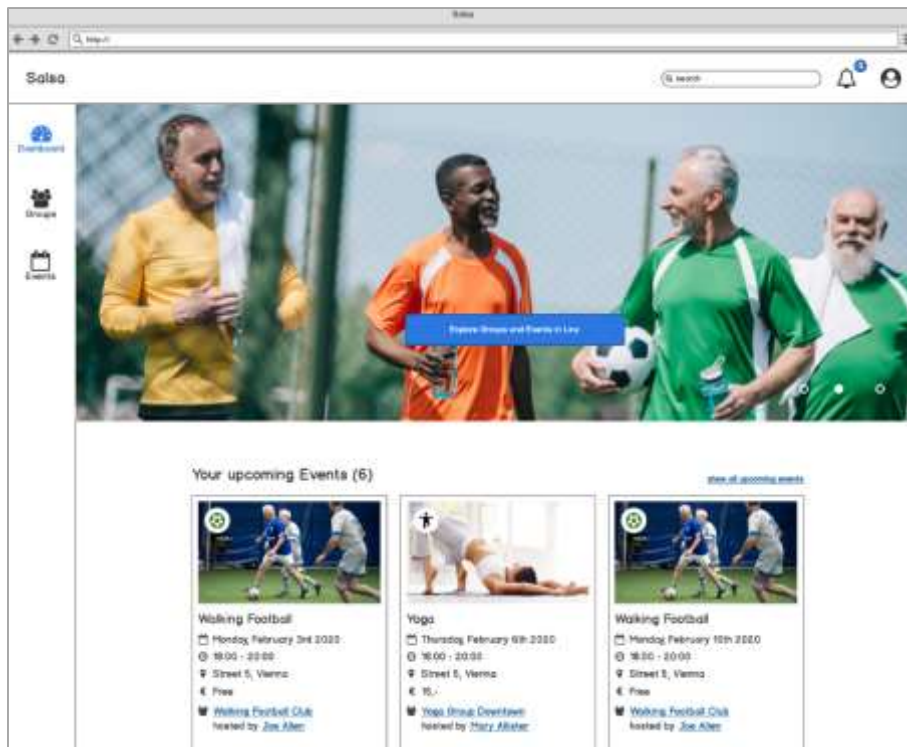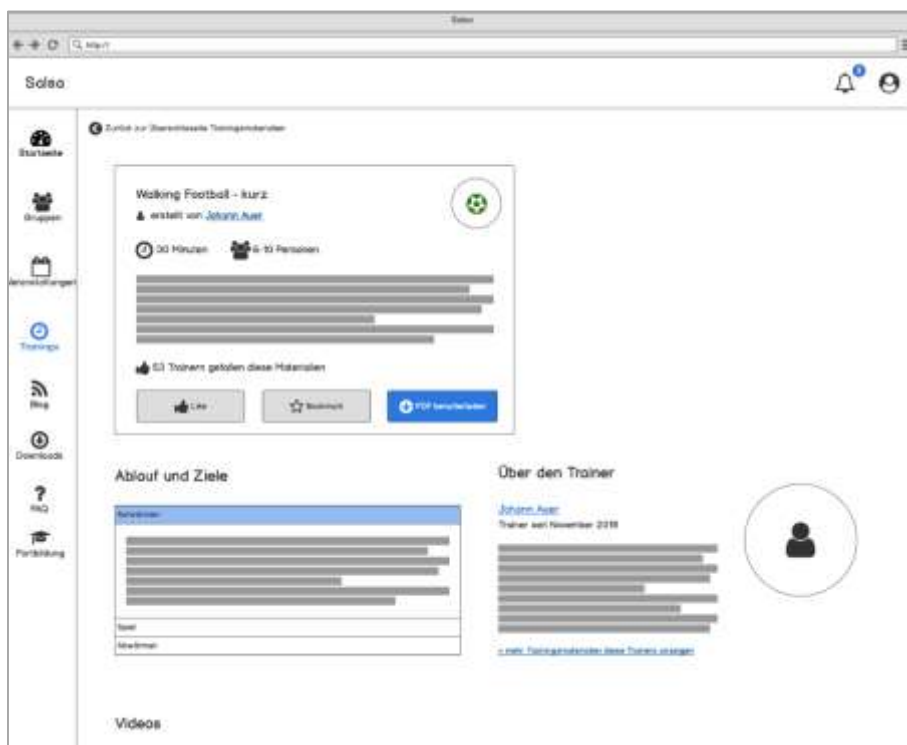

Figure 3: Wireframe SALSA Fun Dashboard

Figure 4: Wireframe SALSA Fun Training Template

These were not the very first wireframes and already included a lot of feedback gathered in evaluation settings, especially with Oldstars Walking Sports trainers in the Netherlands. They already mocked simple functionalities, mostly navigation.

The following tasks were prepared to do in the wireframes:

For users 55+:

1. *Task - DE - Search for walking football groups in Linz*
2. *Task - DE - Join walking football group Linz/Urfahr*

1. *Task - NL - Search for walking football groups in Rotterdam*
2. *Task - NL - Join Oldstars walking football group Rotterdam*

For trainers:

1. *Task - DE/NL- Create a new group*
2. *Task - DE/NL- Create a new event*
3. *Task - DE/NL - Search for walking football trainings material*

3.1   Heuristic evaluation

The heuristic evaluation among project partners and experts provided the following feedback:

Visibility of system status

- No - In the current prototype you do not see any hover effects on buttons (change in colour etc.)
- No - No progress bars/waiting icons when loading information (although this might not be possible with the tool the mockups were made in…)
- Yes - Headers for each element of the website (e.g. under 'Gruppen' I see 'Deine Gruppen' and 'Empfohlene Gruppen'.
- Yes - The system is quick and results are what I expect.

Match between system and real world

- Yes - Words used on buttons and near icons are understandable for the user and connected to the action of the button/icon.
- Yes - Wording seems simple and understandable, the order is logical as well, with upcoming events / my group on top, as those are probably most relevant for you.
- No - The person who will contact me is the "administrator"?
- No – "Events near me" - it is not clear where I am - is this derived from profile information or the device location?
- No – Activities, for which I have already registered, should be more distinguishable at first view.

User control and freedom

- Yes - Everywhere you can go back to the previous page using a back arrow, pop-ups can be closed using a cross, you can clear a search in the groups with a separate cross.
- Yes - After entering a group you can easily get back to the group overview. Getting back to the homepage is possible from every screen.
- Yes - This is a simple and straightforward interface. Back key works.

Consistency and standards

- Yes - Same words have same meaning.
- Yes - Same words (groups and events) are used consistently.
- Yes - It is very consistent.
- No - The use of "Dashboard" instead of "Home" is confusing.
- No – Blog icon looks like W-Lan signal strength icon.
- No – Show more ... sometimes displayed as text with link, in one place as button.

Error prevention

- Yes –Search bar shows an example text in grey of what the user might use the search bar for.
- Yes - You can see group and event details before you join, which prevents accidental signing up. For the events you also see information like time and date beforehand.
- Yes - I made no mistakes.

Recognition rather than recall

- No - Can be improved by using breadcrumbs to show user clear overview of how he or she navigated.
- Yes - Events show the organizing group as a link, so you can 'recognize' that.
- Yes - Even though the content is supplied from outside clubs and therefore is "unexpected" this is only where expected.

Flexibility and efficiency of use

- No - No accelerators, users cannot skip actions.
- No -Maybe there could be an option of 'favouring' an event so that you for example always see the walking football in your list and do not always have to sign up, but rather sign off if you can't join.
- Yes - The interface is very straightforward and no need to have shortcuts.

Aesthetics and minimalistic design

- Yes - Use of icons and short excepts of information. Good.
- Yes - The short descriptions only show basic information and if you want more info you can get more details.
- Yes - The design is not confusing.

Help users recognize, diagnose and recover from errors

- Not applicable, as not implemented yet.

Help and documentation

- Not applicable, as not implemented yet.

