

D2.4: Evaluation results of first functional prototype and updated requirements

Dissemination level: Public

Document type: Report

Version: 2.0.0

Date: August 27, 2018 (Original Version)

March 5, 2019 (Updated Version)

November 19, 2019 (This Version)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement #769553. This result only reflects the author's view and the EU is not responsible for any use that may be made of the information it contains.

Document Details

Project Number	769553
Project title	Council of Coaches
Title of deliverable	Evaluation results of first functional prototype and updated requirements
Due date of deliverable	August 31 st , 2018
Work package	WP2
Author(s)	Tessa Beinema (RRD), Harm op den Akker (RRD), Marijke Broekhuis (RRD), Gerwin Huizing (CMC), Lex van Velsen (RRD), Silke ter Stal (RRD), Jorien van Loon (CMC)
Reviewer(s)	Konstantina Kostopoulou (iSPRINT) (Original Version) Jorien van Loon (CMC) (This Version)
Approved by	Coordinator
Dissemination level	Public
Document type	Report
Total number of pages	39

Partners

- University of Twente – Centre for Monitoring and Coaching (CMC)
- Roessingh Research and Development (RRD)
- Danish Board of Technology Foundation (DBT)
- Sorbonne University (SU)
- University of Dundee (UDun)
- Universitat Politècnica de València, Grupa SABIEN (UPV)
- Innovation Sprint (iSPRINT)

Abstract

This deliverable reports on the first round of usability evaluations for three different functional prototypes of the Council of Coaches system. A focus group was held in which participants were able to experience these three different variations. During the focus group, participants were asked for their feedback on the prototypes and were additionally asked to perform two card-sorting tasks to get more insight into their preferences for coaches' characteristics and coaching topics. We conclude this deliverable with a discussion of the results and how they serve as input for the next design cycles in the Council of Coaches.

Corrections

v1.0.1 Correctly applied EU logo on header page.
Fixed footer.

v2.0.0 This version of the deliverable was updated as a response to the project's first periodic review in April of 2019. The following changes were made:

- Changed the title 'Demonstrator evaluation study' into 'Usability evaluation study'.
- Additional content added:
 - Table in the introduction that describes the setting, sample size and type of participants in the usability evaluation study
 - Abstract in section 4, the description of the usability evaluation study.
- Restructuring of the document:
 - Overview of functional demonstrators (in version 1.0.1 part of the methodology of the 1st evaluation study) is now a separate section, after the objectives.
 - The outline of the usability evaluation study; it now consists of the following subsections: abstract, introduction, method, results, discussion, future research
 - In v1.0.1 there was a separate section "expert review on the demonstrators". This review was done to finalise the functional demonstrators for the first evaluation study. Therefore, we integrated the results of this review in the descriptions of the functional demonstrators (Section 3).
- Removed "Future Research" section (originally section 6 in v1.0.1).
- Added section with Updated Requirements according to the FICS scheme, in line with other related deliverables D2.3, D2.5 and the upcoming D2.6.
- This document now uses the new, updated deliverable template format.

Table of Contents

1	Introduction.....	7
2	Objectives	8
3	Overview of functional demonstrators.....	9
3.1	The PowerPoint demonstrator	9
3.2	The 2D/Web demonstrator	10
3.3	The official Council of Coaches functional demonstrator.....	12
4	Usability evaluation study	14
4.1	Abstract.....	14
4.2	Introduction.....	14
4.3	Method	14
4.3.1	Participants	15
4.3.2	Study procedure	15
4.3.3	Data analysis	17
4.4	Results.....	17
4.4.1	Participants	17
4.4.2	Feedback on the PowerPoint demonstrator	17
4.4.3	Feedback on the 2D/Web demonstrator	18
4.4.4	Feedback on the official Council of Coaches functional demonstrator.....	19
4.4.5	Differences and similarities between the demonstrators	20
4.4.6	Results of card-sorting method	21
4.5	Discussion & Conclusion.....	23
5	Updated requirements.....	25
5.1	Functions and Events	25
5.2	Interaction and Usability	26
5.3	Content and Structure	27
5.4	Style and Aesthetics	29
6	Bibliography	31
A.	Focus group protocol	32
B.	Demonstrator evaluation table	34
C.	Card-sorting sheet	35
D.	Cards used in the card-sorting task – Characteristics	36
E.	Cards used in the card-sorting task – Coaching topics.....	37

List of figures

Figure 1: A screenshot of the PowerPoint Demonstrator.	9
Figure 2: A screenshot of a PowerPoint slide containing the 'select an answer' icon that was added to the PowerPoint demonstrator.....	10
Figure 3: A screenshot of the 2D/Web demonstrator	11
Figure 4: A screenshot of the updated 2D/Web demonstrator.....	12
Figure 5: A screenshot of the official Council of Coaches functional demonstrator.....	12
Figure 6: A screenshot of the updated interface for the 3D demonstrator.....	13
Figure 7: Results from the first card-sorting task for all ten participants per topic.	22
Figure 8: Results of the second card-sorting task for all ten participants per characteristic.....	23

List of tables

Table 1: Overview of studies in deliverable D2.4 (name, method, setting, N, participants).....	7
Table 2: Council of Coaches evaluation cycles.	8
Table 3: Focus group framework of (sub)-themes.	16
Table 4: List of coach characteristics and topics for card-sorting method.....	17
Table 5: Positive, neutral, and negative comments on the PowerPoint Demonstrator.....	18
Table 6: Positive, neutral, and negative comments on the 2D/Web Demonstrator.	19
Table 7: Positive, neutral, and negative comments on the PowerPoint Demonstrator.....	20

Symbols, abbreviations and acronyms

CMC	Centre for Monitoring and Coaching
COUCH	Council of Coaches
D	Deliverable
DBT	Danish Board of Technology Foundation
EC	European Commission
FICS	Functions, Interactions, Content, Style
ISPRINT	Innovation Sprint
M	Month
MoSCow	Must have, Should have, Could have, Won't have
MS	Milestone
RRD	Roessingh Research and Development
SU	Sorbonne University
UDun	University of Dundee
UPV	Universitat Politècnica de València
UT	University of Twente
WP	Work Package

1 Introduction

This deliverable describes the first usability evaluation of the first set of functional prototypes for Council of Coaches. We start by listing the objectives in Section 2 and describing how this evaluation round fits in the four evaluation cycles planned to be performed in the project.

We start with the description of three versions of the functional demonstrator that were used in this first usability evaluation round in section 3. In addition to an early prototype containing fully 3D, animated embodied conversational agents, two other functional demonstrators were designed that were deemed more suitable for obtaining feedback on the contents provided. Thus, allowing us to evaluate some aspects that would be difficult to evaluate using this 3D prototype. This was done because some specific aspects, such as user's preferences for human-like or non-humanlike characters, might be quite abstract to test without an example. The two additional demonstrators also allowed us to give the participants a hands-on experience, since the functional prototype for now requires a powerful computer to run, while the other demonstrators could run on tablets – allowing us to let the users experience interaction first-hand.

Section 4 describes the end-user evaluation of the demonstrators. A total of 10 older adults (55+ years) took part in this study. We asked them for feedback on the interaction with the technology, design of the interface, and the character design and appearance of the virtual coaches. The results indicated that the participants positively evaluated the basic concept of the Council of Coaches system. In the conclusion (Section 5) we describe how the results serve as input for further development of the Council of Coaches system by summarizing the results and providing a list of updated requirements.

An overview of the study described in this deliverable can be found in Table 1 below.

Table 1: Overview of studies in deliverable D2.4 (name, method, setting, N, participants).

Study	Method	Setting	N	Participants <54	Participants >55	Participants with health conditions (DM-II, CP)
Evaluation of first functional prototype	Focus groups	Lab-setting	10	-	10	-
Totals			10	-	10	-

2 Objectives

The objective of this deliverable is to describe the first round of end-user evaluations of the first Council of Coaches functional prototype. In Council of Coaches, there are four “official” cycles of demonstrator releases followed by evaluations, as depicted in Table 2 below.

In this document, the first evaluation study of the functional demonstrator is described. Besides the official Council of Coaches prototype, two additional demonstrators were developed to measure specific aspects such as virtual coach preferences, which are not yet available in the official prototype. This will aid the development process of the Council of Coaches prototype since it generates more insights in the needs, preferences, and barriers end-users have when interacting with virtual coaches about their health, which can serve as input for the future design cycles.

The aim of this first prototype evaluation study is to gain insight in the interaction that end-users have with the Council of Coaches application. These insights are gained by providing potential users three distinct functional demonstrators of the Council of Coaches in a focus group setting. The main focus of this evaluation study is on usability, character design and presentation, as well as interface design of the functional prototype.

Table 2: Council of Coaches evaluation cycles.

Council of Coaches Evaluation Cycles	
Cycle 1	
M9	Milestone 2: First Functional Prototype
M12	D2.4: Evaluation results of first functional prototype and updated requirements
Cycle 2	
M15	Milestone 3: Second Functional Prototype
M18	D2.5: Evaluation results of second functional prototype and updated requirements
Cycle 3	
M21	Milestone 4: Third Functional Prototype
M24	D2.6: Evaluation results of third functional prototype and updated requirements
Cycle 4	
M27	Milestone 5: Technical Prototype
M36	D7.7: Final Demonstration Results

	Completed
	To be done

3 Overview of functional demonstrators

The aim of the first functional demonstrator evaluation was to evaluate the acceptance of the Council of Coaches concept and approach. The three different demonstrators that were used in this evaluation study are: (1) a PowerPoint demonstrator, (2) a 2D/Web Demonstrator, and (3) the official Council of Coaches functional prototype. In the following subsections we describe these three demonstrators and how they represent various parts of the concept. We also describe how we prepared each demonstrator for evaluation.

3.1 The PowerPoint demonstrator

The PowerPoint Demonstrator is an interactive PowerPoint presentation that shows one main concept of the Council of Coaches, that is, the idea of multiple virtual coaches with personalities and backstories. The function of this demonstrator is to know how end-users evaluate these rich character designs, the interaction options (buttons), and the multiple-coach setting. A screenshot of this demonstrator can be found in Figure 1. In this dialogue, the users could get to know Francois, where he is from, how much he loves cheese, and what his role in the council of coaches is. The Francois introduction dialogue consists of 62 different “nodes”. In the conversation that this demonstrator contains the user can talk to the diet coach François. This demonstrator was translated to Dutch for the focus group.

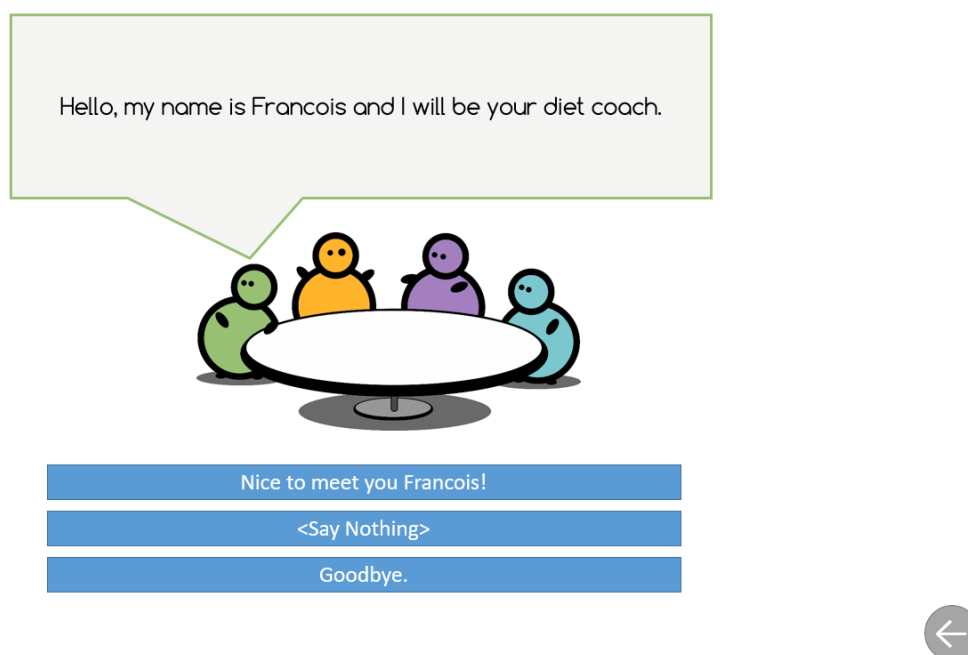


Figure 1: A screenshot of the PowerPoint Demonstrator.

Before presentation to the focus group, we performed a pre-evaluation of the PowerPoint demonstrator as a final check. We showed it to a number of HMI-researchers and asked them for their feedback and thoughts. The points for improvement provided for the PowerPoint demonstrator were:

- It was not directly obvious that the blue buttons at the bottom of the screen represented “things that you can say to the coach”.
- One expert expected that the coach would ask more about her – the dialogue focuses more around “getting to know François” – *“It’s a coach, right? So, I expect him to ask about what I like. Like when he asked if I’d like cheese.”*
- When only one “reply option” button is available, it is not directly obvious that you have to press this as a “Continue” option.
- It might be unclear which coach says what.

Based on these comments, we made the following adjustment to create the final version of this demonstrator:

- Instruction icons were added to the first few slides ('select an answer' and 'press to continue'). An example of such an icon can be found in Figure 2 below.

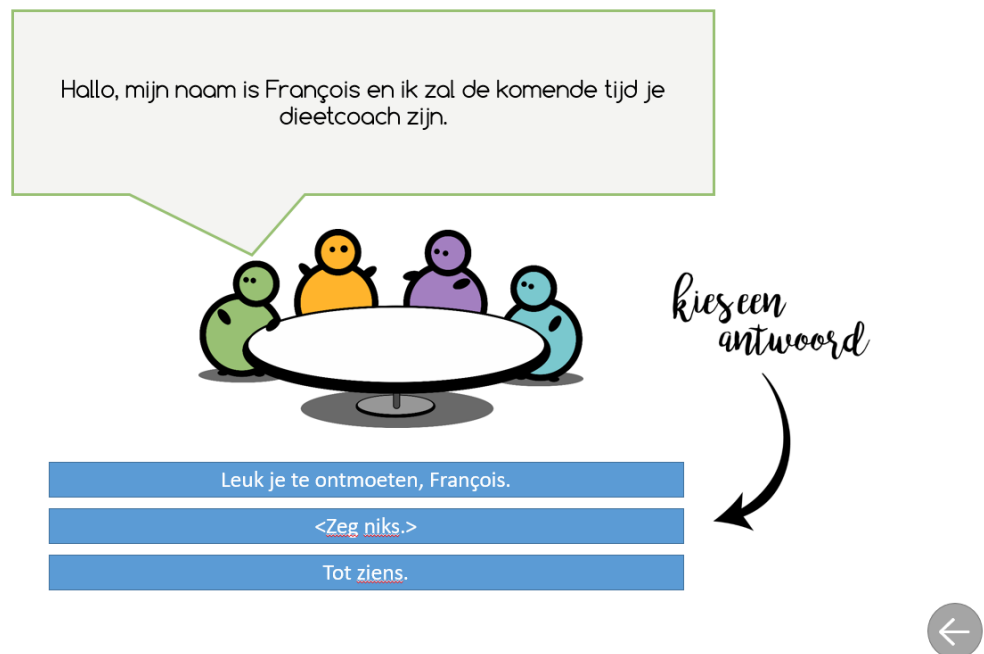


Figure 2: A screenshot of a PowerPoint slide containing the 'select an answer' icon that was added to the PowerPoint demonstrator.