3.2   User Testing in the Netherlands

In the Netherlands the wireframes were tested with 5 trainers, of which two were senior volunteer trainers without a professional sports background and three trainers with a professional sports background. All trainers have followed an OldStars course, which means they are familiar with the vision of OldStars and the ASM method. The trainers were visited on their location of choice and asked to perform the three tasks while verbalizing their thoughts.

Professional/volunteer	Sport	More information
Professional background	OldStars Table Tennis	Trainer is at the end of a pilot phase with OldStars Table Tennis.
Professional background	OldStars Tennis	Also owner of a sports hall.
Professional background	OldStars Walking Handball	Trains 5 different sport groups.
Senior volunteer	OldStars Walking Football	Senior volunteer, who also shares experiences with other trainers.
Senior volunteer	OldStars Walking Basketball	In the starting phase of setting up an OldStars group.

Differences between professional and senior volunteer
During a test we found a difference between professional trainers vs. senior volunteer trainers. The main differences are:

- **Wanted functionalities**
  Professional trainers value functionalities differently than senior volunteer trainers, because of the amount of groups they train.

- **Usability**
  The professional trainers were more known with using digital appliances, which made them faster in executing the tasks than the senior volunteer trainers.

- **Willingness to share training material**
  Senior volunteers were happy to learn more from others. Professional trainers were careful with what kind of training material they would share as their training methods are what makes them unique as a trainer. Although they are willing to share, they will be careful with what they are going to share and take into account how fast they can come up with different materials to stay valuable as a paid trainer.

Functionalities of SALSA
- **Using SALSA as a communication tool towards the group**
  Senior volunteer trainers did not see the added value of using SALSA as a communication tool as they already used WhatsApp. Using SALSA for communication would lead to an additional communication channel, which need to be used by their participants. WhatsApp is seen by the senior volunteer trainers as the best communication method with participants as mobile phones are more accessible than email or websites for which you need to log in. Professional trainers did see the benefit as they train multiple groups and reactions in WhatsApp group messages can be confusing. A list of who is coming or not, prevents scrolling up and down to see who is not coming to a training. Besides this, it filters information that is shared in group chats, that are not valuable for the professional trainer to know about. Professional trainers also raised questions on who receives certain questions. E.g. questions about the location or schedule can go to the secretary of the club, but questions on content of the trainer can go to the trainer him or herself.

Next to this, professional trainers wanted to be able to access SALSA on their mobile phone when they are at the field or sports hall to send messages or see who is joining.

- **Managing group function**
  Managing groups is especially interesting for trainers who train multiple groups. Senior volunteers often only have one group, which makes them question why they need such a function. Professional trainers or coaches of municipalities often train multiple groups and sometimes rotate with other trainers on who trains the group. It is therefore convenient to see for a specific training who is joining or not, and that you can see information for the specific training who are giving and have a dashboard for the group in general with co-trainers. Besides that, this is valuable for training preparation, it also gives social control on which participant has not showed up a couple of time to check how this participant is doing.

- **Missing function: see and approach trainers in the neighbourhood to join forces**
  Trainers indicated that it would be nice to see which Walking Sports trainers are in the neighbourhood to join forces to rent a sports hall, organize an event, organize a fittest, get EHBO trainings etcetera. If they join forces, they can share costs, to meet the needs of their participants or their needs as a trainer for seniors.

Create a new group
The following should be added to make this task more suited for the process of trainers:

- Clickable link to the sports club.
- Add address of the sports club.
- The blue button of creating a new group was not spotted. Trainers would go to the menu on the side to add an event.

Create a new event
In a survey among 24 trainers seeing how many participants are joining selected by 9 trainers and an agenda with relevant workshops such as EHBO by 12 trainers. The following elements should be changed to make this task more suited for the process of trainers:

- The wording of event is unclear. Trainers prefer to use the word 'training'.
- The blue button of create a new event is not spotted by 4 out of 5 trainers. They would go to the menu on the side to add an event.
- Per training the trainers want to be able to add who is standing in front of the group, as this could make a big difference to participants.
- Field of indoor or outdoor should be added to indicate of the event takes place indoors or outdoors. This can make a difference for seniors, especially in relation to the weather.

Search for Walking Sports training material
For searching exercise or training material the following elements should be changed. In a survey among 24 OldStars trainers videos on age-friendly exercises were selected by 12 trainers as a wanted function. 14 trainers selected staying updated on age-friendly sports by 14 trainers, sharing training material was selected by 8 trainers, sharing experiences among trainers by 9 trainers.

- **Length of the video clips**
  All 5 trainers preferred short video clips of exercises of 5s to 30s instead of searching for full training templates. The exercises are used as inspiration. Therefore, short clips suffice.

- **Tagging system for video clips.**

For OldStars Sport they suggested a tagging system of: warming-up, movement-form, game-form (in Dutch: warming-up, beweegvorm & spelvorm) . As this is connected to the OldStars and ASM vision of how trainings are structured, for other sports a different tagging system is necessary. E.g. Aikido for senior does not work in the same manner, while this is an age-friendly sport. Therefore, it is suggested that SALSA integrates a flexible tagging system for categorizing training video clips.

- **Learn from other sports**
  Trainers were aware that they could learn from other sports. Therefore, trainers suggest to also show clips from other Walking Sports or that you can click on an icon to see more exercises for this sport.

Conclusions

The main conclusion of this user-test is that the SALSA consortium must be aware that there are two different types of trainers with different needs: professional trainer and the senior volunteer trainer. It would make sense to have the option to not use certain functions. Nevertheless, SALSA seems to contain all required functions at the moment, except for the wish to link with other trainers in the neighbourhood to join forces to save costs.

3.3   User Testing in Austria

Users 55 +

In Austria three participants took part in the first round of wireframes testing, one male (57) and two females (57 and 71). The testing sessions took place at the homes of the participants. All three participants were able to finish the tasks easily.

Figure 5:: User testing in Austria

Observer's notes on the test sessions:

LFTL Participant 1: Male, 57 years old. The two tasks were completed easy. The user liked the elements on the side navigation board and they were clearly understandable. The text "show more groups" was difficult to recognise for the user at first view. The font size was too small for an enjoyable read. The short descriptions, details and language were clear and understandable but sometimes the headlines of

the groups switched to an English version and showed a lack of consistency. The user expects further confirmation after joining the group. Returning to the main page was possible. The search function was used to search for "walking football groups" in Linz. The user expects some kind of arrows to navigate through further events (moving left and right). I observed the preference of the user to click first on details and in the next step to attend the activity or group. For the user it was recognisable if he already joined a group. Especially for group activities the user could not identify if the activities are only a one-time activity or a weekly activity - so further information is needed here. The task "searching for" and "join the group" was feasible for the user but there was no possibility to unsubscribe from a group. The user liked the short basic information and important facts at the first view and the option to open more details.

LFTL Participant 2: Female, 57 years old. The user performed the two tasks quickly and without errors. The user suggested to highlight already attended groups in a visual way. The font size is very small. The short description, detail information and language were clear and understandable. After joining the group, the user expected to return automatically to the main page, but using the side board navigation was feasible. The detail information about the date, duration, meeting point and price is not shown on group activities. The user wants to be further informed if the activity takes place outdoor or indoor. It is important for her that the shown picture reconciles with the further activity (facts like: male/female; indoor/outdoor). Furthermore, information would be important for the user to see for example "how many participants are planned for the group or event" (limited number of participants); "which participants are planned for the group or event" (only male, female or mixed-participants; for beginners, intermediate or advanced players) Before joining the activity or group the user preferred to click at first on the detail information. Furthermore, the user could not identify how often the group activities will take place. (only once or on a weekly basis). The user liked the simple structure of the page.

LFTL Participant 3: Female, 71 years old. Task 1, search for walking football group: clicked instantly on "Groups" in side bar, found yoga group and liked it better, found football group and thought, that she had finished the task (instead it was a group, she was already a member of), eventually found the search field and clicked on it, got three results. Task 2: Join Walking Football group: Clicked on "More Details" first and found the join group button. Was slightly confused by the list of events and the possibility to register for the events. After some time clicked on the join button, got the confirmation and was satisfied.