3.2 The 2D/Web demonstrator

The 2D/Web Demonstrator is an interactive 2D demonstrator that runs as a website. A screenshot of this demonstrator can be found in Figure 3. This demonstrator contains the same dialogue with diet coach François, but the user now also has the possibility to also talk to physical activity coach Alexa. In this dialogue, the users are introduced to Alexa, her uneasy youth in Brazil, her short Olympic career, and here style as a physical activity coach (97 dialogue nodes). In addition, the other coaches (social coach and cognitive coach) are introduced. This demonstrator has a more dynamic interaction than the PowerPoint demonstrator, since other coaches can interject when François and Alexa are speaking. Furthermore, the coaches do not sit together behind one table but are spread around in the room with attributes that illustrate their coaching domain (e.g. kitchen for nutritional advice, bikes for activity coach). This demonstrator was also translated to Dutch for the focus group.

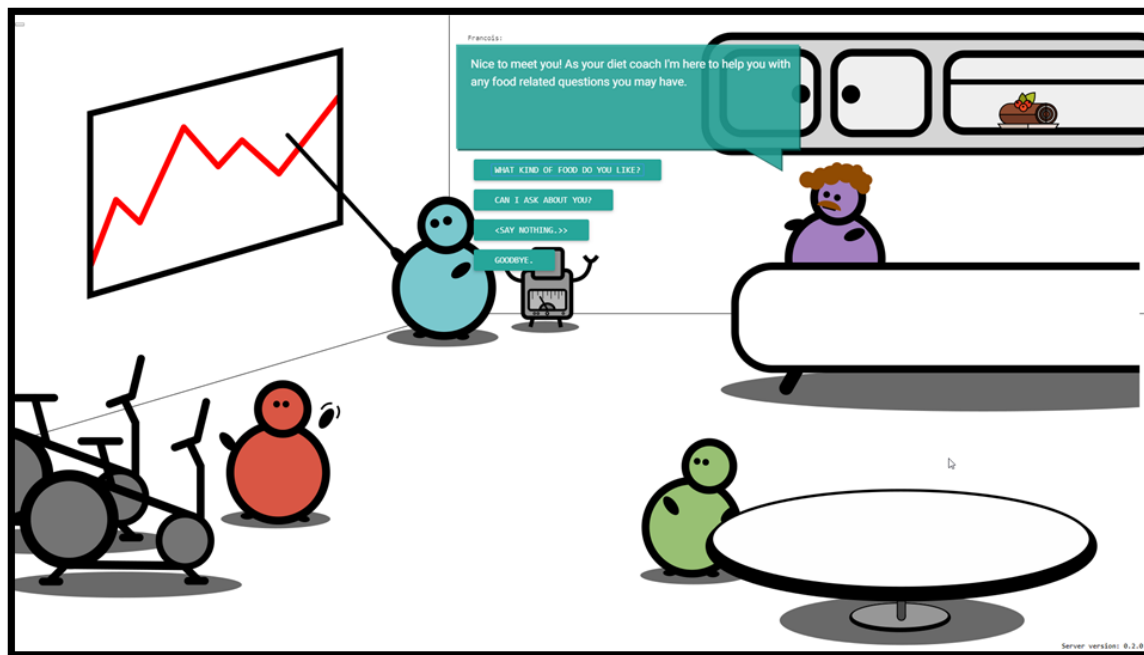


Figure 3: A screenshot of the 2D/Web demonstrator

Before presentation to the focus group, we also performed a pre-evaluation of the 2D/Web demonstrator as a final check. We again showed it to a number of HMI-researchers and asked them for their feedback and thoughts. The points for improvement provided for this demonstrator were:

- François has a different colour in this demonstrator than in the PowerPoint slides.
- Again, the 'continue' option might not be clear.
- A short introduction to how the demonstrator can be used might be useful, since it might not be clear that the user can click on a coach to start a dialogue.

Based on these comments, we made the following adjustments to create the final version of this demonstrator:

- François was changed to the same colour as in the PowerPoint demonstrator. A screenshot of the updated interface can be found in Figure 4.
- An introduction of the coaches and short instruction on how to select a coach for a conversation was added.

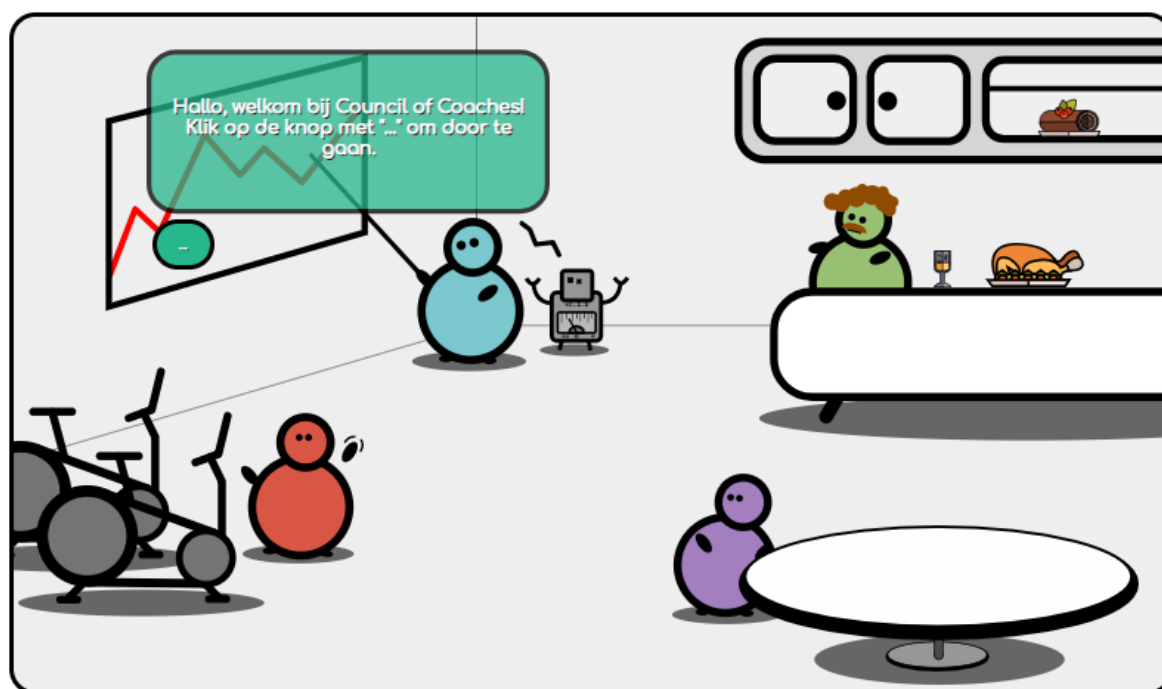


Figure 4: A screenshot of the updated 2D/Web demonstrator.

3.3 The official Council of Coaches functional demonstrator

The Council of Coaches functional prototype is a demonstrator that runs on Unity. A screenshot of this demonstrator can be found below in Figure 5. The dialogues in this demonstrator are in English, since it currently only supports text-to-speech for the English language. This feature and the fact that it needs a powerful computer to run are the main reasons for showing this demonstrator plenary.



Figure 5: A screenshot of the official Council of Coaches functional demonstrator.

- The coaches' speech might not be optimal for understanding since there were strange pauses due to the text-to-speech software.
- The answer buttons might be quite small.

- The background is a green 'box' with red logo, which is an unnatural environment for humans.
- The coaches do not have much movement/expression yet, so they can be perceived as staring or unfriendly.
- The demonstrator is in English, which is not optimal for the Dutch target audience.

Based on these comments, we made the following adjustment to create the final version of this demonstrator:

- The background was changed to a café-like setting (see also Figure 6).
- The input for the text-to-speech was adjusted so that the speech/pauses were more natural. Punctuation and spaces were adjusted and some words were written phonetically to improve overall text to speech output.



Figure 6: A screenshot of the updated interface for the 3D demonstrator.

4 Usability evaluation study

4.1 Abstract

Background: We evaluated the first functional demonstrator in order to get input on the acceptance of the Council of Coaches concept by potential end-users, and in addition to receive insights into their preferences for coaching and the representation of the coaches.

Methods: We held a focus group with 10 older adults. In this focus group the participants got to try out a PowerPoint and 2D/Web demonstrator in addition to the official 3D-demonstrator. The participants also completed two card-sorting tasks in on the topics of coaching content and coach characteristics.

Results: Participants seem to accept the Council of Coaches concept. They showed a slight preference for human-like coaches. They seemed to like the coaches' backstories and the interaction between the coaches. Some points of improvement were identified, such as a better representation of which coach was talking, and a better balance between coaching and social conversation. Topics for coaches to discuss that were appreciated in the card-sorting task were physical activity, nutrition and social advice. Examples of characteristics that were appreciated were expertise and humor. Participants also responded that a coach being 'a friend' would be too much, but 'friendly' would be appreciated.

Discussion & conclusion: We learned that the interaction with multiple coaches and the dynamic parts of the conversations were appreciated. The balance between social conversation and coaching content needs to be improved, since participants indicated that there was too much small talk. However, since participants also seemed to be engaged in the background stories and the content consisted mostly of introductions to the coaches, it might also be the case that this is a case of too little coaching content. Participants liked the characters, but they should be easy to distinguish from each other, and their coaching domain should be clear. Overall, the Council of Coaches concept seems to be accepted by participants and development should focus on further definition of the content to be discussed and the usability of the interface.

4.2 Introduction

The aim of this study is to collect opinions and thoughts of potential end-users (older adults of 55 years or older) on the basic concept of the Council of Coaches system with an emphasis on acceptance and preferences. That is, we wanted to know how users evaluate the system interaction, the User Interface (UI) design, the content of the coaching, and the appearance and character design of the coaches. These topics were selected since they could provide us with feedback that serves both the design process of the interaction and the coaches, and helps us direct the focus on in the process of creating content, that is, specifying the initial coaching topics and coaches' personalities. The results of this evaluation will serve as input for the following design cycles and other work packages. Input on coaching content and strategies will benefit the process in WP3, preferences on topics can also provide input as to what needs to be sensed in WP4, feedback on the type and flow of dialogues that are preferred helps WP5, and WP6 benefits from the participants views and suggestions on the interface and characters.

4.3 Method

We performed a focus group in which participants interacted with three versions of the demonstrator. The use of three versions of the demonstrator (as introduced in section 3) allowed us to present the various elements that make up the Council of Coaches concept in different forms. It also allowed us to shift the emphasis between different parts of the concept, which improved the specificity of feedback.

The PowerPoint version is a clean interface with minimal distraction from the content of the coaching and emphasizes the concept of backstories and personalities. It also introduces the interaction with the coaches following a speech-bubble and reply-button paradigm. The 2D/Web version focuses on a more dynamic interface design (while following the same basic interaction paradigm) with more interaction between the coaches. It also features more content for coaching and a scene that connects the coaches to their domains. Last, the official demonstrator emphasizes the appearance of the coaches in a more realistic human-like 3D form with non-verbal behaviours and text to speech. The first functional demonstrator for Council of Coaches is technically complex, and has relatively high hardware

requirements to run smoothly. Therefore, this prototype was shown plenary to participants to avoid biases regarding the prototype that are caused by the use of lower-end devices.

Once participants had interacted with the demonstrators, we asked them to perform a card sorting task that provided more insight into the acceptance and preferences for various topics and coach characteristics, which benefits the further design and development process.

4.3.1 Participants

As the focus of this evaluation is on interaction, user experience, and character design aspects of a functional prototype, end-user inclusion criteria should match those that are described in D2.2 (Broekhuis, van Velsen, op den Akker, Øjvind Nielsen, & Andersson, 2018). The initial target groups for the final Council of Coaches are older adults who are 55 years or older, and have one or more age related impairments (ARI), chronic pain (CP) or diabetes mellitus type 2 (DM2). Other than that, the participants should be able to understand Dutch; no additional inclusion criteria are present, since the content of the first functional demonstrator is not tailored to individuals or specific conditions.

In order to recruit participants, an invitation was sent to the members of the 'RRD Onderzoek Panel' (RRD Research Panel) that fit the criteria. The RRD Research Panel is a collection of individuals that have voluntarily signed up to be contacted for research studies. A total of ten respondents agreed to participate in this study. Additional information on the recruitment of participants can be found in Deliverable 9.1 (van Loon, et al., 2018), which also discusses preventative measures to protect potentially vulnerable target groups during evaluation processes.

4.3.2 Study procedure

The focus group had the following general outline (see appendix A for the complete procedure):

- Short introduction on the project and the goals of the Council of Coaches application.
- Each participant received a tablet to try out two different demonstrators. After each demonstrator, they filled out the form as included in appendix B and a plenary discussion was held.
- A third version of the demonstrator was shown plenary (the "official" Council of Coaches functional demonstrator). Participants were again asked to fill in the form and the demonstrator was discussed plenary.
- Once participants had interacted with the demonstrators, they were asked to perform two card-sorting tasks on the characteristics of a coach and coaching topics.

To summarize, the data that was collected were the following:

- Audio recordings of the full session (using two recorders). These recordings were transcribed.
- Feedback-forms for all of the participants for each of the demonstrators.
- Card-sorting sheets for the topic and characteristics tasks (for each participant).

4.3.2.1 Discussion topics

The plenary discussions during the focus group were semi-structured, meaning that while the discussions had an open nature to allow for new ideas and opinions brought up by the participants, there were a clear set of themes to be explored during the focus group.

The main themes were: (1) interaction with the technology, (2) opinions on character design, (3) appearance of the virtual coaches, and (4) the content of the coaching.

Table 3 shows the framework of the themes and subthemes for the discussions in the focus group. These (sub)-themes were not explicitly asked in the discussions. Instead, follow-up questions were formulated for each theme when the discussions slowed down.

Table 3: Focus group framework of (sub)-themes.

Theme	Subtheme
Interaction	<ul style="list-style-type: none"> Do the participants 'get' that there are different agents talking? Do the participants 'get' that they can select an answering option? Are subtitles needed? Do the buttons have a reasonable size? Is the process of selecting answers through buttons workable?
Concept	<ul style="list-style-type: none"> When do the participants expect to talk to the coaches? (At which moments could there be a need for support by coaches?) What do the participants think about having multiple agents to talk to? What do the participants think about representing different health domains through different coaches?
Content/characters	<ul style="list-style-type: none"> What do participants think about the backstories of the coaches? What do participants think about small talk with the coaches? Do participants think that characters are allowed to make jokes/user humour?
Looks	<ul style="list-style-type: none"> Do participants prefer 2D or 3D agents? Should agents have a neutral look or can their look reflect their personality? Do participants prefer humanlike or non-humanlike characters?