Most important issues/findings
Some texts/links like "Show more ..." are too small and not so easy to find.
In some places a confirmation of actions is required (e.g. after joining a group).
More information on events requested, e.g. indoors/outdoors, one-time event/multi-event, required fitness level, ...

Trainers

In Austria the wireframes were tested with one trainer (male, 82 years old). Again, the user did well and could finish the first two tasks quite easily. The third task could not be completed as he did not find the relevant section (Trainings). Again some functionalities/texts were a bit hard to find due to small font size.

Observer's notes on the test sessions:

LFTL Participant 3: Male, 82 years old. The user needed some time to find the "create a group" option. The participant suggested a kind of template for the group description. This would be helpful not to forget important information like (is the group for beginners of for advanced players). For the user it was not clear if the mentioned address is also the meeting point. The function "create a new event" has been

found by the user. In the German version "Bezeichnung" should be changed to "Name der Veranstaltung". The other options were clear for the user. The user didn´t find the search function for the training material. He was looking on the main page.

Most important issues/findings
Some texts/links like "Create a new …" are too small and not so easy to find.
More support/guidelines on how to describe a group when creating a new group would be helpful.
Address in group description: Is this the training site?
Some wording problems, better translation needed.

3.4   User Testing in Switzerland

In addition to the employees of the terzStiftung, the test participants included two primary end users. Both were men aged 60 and 79. They were shown the clickable mock-ups and the related tasks on the computer. Both worked on the tasks independently. Participant 1 was able to solve the tasks without any problems. He praised the good menu navigation. Participant 2 was also able to solve the tasks well. He liked the design of the website. He first clicked on the details before joining the group. He suggested to have a section on the main page that says something like "my groups".

4   Pre-Prototype Testing SALSA Fun

4.1   Round 1

The feedback gathered with the wireframes was processed and evaluated in separate project meetings by LFTL and NET. Some issues were minor, some feedback, especially from Oldstars trainers, resulted in new functionalities. Start of August a new version of SALSA Fun was available, this already being a first fully functional and designed prototype.
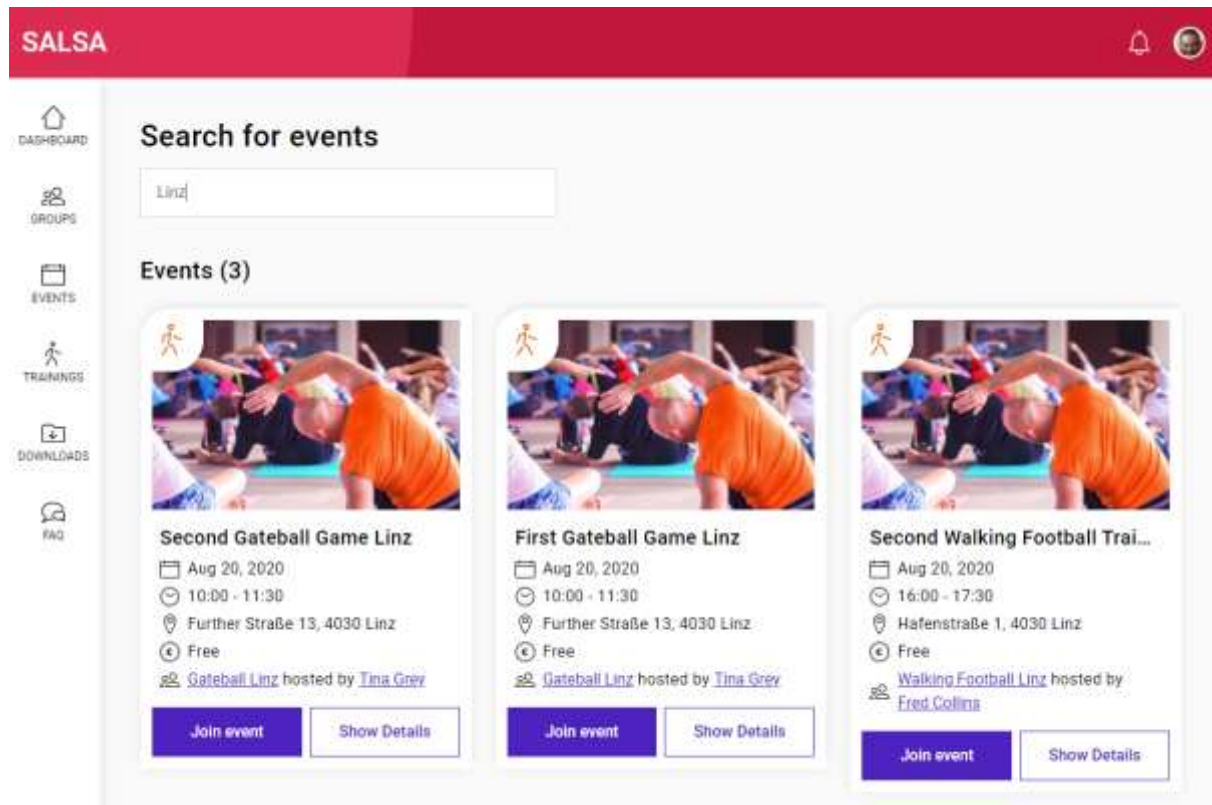
Figure 5: First prototype SALSA Fun, used in the 2nd test round

Again this version was to be tested with end users, people 55+ and sports trainers. This time, there was no heuristic evaluation by experts. We asked the participants a few personal background questions, mainly to assess their digital literacy. After the tests, they were additionally asked to fill out the validated User Experience Questionnaire (UEQ).

The Covid-19 situation however made this testing round more challenging, as it was not that easy to involve (vulnerable) participants face-to-face. So many of the test sessions were done remotely using conferencing tools.

These were the tasks given to the participants:

Tasks for users 55+:

1. *Task – Register.*
2. *Task – Log in.*
3. *Task – Search for events in your vicinity.*
4. *Task – Search for a group in your vicinity and inform yourself about it.*
5. *Task – Join the group and check the number of members.*
6. *Task – Subscribe for an activity of that group.*

Tasks for trainers:

1. *Task – Identify which groups you are a trainer of and look for more details on that group.*
2. *Task – Check your upcoming events as trainer and look for more details on that event.*
3. *Task - Create a new event.*
4. *Task – Go to the FAQ section and read a topic of your choice.*
5. *Task – Go to the download section and look for provided material.*
6. *Task - Create a new group.*

4.1.1   User Testing in the Netherlands

In the Netherlands the prototype was tested with 5 trainers (age from 64 to 72), all male, four of them with a professional sports background. They use either desktops or laptops, mostly on a daily basis. Mobile devices like smartphone (3) and tablets (4) are also used on a daily basis (4). Three of them also use social platforms like WhatsApp and Facebook. So it is fair to assume, that these participants know the needs of walking sports trainers and have basic digital competences.

The participants were quite positive about the concept and the look & feel. The tasks could be completed easily. Most of the feedback are suggestions of improvement in terms of additional functionalities of the platform. Minor issues concern wording, which will be fixed by the Dutch partner NFE, these issues are not listed here. Also one technical issue was reported.

It was indicated by the participating trainers in this round, that it is most important to them to communicate with other trainers and find/share training material. Group management or activities with registration options are of less interest to them, as events normally are serial events (same day, same location). Dedicated (co-)trainer groups and publishing special events are interesting though.

Technical Issues
When creating a group, the zip code option does not work.

Usability Issues
Introduction text / video explanation on how to make trainings/exercises missing.
Automatic filtering of search results based on location (show groups, activities in vicinity).
Change sequence of buttons: First trainings, then exercises.
When adding information for training/exercise, provide as few text boxes as possible (but more small steps). Similar to Vinted. Current setup too long and not desirable.
2 or 3 extra buttons in the dashboard to directly go to create a training or exercises.
How do I learn about new stuff? E.g. when there is a new document, exercise video, …?

Suggestions for Improvement/Additional Functionalities
More information on landing page before log in. Also concerning responsibilities, contacts: who is the provider/admin?
Possibility to easily contact a responsible person (administrator) or organization like NFE, Oldstars, …
Change photo in background. Mix of ladies and gentlemen in the picture.
Separate dashboard from general information.
How can a trainer remove a person from a group/activity?
Filter options for searching exercises and trainings missing.
When creating own content, the trainers indicate that the text blocks may be shorter and that the focus should be on video and images.

Possibility to download training forms.
Add possibility to create serial events (weekly, monthly, ...).
Drag&Drop functionality to compose a training from content found on SALSA.
Add information about what the ASM / OldStars vision is for training and exercises.
Add ethical code for platform. How do users deal with this platform?

The UEQ yields a good result and reflects the general positive feedback of the participants. All values are in the green (good) area. Due to the small sample number the confidence interval is rather big for some dimensions, especially for perspicuity, efficiency and novelty. Which also hints that the given answers differed noticeably among the participants. Attractiveness, dependability and stimulation rated high and good, even considering the confidence interval.

The dimension perspicuity (How easy is it to get familiar with the product and to learn how to use it?) is comparatively low and may be caused by wording problems and the lack of supportive information.
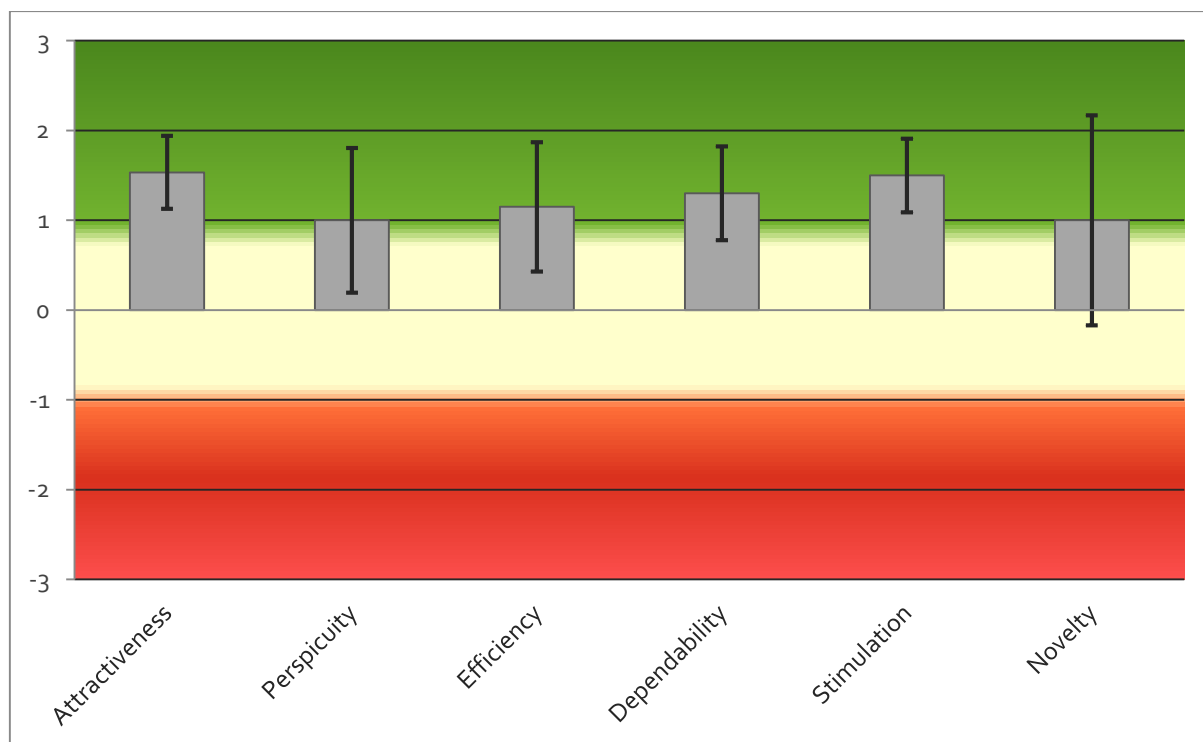
Figure 6: Results of UEQ with trainers in the NL on SALSA Fun, n=5

To get a better picture on the quality of a product it may be helpful to compare the measured user experience of the product to results of other established products, for example from a benchmark data set containing quite different typical products. The UEQ offers a benchmark (it contains at the moment of usage the data of 452 product evaluations).

The benchmark classifies a product into 5 categories (per scale):
  • Excellent: In the range of the 10% best results.

  • Good: 10% of the results in the benchmark data set are better and 75% of the results are worse.

  • Above average: 25% of the results in the benchmark are better than the result for the evaluated product, 50% of the results are worse.

- Below average: 50% of the results in the benchmark are better than the result for the evaluated product, 25% of the results are worse.

- Bad: In the range of the 25% worst results.

The evaluation shows, that SALSA Fun is doing fine in all dimensions, performing 'good' or 'above average'. The exception is perspicuity, again pointing out, that some more efforts have to be put there, to help users get familiar quicker.
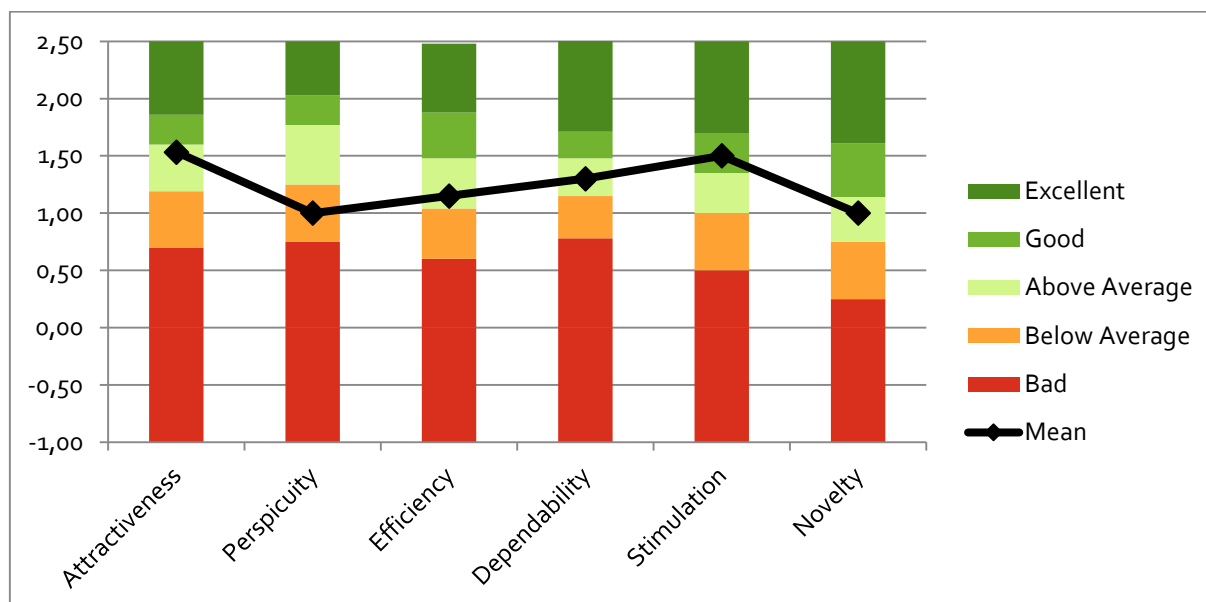

Figure 7: Benchmark results of SALSA Fun, compared to 452 products

4.1.2   User Testing in Austria

The following feedback was collected with one professional sports trainer in Austria.

Usability Issues
Creating a group was not successful, as there was no hint that uploading a picture is mandatory.
Button 'New Group' was hard to find.
Activity list is empty, because of missing content. When searching with ZIP code, it would be helpful at this stage to display results from other areas also.
Icon: most trainers did not associate the FAQ area with the speech bubble icon.

Suggestions for Improvement/Additional Functionalities
Trainers want a way to invite people and to import email and phone numbers, e.g. over established channels like WhatsApp.
Allow serial events.
Allow definition of min and max number of participants for an event.
More functionalities for member management needed.
Chat functionalities missing.
Waiting list functionality missing when an event is already fully booked (compare app "No Excuse").