4.3.2.2 Card-sorting

After having seen all three demonstrators, participants performed a card-sorting task to gain insights into the properties that the participants would like the coaches to have and the topics for the coaching that the participants would like to have the coaches provide. The participants were asked to map the cards on a sheet with a scale of 'I think this is important' to 'I think this is less important' (this sheet can be found in appendix C). There were two rounds of card-sorting. In the first round, participants mapped variables for coach characteristics, such as expertise, friendliness, and gender. These characteristics were a combination of roles that a coach could take on (expert, experience expert, companion / friend), characteristics that would influence a coach's looks (gender, age, appearance / looks), and characteristics that influence their behaviour otherwise (backstory, has humour, personality / mannerisms, and friendliness). In the second round, the cards contained various health topics for coaching, such as recipe ideas, brain training, and medical reminders. The topics were selected based on the coaching domains. To give some examples: recipe ideas and food diary would be possible topics for a diet coach, brain training would be a topic for a cognitive coach, and talking about emotions or stress would be part of a mental health coach.

The original Dutch card-sorting sheets and cards can be viewed in appendices D and E. In Table 4 below, the English translations for both card-sorting tasks are shown.

Table 4: List of coach characteristics and topics for card-sorting method.

Coach characteristics	Coaching topics
Expertise	Recipe ideas
Experience expert	Tips for physical activity
Friendliness	Tips for social activity
Personality/mannerisms	Brain training
Gender	Reminders for appointments
Appearance/looks	Medication reminders
Age	Ideas for things to do/places to go
Backstory	Food diary
Companion/friend	Smalltalk
Has humour	Talking about hobbies
	Talking about emotions or stress
	Talking about the weather

4.3.3 Data analysis

The audio recordings that were made during the focus group were transcribed. For each demonstrator, the comments and responses of the participants were extracted from the transcript and combined with the comments that the participants wrote on the demonstrator evaluation tables. For the card-sorting tasks the responses for each participant for each card were coded. Placement of a card on the 'important' side of the scale was coded as (1) important, placement of a card on the 'less important' side of the scale was coded as (3) less important, and cards that were placed in between were coded as (2) neutral. Following the coding, the number of responses per card for 1, 2, and 3 were counted.

4.4 Results

The focus group resulted in several types of feedback and information for improvement and further development of the Council of Coaches system. We will discuss these results below.

4.4.1 Participants

The participants of the focus group were 10 older adults aged between 58 years old and 80 years old ($M = 71$ years, $SD = 5.8$). Of the 10 participants, seven (70%) were women and three (30%) were men. They had various backgrounds (e.g. technical and health care) and most participants stated to have active lifestyles and took part in different activities such as volunteer work.

4.4.2 Feedback on the PowerPoint demonstrator

The first demonstrator discussed during the focus group was the PowerPoint demonstrator. The participants got the opportunity to use this demonstrator on a tablet and were asked to fill in tables on what they thought was positive, neutral and negative about the application.

In general, respondents especially liked the characters. They were considered nice, amiable and humorous. As one said: *"It is a happy character and he has humour!"*. Respondents enjoyed interacting with the coaches. They liked the manner in which the coaches posed questions and that the topic was not just about food, but had a broader character. However, they did think the conversation was sometimes lacking direction and that they did not receive much useful information from the coaches. The main consensus seemed to be that they wanted the conversation to be more to the point, and that they felt that, while entertaining, it went too much off topic, which made it unpredictable what would be next. This made it for some participants unclear, because they did not understand the underlying goals

of the conversation. Further, the multiple-agent setting was at times a bit confusing for the respondents and they would like a return-option to go back to previous questions. As one respondent said: *"Once you have selected an answer you cannot go back. (I started with fish and he keeps talking about fish)"*. Some respondents also indicated that they would prefer something they would not need a computer for, as one mentioned: *"This is quite fun, but I am not in front of the computer that often. I solve this by talking a lot with friends, so I cannot do much with this."* Table 5 below provides an overview of all the positive, neutral, and negative comments on the PowerPoint demonstrator.

Table 5: Positive, neutral, and negative comments on the PowerPoint Demonstrator.

Positive	Neutral	Negative
<ul style="list-style-type: none"> Formulation of questions Enjoyable interaction Broad themes (e.g. not just about food) Multiple answer options Humour of coaches Amiable characters 	<ul style="list-style-type: none"> Lack of direction in conversation Not receiving useful information Need for more facial expressions of coaches Runs on computer 	<ul style="list-style-type: none"> Not understanding/unclear of system Confusing with multiple coaches talking No option to return to previous options Childish appearance

4.4.3 Feedback on the 2D/Web demonstrator

The second demonstrator shown to the participants was the 2D/Web demonstrator. As with the previous demonstrator, the participants got the opportunity to use this demonstrator on a tablet and were asked to fill in tables their thoughts (positive, neutral, negative) on the application.

Participants were positive about this prototype because there was more interaction with and between the characters. One participant stated that *'it was like a game'* and that she liked it, even though she had disliked the previous demonstrator somewhat. Further, they were positive about the context in which the coaches were placed. This created better distinctions between the coaches (e.g. physical activity coach next to the home trainers). Respondents evaluated Alexa as having a nice character and they liked that she starts by suggesting small steps, such as talking a walk through the park, to heighten physical activity (*"Do you like to walk! Oh, good. Taking a walk in the park is excellent' – That, for me, is a decent advice in terms of content."*).

Participants were less positive about the background stories of the coaches that are interwoven in the conversations between user and coach. They feel these backstories are quite elaborate and distract them from the goals of the conversation. One participant stated that the realization of *"it is just a computer"* made the backstories less relevant (*"it is a computer and it is not born in Brazil"*). Suggestions were made to let the user choose between 'hard' vs. 'soft' conversations: 'hard' conversations that are direct and to-the-point and 'soft' conversations with more background information. It was also mentioned again that having a coach who looks away is less believable as a coach.

The participants would again have liked an 'undo' button. They would also have liked more answering options. They also advised us to be careful with adding opinions on topics to the content (e.g., fish is better than meat). In addition, the use of difficult words was also noted and they advised us to keep the language from being too difficult. For example, the term 'cognition' could lead to confusion, as one respondent commented on: *"Do people know what a cognition coach is? Perhaps memory trainer? That means something different, but normal people do not make that difference."* Other advises that were given were to make it clearer who was talking (e.g. make the colours of the speech bubbles match the coach or zoom in on the coach that is speaking). Positive, Neutral and Negative comments on the 2D/Web demonstrator are summarized in Table 6 below.

Table 6: Positive, neutral, and negative comments on the 2D/Web Demonstrator.

Positive	Neutral	Negative
<ul style="list-style-type: none"> ▪ Engaging conversation ▪ Context ▪ Mutual interaction between the coaches ▪ Gaming element ▪ Multiple step goal setting: starting with small goals ▪ Gaining some useful information ▪ Pleasant character (Alexa) ▪ 'Fit' with end-user ▪ Clear distinctions between the coaches (e.g. colour and attributes) 	<ul style="list-style-type: none"> ▪ Curiosity to interaction for conversations with other coaches ▪ Not always interested in background ▪ Preference to adapt conversions: 'hard' vs. 'soft' conversations ▪ Coaches make judgements (e.g. fish is healthier than meat). ▪ No option to add recipes or tips of your own ▪ Preference for human coaches ▪ Coaches should not look away whilst talking ▪ Connection between social coach and 'scientist' like attributes not clear 	<ul style="list-style-type: none"> ▪ Lack of relevance in conversations ▪ Connection between speech bubble and coach not always clear ▪ Definition of cognition coach is unclear ▪ Language level is too high ▪ No option to return to previous question ▪ Underlying assumption, that end-user can be persuaded to healthier choices. ▪ Coaches have little emotions ▪ No option to talk to one or two coaches

4.4.4 Feedback on the official Council of Coaches functional demonstrator

Lastly, the Council of Coaches functional prototype was plenary shown to the participants. They were again asked to fill in tables on what they thought was positive, neutral, and/or negative about the application.