Detailed remarks on waiting list functionality by sports trainer:

- There is no waiting list when a course is fully booked and that people can register on the waiting list.

- If someone drops out, which happens again and again, a push message is sent to the first person on the waiting list with the content "You are now registered", short decision whether he/she will Take the place Yes or No. If no, the message is sent to the second person on the waiting list etc. until the course is full again.

- This should also be possible in case of short-term or long-term cancellations, depending on how long in advance you have to register for a course (often 4 days). It is essential that the course is refilled in this way. In the case of yoga courses, people often decide to take part within an hour if a place becomes available.

- It is also good if there is a minimum/maximum number of participants, e.g. minimum number for a course is 6 participants or maximum 10 or 12 participants.

- And that in between there is a message to the participants, "Sorry, the course does not take place" or "Registration only possible for the next time".

- Trainer likes the visual presentation very much, namely that a green or red bar indicates that there are still many places available or, if it is red, that it is already short or full.

4.1.3   User Testing in Switzerland
In Switzerland in total 8 persons participated in this second test round, 3 trainers 5 users 55+.

Trainers

The participating trainers (2 females, one male, age from 57 to 61) all use smartphones and laptops on a daily basis for standard applications. None of these trainers is a professional sports trainer but experienced in guiding senior sports groups (gymnastics, fitness training).

General feedback: "Attractive page with very good overview. Design is appealing, use case meaningful."

Usability Issues
Create a new group was expected on the "Groups" page and not on the Dashboard.
User was confused about the division of event and group. Suggestion: Courses that follow a course plan and where the meetings take place regularly and single events, such as separate training or competition.
It was not clear whether the trainers page is an exchange platform for trainers or the trainer 'backend' for the primary users page.

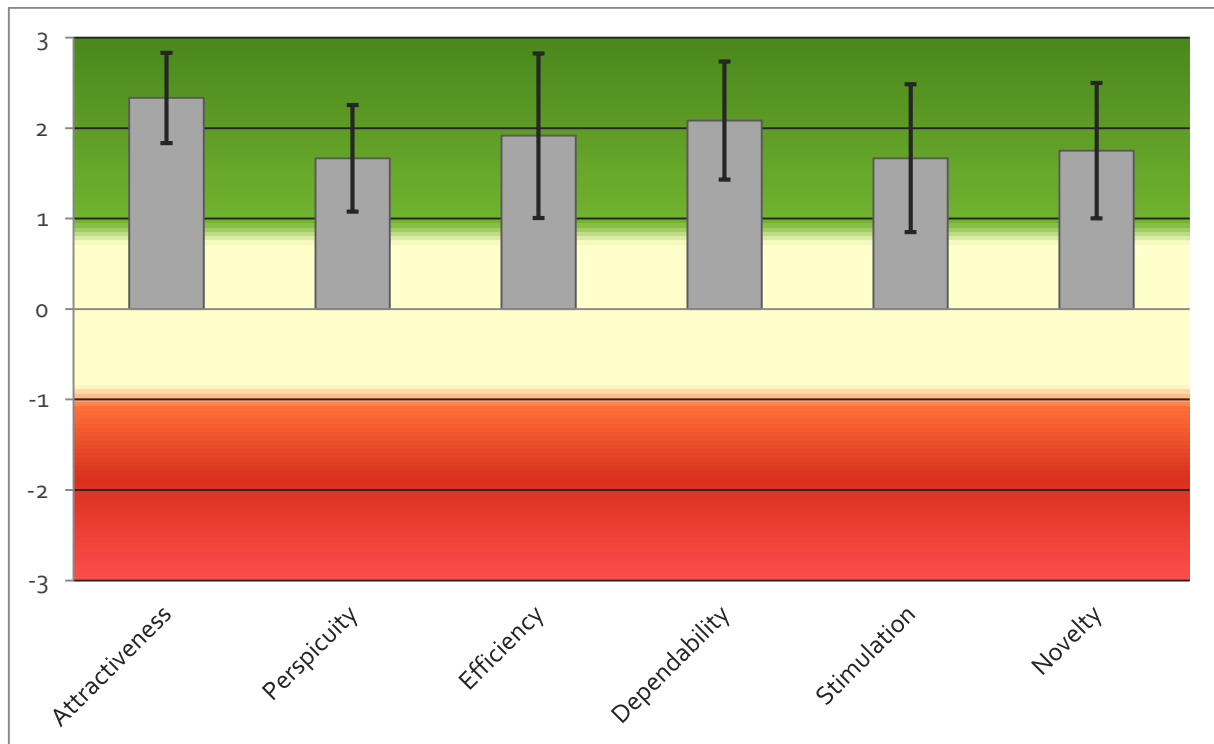| Suggestions for Improvement/Additional Functionalities |
| --- |
| When creating new group/event, give possibility to select image from a list. |
| Add the function "Forum" where trainers can ask for assistance and discuss exercises. |
| Print function for exercises/trainings. |

Figure 8: Results of UEQ with trainers in Switzerland on SALSA Fun, n=3

The results as can be seen in the figure above are very good, even considering the large confidence intervals.

Users 55+

From the five participants three were female, two male, the age ranging from 57 to 65. All but one use a laptop on a daily basis, for work or private matters. All have smartphones, which they use daily or at least several times a week. WhatsApp is used by four, two are also on facebook. The digital competencies are therefore at least basic.

| Usability Issues |
|---|
| Ask to repeat password when registering and offer possibility to display written password. |
| When searching for a group or event, show also those who are near the city not only those exactly in the city. |
| Filter function too sensitive: E.g. when searching for 'walking', Nordic-Walking is not shown. When entering '4020 Linz' no results are found. |
| Use more confirmation pop-ups, e.g. when joining an event/activity, display a pop-up "successfully joined". |
| The currency in Switzerland is CHF not Euro. Needs to be changed. |
| 'Create Group' and 'Create Event' should move to the respective pages. |
| Users and trainers are shown twice (group detail view). |
| Single events/serial events? |
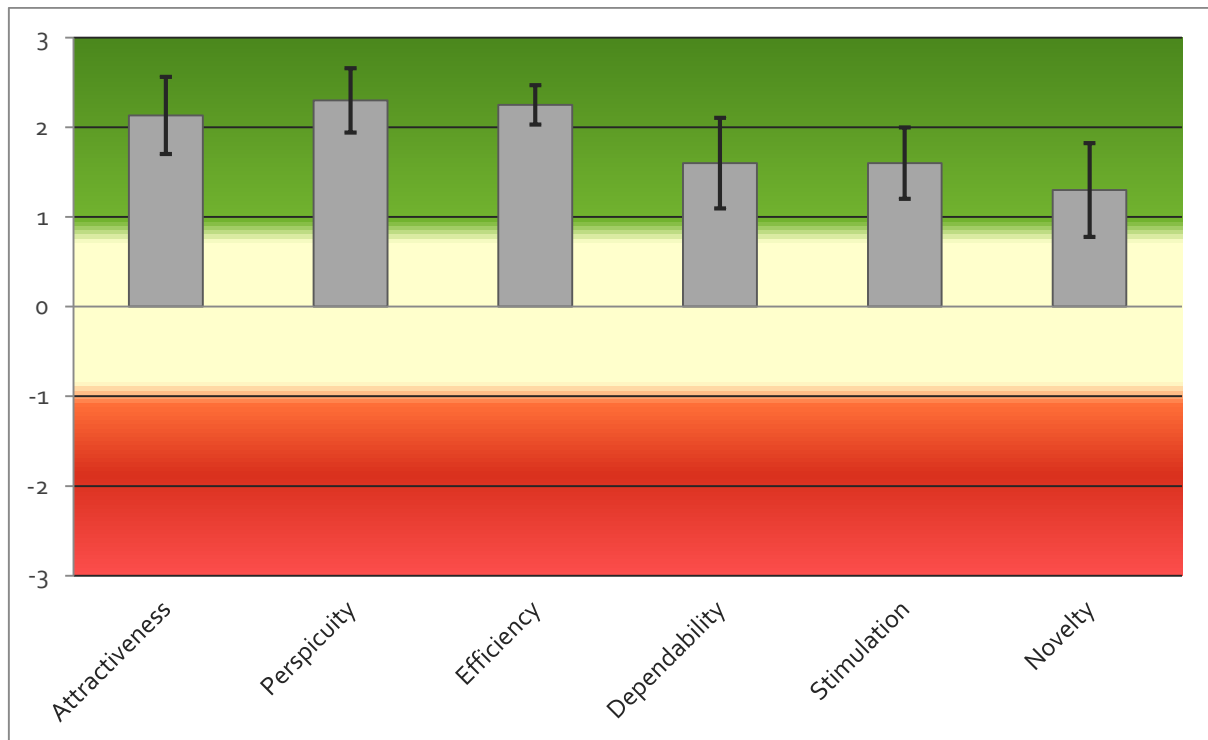
Figure 9: Results of UEQ with users 55+ in Switzerland on SALSA Fun, n=5

Again, very good results and even the confidence interval is acceptable for most of the dimensions.