Respondents liked the human-like appearance of the coaches and the personal touches in their stories. A reoccurring issue with this however is that the functional goal of system (assisting people with coaching) feels less of a priority due to the rich and elaborate background stories. Respondents indicated that there is room for improvement, such as showing exercise examples. When the physical activity coach gives the user physical exercises, that she shows how these should be performed or gives the end user a video of the exercise. As one participant stated: *"That is very important for physical activity, that you have an example."* Another participant stated that for the diet coach having links to recipes might be useful: *"For the diet coach you can have a link to certain recipes. [...] You just have to search for what is healthy and what isn't."* In addition, besides the spoken language, users would prefer if the text of what the coaches say can be shown on the screen, e.g. for people with hearing impairments (*"Sound with text. Because there are people who have issues with their ears, and who cannot turn the sound of the computer to a volume that allows them to follow the conversation."*). A negative comment was made on the speech-face synchronization (speech is not synchronized with facial expressions). In addition, the facial expressions of the diet coach could have used a little more contrast to improve their recognition. Again, the positive, neutral and negative outcomes of the demonstrator are summarized below in Table 7.

Table 7: Positive, neutral, and negative comments on the PowerPoint Demonstrator.

Positive	Neutral	Negative
<ul style="list-style-type: none"> Human-like appearance Personal background of the coaches 	<ul style="list-style-type: none"> Need to inspire end-users to work on their health Lack of direction in conversation Preference to adapt conversions: 'hard' vs. 'soft' conversations Need for showing exercises Show the text of what the coaches say on the screen. 	<ul style="list-style-type: none"> Speech-face synchronization Contrast of facial features

4.4.5 Differences and similarities between the demonstrators

In this section, we will discuss some comments that were made across demonstrators and differences between demonstrators for each of our four themes.

Interaction with the technology

The interaction through buttons with the demonstrators worked quite well. Participants were confused by the '...'-continue option when they encountered it at first, but after some assistance they got more acquainted with it. The wish for an undo button was one that continued for the first two demonstrators. The wish for an undo button was not expressed for the third demonstrator, but participants indicated that the addition of text (e.g. subtitles) would be a good addition to the coaches' speech output.

In terms of interaction experience of the conversation, participants seemed to prefer the demonstrators in which more coaches joined the conversation. This might be due to the 'it is like a game' effect that one of the participants stated she experienced for the second demonstrator. This effect together with the suggestions that there should be some sort of reward system, for example, through points or a logbook might be good directions for creating an engaging system.

Opinions on character design

While participants did think that there was too much small talk going on, they liked the characters. Even though some participants stated that they did not care for the background story (*"it is just a computer"*), participants were found to have remembered what the coaches told them. When discussing the difficulty of the language used by the coaches, and more specifically the types of cheeses François was listing, one participant stated *"It makes me think: all those cheeses... they have these names; where do they come from?"*. The other participants immediately stated *"From François!"* and started to refer to the coach's French background as an argument that it made sense for the coach to talk about French cheese. This indicated that they did identify with the character, and that there might be advantages to having backstories, but that they might simply be overrepresented in the dialogues for these three demonstrators.

Appearance of the virtual coaches

Participants were first introduced to the two demonstrators that had the quite neutral 2D "blob men" (the round and friendly mascots of the Council of Coaches project). While they liked them and were pleased with the fact that they were quite neutral, they did prefer the 3D human-like characters in the third demonstrator more. Our hypothesis is that this is mostly because it is easier to recognize the character that is speaking due to the increased expressional capabilities of the 3D characters. Another effect that might be in play is that the 3D characters are also more comparable to actual human coaches.

Content of the coaching

Participants stated that they would like their coaches not to drift too far or too long from the supposed coaching topic, which was an important insight. Although they did indicate that the small talk and getting to know conversations were fun, they also stated that in these cases there was too much of it in comparison to the serious coaching conversations. Participants suggested that the small talk conversations could be part of an entertainment coach, whilst the other coaches could then focus on their respective topics. A participant stated that she thought that it might be a way to get people to get to know the coaches, so that they might follow the given advice faster later on. The personal stories and small talk topics stayed topics on which opinions varied. Participants did like the diet and physical activity topics and stated that they thought these were very important. The results of the card-sorting exercise on topics that will be discussed as part of the following section reflects these statements.

4.4.6 Results of card-sorting method

As described in Section 4.3.2.2, after the hands-on evaluations with the demonstrators, the participants were asked to perform two card-sorting tasks in which they could indicate their preferences for topics that the coaches might address and for properties of the coaches.

4.4.6.1 Preferences for coaching topics

The first card-sorting task that was performed during the focus group was on coaching topics. Participants were asked to place cards on a scale of 'important' to 'less important'. Cards that have been placed in the middle of this scale are interpreted as neutral. An overview of the results can be found in Figure 7.

The topics that were found most important were 'recipe ideas', 'tips for physical activity', and 'brain training'. Topics that were considered less important were 'small talk', 'talking about the weather' and 'talking about hobbies'.

On the paper on which the cards were mapped, some participants wrote down additions. These were the following:

- For 'tips for physical activity', a participant wrote down that the coach needs to know and take into account any conditions the user might have.
- For 'ideas for things to do/places to go', a participant wrote down that these could be part of a reward system.
- For 'medication reminders', a participant wrote down that that should be the case when the user indicates that they want that.
- For 'brain training', a participant wrote down that these should be short exercises.
- For 'reminders for appointments', a participant wrote down that they should not be patronizing.
- For 'talking about the weather', a participant wrote down that a user cannot change the weather.
- For 'small talk', a participant wrote down that there is no time to do so.
- For 'talking about emotions or stress', was dependent on the type of coach.

A participant also wrote down two additional topics that were perceived as important. These were 'interaction' and 'reward'. While not directly topics that can be the subject of coaching, these are two themes that can be taken into account for the design of the system.