4.2   Round 2

4.2.1   User Testing in the Netherlands
In January 2021 another round of user testing was done. This time we asked a number of Oldstars trainers (n=4) in the Netherlands only, as SALSA Fun is largely influenced by the needs of Walking Sports trainers and we wanted to get feedback of active and experienced trainers.

Again we had prepared tasks to do in order to check the newly implemented functionalities.

Tasks for trainers:

- *Log in with the provided trainer credentials.*
- *Check your groups (you are the trainer of) and edit the group information.*
- *Remove a member of your group.*
- *Add a comment to your group.*
- *Add an event.*
- *Search for a warm up exercise and use the filter function.*
- *Search for a training plan and use the filter function.*
- *Add an exercise and add a youtube link for it.*
- *Explore freely.*

Trainer 1

General impressions

- It would be nice to make trainings examples and forms more prominent. That is what you are aiming for, searching for. It would be nice if it were visible right away.

- The style of the platform could be more attractive, less clinical. They give the example of OldStars website.
- Please add in dashboard a short description about the platform. How is it used, maybe a tutorial.

<u>Testing predetermined tasks</u>

- Login went very easy and quick.
- It took quite some time to fill in the group activity.
- Loading the remarks took quite some time, this could also be the internet connection.
- The emoticons could be more refined and focussed on sports and activities. At this point there are a lot of emoticons which are not applicable for their situation.
- If you are in the menu of activity you can search for activities, but not add an activity. They found this very confusing. After a long search they found the option to add an activity in the dashboard menu.
- To add a photo in the composing of an activity is unnecessary.
- When searching for warming up exercises the search subject *training* is not clear.
- Adding filters in the search option is useful. They want to select warming-up or game form or ground exercises from ASM trainings. A drop-down menu was mentioned.
- A reaction to a trainings form is not a valuable addition.
- When adjusting the trainings form a short description would be useful and also to add videos and photos/images.
- When adding an exercise, it is not clear when adding the duration. Is this hours or minutes.
- To obligate a trainer to fill in the size of the group is not needed.
- When adding an exercise, it was not possible to upload a youtube link. Information about adding a video was missing (for instance what kind of file format, size, etc).

Trainer 2

Comment by the researcher: *The digital literacy of this participant was low. Logging in took a lot of time, not only in teams but also on the platform. We were forced to end the exercise half way. He was given the login details to walk through the steps on his own, but this also didn't work. In a normal situation we would be alongside him to guide him through the first steps, but COVID made this impossible. This is a con of taking these tests online.*

<u>General impressions</u>

- ASM platform is mentioned as an example of a comparable platform.
- First impression is good. Not really enticing.

<u>Testing predetermined tasks</u>

- Group page is easy to find. Adjusting the page isn't that clear. It is mentioned *show details*, this is not logical when wanting to adjust something.
- After finding the page where you can adjust the page it's easy to follow.
- A dropdown filter menu is useful. There need to be more filter functions.

This was how far we good get with this participant.

Trainer 3&4

Comment by the researcher: *We interviewed these two participants together because they work together and are both trainers.*

<u>General impressions</u>

- The first impression is good, it is very clear how to use it.

- It would be very useful to work with such a platform for inspiration, ideas for warming-up and other exercises.
- It would be nice to have a lot of set trainings and exercises here for table tennis and other sports. Like a library of trainings. Per sport a set of documents / films / photos / images.
- They mention the people need to be willing to make exercises. They are a bit in doubt if everyone is willing to do this.
- Safety is very important. Data control of all the files.
- Participants will need to give their consent to be filmed. They are not sure if everyone will do this.

Testing predetermined tasks

- Very easy to find the login option.
- When editing a group, the details of contacts of all the groups are missing. And of other trainers. It would be nice to have a contact list of all your participants and other trainers.
- To add a comment, they found it difficult to see where you can do this. It is all the way down, very low on the page. It needs to be placed higher on the page and also it needs to be clear where you are responding to. Are you responding to a group? Or a training? It is not clear
- They do not use emoticons. Love and angry is very strange to use in this type of group. They see this as a formal atmosphere and emoticons are very informal.
- They wanted to add an activity on the activity page, but they didn't find the option here. Only at dashboard. This feels a bit strange. Maybe on both pages?
- When they find the option everything is clear.
- When filling in the postal code you cannot fill in the two letters. You can only fill in the 4 numbers. In the Netherlands you need to fill in 4 numbers and 2 letters.
- It is a bit strange to add a photo to the activity. It would be very useful to add the size of the image (how big can the image be).
- When clicking on the costs for the activity, everything is stopped. They needed to logging in again. It could be their internet connection or a bug.
- Warming-up is easy to find. The filter function is a bit more difficult to find. It could be useful to add a symbol of a magnifying glass.
- Also the options in the filter function or not clear. Name is not very often used. Also the margin of error is very high.
- The filter function needs to include a dropdown option. It needs to include options of warming up, different spots, maybe that you can search / tick on a few boxes for searching.
- Searching for number of participants and duration is not useful.
- For training plans there is a drop down filter menu, it is more easy to use.
- Maybe include YouTube for training plans? To add films would be very heavy, size wise. Would the server be able to run smoothly when everyone is downloading heavy documents?
- When adding an exercise, it is clear what to do.
- For the duration they don't know if its hours or minutes.
- Intro is also strange to fill in here.
- The different options at the training plan are a bit strange. They never use this with their trainings. It would be useful to add videos / images or photos.
- Video is essential to explain the training.

5   Wireframes Testing SALSA Health

In August 2020 the first official round of SALSA Health wireframes testing started. Feedback on the first sketches had been gathered before, but not in a coordinated and structured way as the focus then lay on SALSA Fun. For the test sessions in August, a click path was defined to allow basic navigation. The wireframes were not translated (to Dutch or German) and we only asked physiotherapists to do the prepared tasks. Still being wireframes without design we nonetheless used the UEQ for a first assessment of user experience.
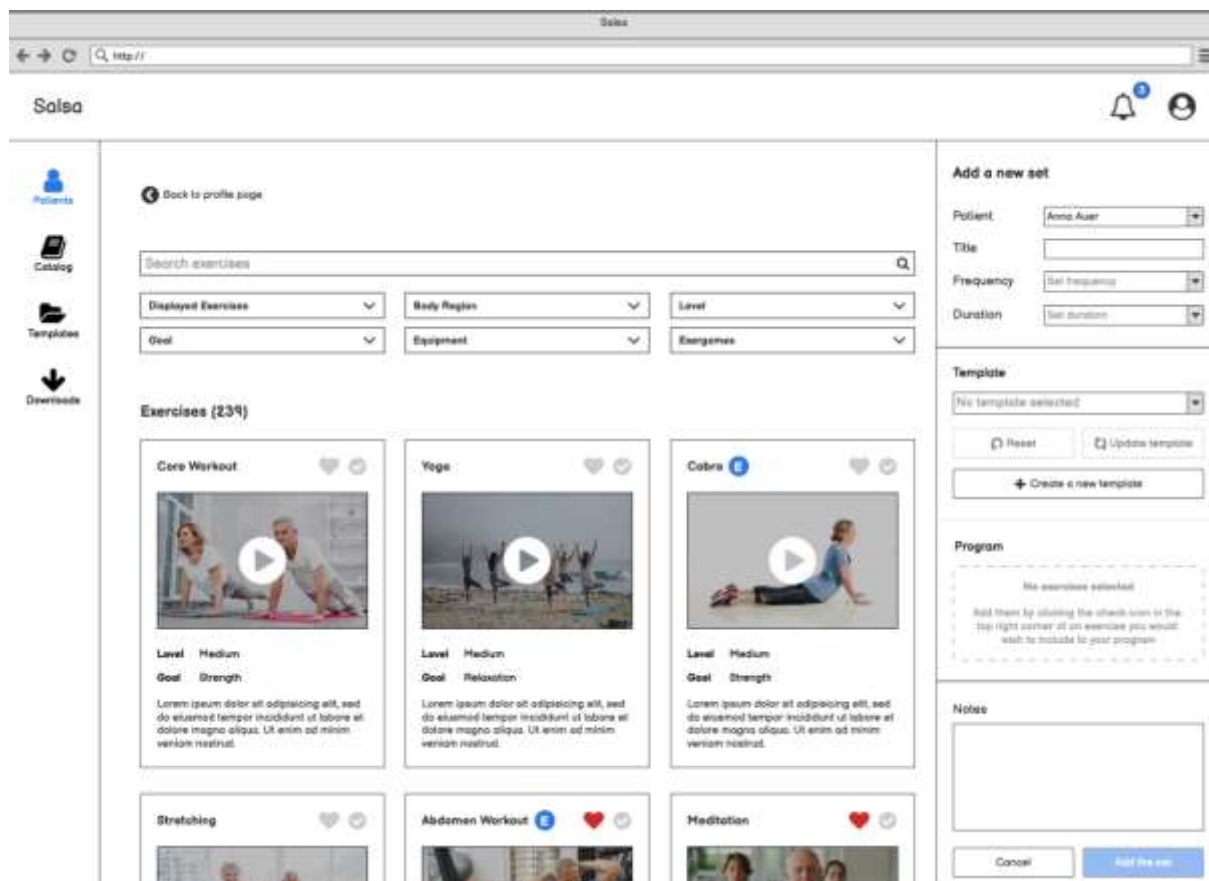

Figure 10: Wireframes SALSA Health used in user testing with physiotherapists

Tasks for physiotherapists:

1.   *Task – Add new patient - Fill in details.*
2.   *Task –* Assign a new exercise set - Apply some filters options (medium, beginner, relaxation, …).
3.   *Task – Add yoga and meditation to the program – Add Set.*
4.   *Task – Go back to patients' overview.*

5.1   Results

In total, 10 physiotherapists participated in Austria, the Netherlands and Switzerland (4 female, 6 male, the age ranging from 26 to 67). In general, the feedback was positive, most participants like the idea of having videos of exercises, which can be given to patients for home training. (The feedback of therapists in Austria and Switzerland being more positive than in the Netherlands.) It's important to remark, that not the whole app with all functionalities has been tested, but that the focus was on the video exercise functionality.

Doing the tasks was not easy for some participants, because the mocked functionality of the wireframes was very limited, which was confusing to them. Other participants were not confused by this and commented more generally on what they saw in the wireframes, liked, disliked, missed, ...

Some problems were caused by wording and the fact, that everything was in English, like 'template' or 'set'. This can easily be solved, when the prototype is localized.

One major topic that needs to be discussed are GDPR issues:

- It must be clarified whether and which sensitive data may be entered by the patient (software provider must meet GDPR)
- Notes on patients have a documentation character (legal documentation obligation, notes on the course of treatment are considered sensitive data)
- It should be considered whether a minimum standard for reporting and documentation should be introduced in compliance with the GDPR - thus, the practice would not need an extra patient documentation software.
- The functionality to share patient data among different physiotherapists needs to be checked with the legislation.

Usability Issues
Wording: Button 'Enter new patient' should be renamed to e.g. 'Save' on template for patient data.
Area with filter options should be named 'Filter Options'.
Patient overview page: Patient info is too big compared. Make it smaller and enlarge the other areas like assigned program, ...
Filter functionality unclear: chooses exercise video first and wants to apply filters afterwards.
Assigned programs listed just by name is not sufficient - It was not understood, that you can (later) click on an assigned program to see the contained exercises.
The navigation bar on the left is not visible enough.
Filters in exercise selection are unclear (What are displayed exercises?).
What do the minutes for exercises mean - does the video for yoga take so long and contain different exercises or should one exercise be performed several times?
What does the blue E mean e.g. for the exercise Cobra?
What are templates?

Suggestions for Improvement/Additional Functionalities/Other Remarks
Patients section should also include: health status, pain status, medication, blood pressure, other risk factors...
Possibility to film patients doing the exercises (or make a photo) and store this material with the patient data. Also send them to patients for exercising at home.
The functionality to share patient data among different physiotherapists seems troublesome, needs to be checked with the EU legislation.
A possible suggestion for exercise filters would be "normal exercise, exercise with equipment, exergames"; level should be displayed as subcategory; at least filters should be rearranged (displayed exercises should not be listed first).
Set should be differentiated and e.g. frequency should not be set identically for the whole set (intensity, frequency very important).
It should be possible to insert assessments as feedback questions, there should be more choice.
The goals should be reconsidered - very unspecific and no real goals, the mentioned four ones are categories.

"How was your exercise" - very unspecific, does not consider different aspects of exercises.
Notifications from patients should be visible for physiotherapists even later.
Statistics should be reconsidered - maybe instead of sessions it would be better to show a goal achievement (percentage), is a greater motivation factor.
A slider should be used for the feedback question instead of categories (metric scale) --> variability becomes more visible; qualitative demand for good and bad components better separated; for qualitative information, perhaps already a preselection should be given, so that patients only have to choose an answer.
The Physiotherapist would find it great if patients could also be asked a general question about goal achievement, e.g. goal is to achieve increased flexibility - "Has flexibility improved through training?" (Patient is able to assess and recognize the benefits of training - PT can react to feedback accordingly).
It should also be considered how this online care can be charged.
A possible filter could also be "symptom".
Consideration should be given to creating opportunities for patients who have problems with exercises to film themselves doing them and send these videos to PTs.
In addition to the repetitions, series should also be displayed during an exercise.
Standard items should be selected as feedback questions, preferably metric scale.

The UEQ yields a satisfying result and reflects the general positive feedback of the participants, especially in Austria and Switzerland. All values are in the green (good) area, except novelty.

The scales of the pragmatic quality (perspicuity, efficiency, dependability) which describe task related quality aspects, are the best rated dimensions with acceptable confidence intervals.