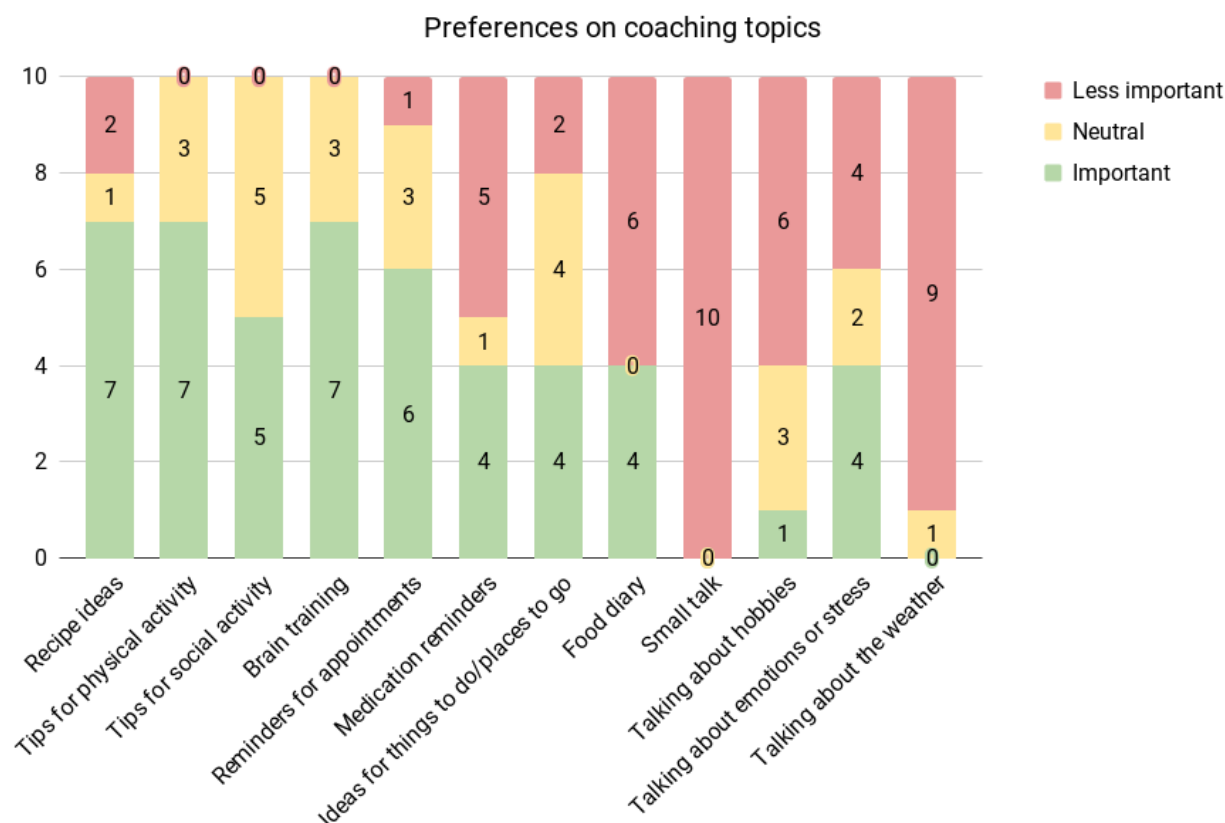


Figure 7: Results from the first card-sorting task for all ten participants per topic.

4.4.6.2 Preferences for coach characteristics

The second card-sorting task that was performed during the focus group was on coach characteristics. Participants were asked to place cards on a scale of 'important' to 'less important'. Cards that have been placed in the middle of this scale are interpreted as neutral. An overview of the results can be found in Figure 8.

The characteristics that were found most important were 'expertise', 'experience expert', and 'friendliness'. Characteristics that were found less important less were 'gender', 'age', 'companion/friend', 'personality/mannerisms'.

On the paper on which the cards were mapped, some participants wrote down additions. These were the following:

- For 'has humour', one participant wrote down that this meant that a coach can laugh about himself or herself.
- For 'experience expert', a participant stated that they also needed to have had a good education.
- For 'gender', a participant stated that a coach should be 'without', while another participant wrote down 'depends'.
- For 'companion/friend', a participant wrote down that that would be too personal.

Participants seemed to agree after completing the task that a coach needed to be professional and friendly, but not too familiar.

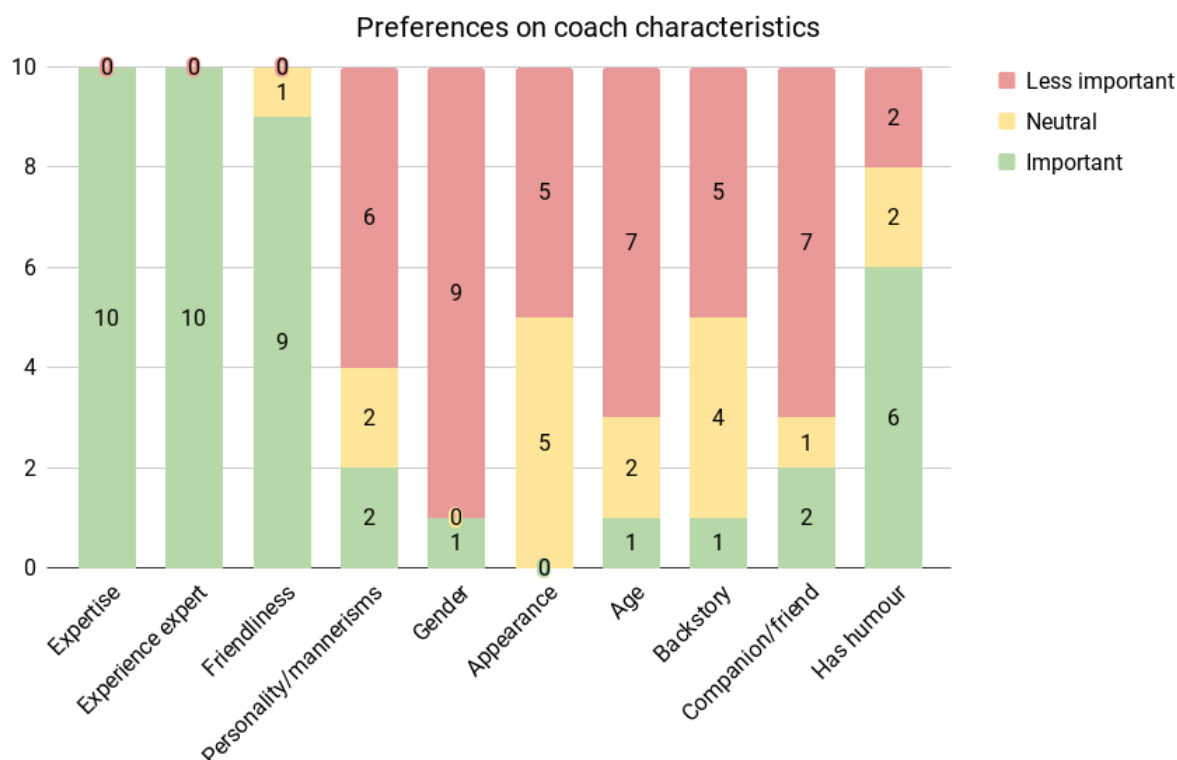


Figure 8: Results of the second card-sorting task for all ten participants per characteristic.

4.5 Discussion & Conclusion

The evaluation of the first functional prototype and the two additional functional demonstrators provided us with new insights and confirmed some assumptions that we already had. We will shortly discuss the main insights below and indicate how we can apply them for the improvement of the system. We will do so by means of the four themes we defined. While doing so we will also list the technical work packages that are the most relevant for the case.

For our first theme ‘interaction with the technology’, there were insights that are helpful for the interface design (WP6). We learned that buttons should be clear (i.e. the ‘...’-continue button was confusing), that an undo button for the answer options was preferred, and that subtitles could make the application more accessible. On the structuring of the conversations, we learned that participants seemed to prefer the interaction with multiple coaches. This is an important insight for the design of the dialogue model (WP5), coaching strategies (WP3), and turn-taking model (WP6).

For our second theme ‘opinions on character design’, we learned that, while participants did think that there was too much small talk going on, they liked the characters. While participants also would like the coaches to come to the point, most of them still seemed to identify the characters as real characters (*‘He is talking about French cheese because he is French off course!’*). The main insight therefore is that the background stories are worth designing (WP6), but that their presence in the conversation should be carefully balanced (WP3, WP5).

For our third theme ‘appearance of the virtual coaches’ the participants indicated that they liked the neutral looks of the blob-men, but they preferred the more human-like 3D coaches. Our hypothesis is that this is mostly because it is easier to recognize the character that is speaking due to the increased expressional capabilities of the 3D characters. Another effect that might be in play is that the 3D characters are also more comparable to actual human coaches. These insights can be used in the character design (WP6).