The confidence interval for novelty is very big, which can maybe be explained, that in the Netherlands digital solution for physiotherapy are more common than in Austria. The confidence interval is also rather big for the dimensions "attractiveness" and "stimulation". These dimensions should also not be overrated at this stage of development, as the tested version was only wireframed and not designed yet.
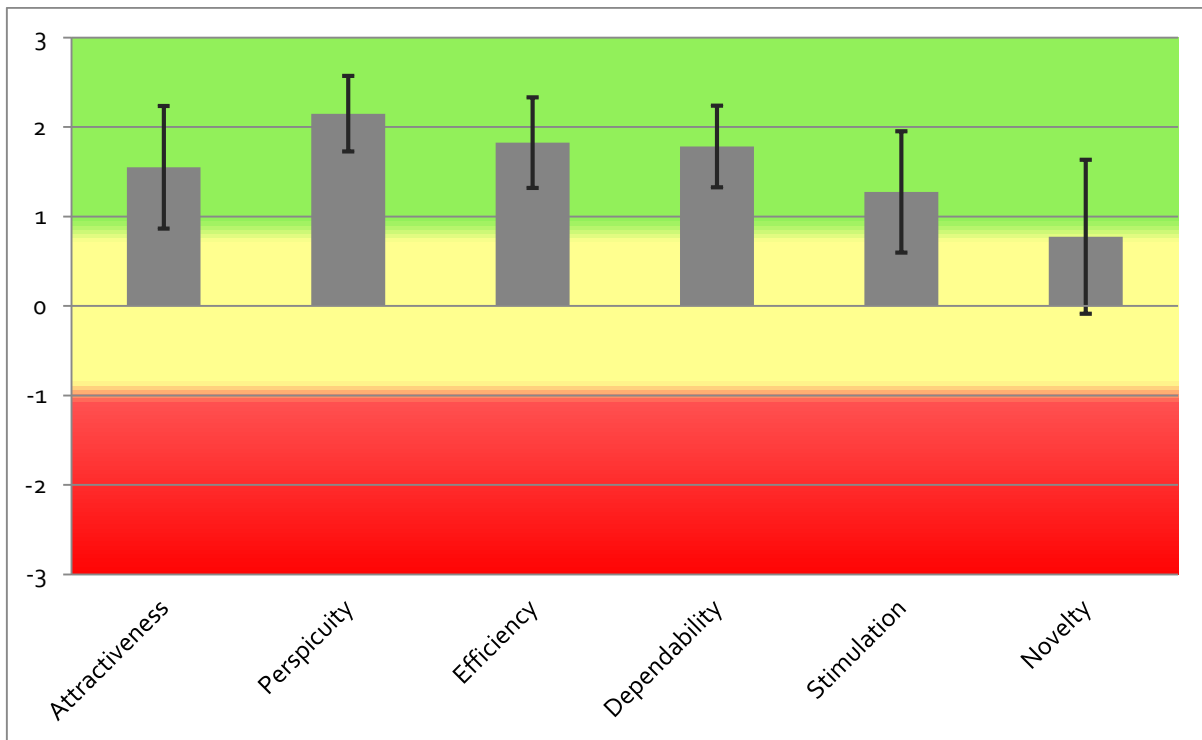
Figure 11: Results of UEQ with physiotherapists on SALSA Health, n=10

The values in the benchmark chart vary between 'above average' and even 'excellent', with the exception of 'novelty', being pretty average.
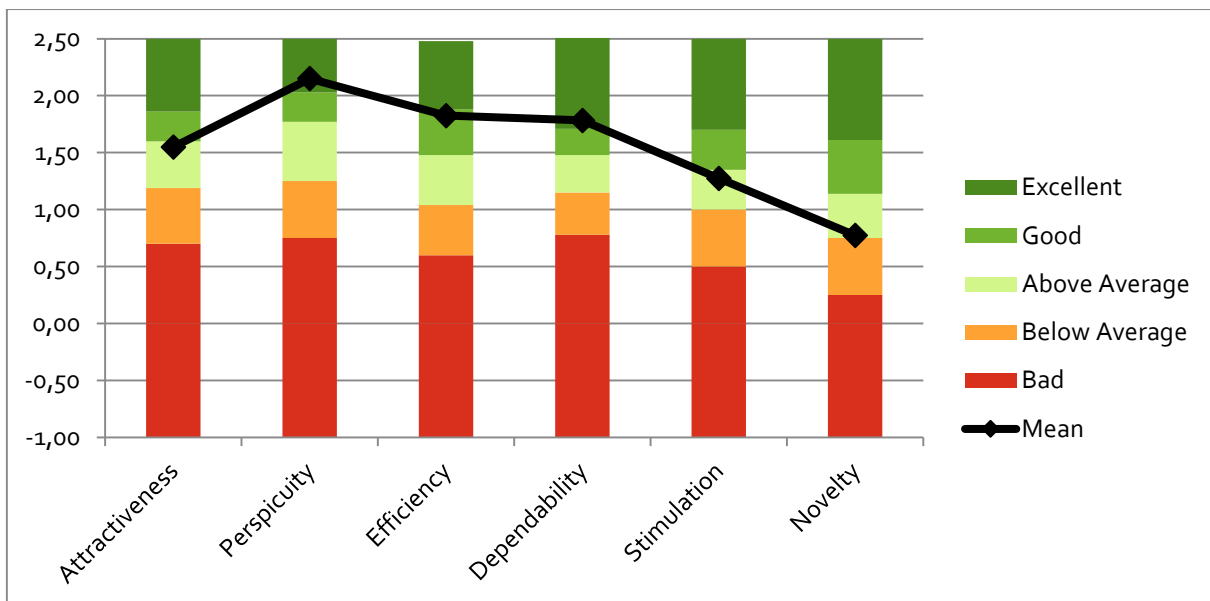

Figure 12: Benchmark results of SALSA Health, compared to 452 products

6   Pre-Prototype Testing SALSA Health

After the mid-term review meeting and in order to follow the recommendations given a number of corrective measures were implemented regarding SALSA Health. Instead of pursuing the initial approach, the consortium considered available alternative systems for tele-rehab as part of the SALSA Health application to expedite its implementation. As a result, the consortium revised the technological architecture of SALSA Health and decided to build on the readily available MIRA platform instead of developing a separate platform from scratch. The initial wireframes were not further elaborated, important feedback and findings of the pre-test phase however used in the following steps.

With this new approach, the testing on pre-prototypes of SALSA Health by the project consortium was ended, as newly developed functionalities were tested through a software verification and validation process already in place at Mira.

6.1   SALSA Health testing procedures

The test methods used at Mira were the following:

- Unit testing, in which certain test suites are included in the automated process of continuous integration / continuous deployment of the MIRA platform.
- Manual testing (including regression testing) of the advanced functionalities in the 'In testing environment' stage within the internal development process based on a Kanban methodology distributed on the following stages:

Backlog	Doing	Review	Ready for testing	In testing environment	Done

- Acceptance testing in Beta.
- Testing of critical functionalities in the stage preceding the deployment in production.
- Continuous monitoring of the solution in production through periodic reports provided by the infrastructure of the MIRA server solution.

Certain difficulties and technical problems (bugs) appeared during the implementation process of the new functionalities, an anticipated fact taking into account the complexity of the solution and the level of innovation (new technologies) involved in the technical development.

The problems and technical difficulties encountered were solved through several implementation interactions, followed by test iterations, thus leading to a stable and widely usable final product.

7   References

[Please put references directly in the document with footnotes and list them here as well, for long links to websites please use a link shortener service e.g. https://bitly.com/]

Why You Only Need to Test with 5 Users	https://goo.gl/X271D3
Online collaboration platform Trello	https://trello.com/
How to Conduct a Heuristic Evaluation	https://goo.gl/hqxqGM
Methods for Successful "Thinking out Loud" Procedures	http://bit.ly/397tFiO
UEQ - User Experience Questionnaire	https://goo.gl/a5jVtC

8   Appendix

8.1   User Testing on Prototypes – Manual

Please see separate document.

8.2   Informed Consent

Declaration of Consent

The project

The aim of the SALSA project is to motivate older adults (55+) to be more physically active and to support them in rehabilitation in case of health problems. For this purpose, SALSA provides a platform for organizations and associations to promote innovative group activities and events (fun) and for physiotherapists to offer therapeutic group activities (health). Furthermore, SALSA can also be used for tele-rehabilitation.

Your participation

Today we invite you to test a first prototype of SALSA Fun/Health. Your experience and feedback will help us to improve the designs. We will ask you some personal questions and to perform some typical tasks in the app. While you do this, we will take notes. You can cancel the test at any time without giving reasons.

Your data

All information that you provide during the interview or during testing will be processed and used anonymously as an overall result. Third parties (outside the project team) will not be given access to individual data sets.

Images that are created during the testing process are used to inform the public about the project and its progress (print media or internet, such as social media like Facebook, …).

Your consent

I have understood all the information regarding the research project and the test procedure and agree to participate under the above conditions. I allow the project consortium to use images on which I am recognizable for the purposes mentioned.

_____

Date   Name   Signature