Finally, for our fourth theme ‘content of the coaching’, we learned that participants preferred to have multiple options to select the direction in which the conversation would continue. We also learned that the conversation should have a language level that is not too high. These insights, and the input that

participants would like to be provided with examples (e.g. for physical activity exercises and recipes) are a good addition to the design of the coaching content (WP3). Another insight that was important in terms of coaching content was the suggested need for reward and feedback functionality, which are confirmations that the behaviour analysis performed in WP4 in combinations with feedback designed by other WPs (e.g. WP3 and WP6) is an important part of the Council of Coaches system.

In addition to valuable insights from the demonstrator discussions, the two card-sorting tasks provided us with some more detailed information on participants' preferences for coach characteristics and coaching topics. For the task on coaching topics, important conclusions were that participants thought some functionality could be useful later on, but that they would like to see it as an option they could 'switch on' (reminders for medication and appointments). The participants also indicated that they thought the diet and physical activity topics were important, as we expected, but they also appreciated the brain-training topic, which indicates that they might necessarily see the coaches limited to roles that can be mapped to human equivalents (e.g. dietician, personal trainer, etc.). These insights are valuable in designing the tailoring options for the overall system and (the content of) the coaches (WP3).

The results from the characteristics card-sorting task confirmed our assumption that coaches should be relatable experts or experience experts. The preferences that were indicated for 'has humour' and 'friendliness' were important insights that can be used in the design of the coaches' strategies, dialogues, looks, etc. (WP3, WP5, WP6).

5 Updated requirements

Below we summarize the set of updated requirements that stem from the usability evaluation performed. The requirements are categorized according to the FICS scheme, proposed by (Benyon & Macaulay, 2002):

- Functions and Events (see §5.1)
- Interaction and Usability (see §5.2)
- Content and Structure (see §5.3)
- Style and Aesthetics (see §5.4)

For each requirement we specify:

- **ID:** A unique identifier for future reference.
- **Requirement:** Description and specification of the requirement.
- **Source:** Where the requirement originates from (i.e., this document and either the “PowerPoint Demo”, “Web Demo”, or “Tech Demo”).
- **Rationale:** Optional further explanation of the requirement.
- **Priority:** According to MoSCoW method: Must have, Should have, Could have, Won’t have
- **History:** Space for annotations in case of changes to this requirement.

5.1 Functions and Events

The following six requirements relate to “Functions and Events”.

ID:	D2.4-F1
Requirement:	There should be an option to add your own recipes or “tips”.
Source:	Usability evaluation of the 1 st functional demonstrator (Web Demo)
Rationale:	This was a suggestion by a participant in the focus group.
Priority:	Could
History:	

ID:	D2.4-F2
Requirement:	The definition of the “cognition” coach should be made clear.
Source:	Usability evaluation of the 1 st functional demonstrator (Web Demo)
Rationale:	The user thought that “cognition” was too difficult a term, and suggested to simplify it to e.g. “memory trainer”, also stating that this was not technically the same.
Priority:	Should
History:	

ID:	D2.4-F3
Requirement:	There should be an option to talk to one or two coaches directly.
Source:	Usability evaluation of the 1 st functional demonstrator (Web Demo)
Rationale:	The demonstrators did not yet allow to select coaches to initiate conversations.
Priority:	Should

History:	
-----------------	--

ID:	D2.4-F4
Requirement:	There could be an option of showing exercise examples, for example through a video.
Source:	Usability evaluation of the 1 st functional demonstrator (Tech Demo)
Rationale:	This was a suggestion by a participant in the focus group.
Priority:	Could
History:	

ID:	D2.4-F5
Requirement:	The diet coach could have links to recipes that are healthy for the user.
Source:	Usability evaluation of the 1 st functional demonstrator (Tech Demo)
Rationale:	This was a suggestion by a participant in the focus group.
Priority:	Should
History:	

ID:	D2.4-F6
Requirement:	System could include a “points” or reward system to increase engagement.
Source:	Usability evaluation of the 1 st functional demonstrator
Rationale:	This was a suggestion by a participant in the focus group.
Priority:	Could / Won't
History:	

5.2 Interaction and Usability

The following three requirements relate to “Interaction and Usability”.

ID:	D2.4-I1
Requirement:	The system should have a way of “returning to the previous option”. After a reply is given, it should be possible to revert this and go back to the previous step.
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo, Web Demo)
Rationale:	Sometimes participants lost track of the conversation and forgot what they actually “said” to the system. Based on this the suggestion to “undo” the move. [Possible it should suffice to have the chosen reply option visible on screen?]
Priority:	Could
History:	

ID:	D2.4-I2
Requirement:	It should always be possible to “escape” from conversational loops (e.g. “I started with fish and he keeps talking about fish”).
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo)
Rationale:	The conversation with Francois contained a loop that would have Francois repeat the same sentences about fish over and over again (dialogue design error).
Priority:	Should
History:	

ID:	D2.4-I3
Requirement:	Spoken text should be clearly “subtitled”.
Source:	Usability evaluation of the 1 st functional demonstrator (Tech Demo)
Rationale:	In the technical demonstrator, the spoken audio was sometimes difficult to understand, and it was noted that people with hearing problems would have particular difficulties.
Priority:	Should
History:	

5.3 Content and Structure

The following seven requirements relate to “Content and Structure”.

ID:	D2.4-C1
Requirement:	Content of the conversations should have a clear direction and should contain useful information.
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo)
Rationale:	Many participants noted that the conversational content went too much off-track and focused too much on “chit chat”.
Priority:	Should
History:	

ID:	D2.4-C2
Requirement:	Conversations should be “to the point”, and shouldn’t go too much “off topic”.
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo)
Rationale:	Many participants noted that the conversational content went too much off-track and focused too much on “chit chat”.
Priority:	Should
History:	

ID:	D2.4-C3
Requirement:	Conversations should be “predictable” in terms of their content.
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo)
Rationale:	Many participants noted that the conversational content went too much off-track and focused too much on “chit chat”.
Priority:	Should
History:	

ID:	D2.4-C4
Requirement:	The underlying goal of the conversations should always be clear.
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo)
Rationale:	Many participants noted that the conversational content went too much off-track and focused too much on “chit chat”.
Priority:	Should
History:	

ID:	D2.4-C5
Requirement:	Conversational content could be optional between “hard” and “soft”, where “soft” includes more background and social talk, as opposed to “hard” conversations that are strictly on topic and to the point.
Source:	Usability evaluation of the 1 st functional demonstrator (Web Demo)
Rationale:	Many participants noted that the conversational content went too much off-track and focused too much on “chit chat”.
Priority:	Could
History:	

ID:	D2.4-C6
Requirement:	There should be “more” reply options in the dialogues.
Source:	Usability evaluation of the 1 st functional demonstrator (Web Demo)
Rationale:	Participants sometimes thought that what they wanted to reply was not an option given to them.
Priority:	Should
History:	

ID:	D2.4-C7
Requirement:	Language use should be simple in all conversations.
Source:	Usability evaluation of the 1 st functional demonstrator (Web Demo)
Rationale:	
Priority:	Must
History:	

5.4 Style and Aesthetics

The following seven requirements relate to “Style and Aesthetics”.

ID:	D2.4-S1
Requirement:	The coaches should be able to express different facial expressions or emotions.
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo, Web Demo)
Rationale:	
Priority:	Could
History:	

ID:	D2.4-S2
Requirement:	The overall appearance of the system should be less “childish”.
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo)
Rationale:	The participants preferred the more “human” style of the coaches in the technical demonstrator.
Priority:	Should
History:	

ID:	D2.4-S3
Requirement:	The coaches should be facing the user when talking (i.e. don’t look away within the scene).
Source:	Usability evaluation of the 1 st functional demonstrator (Web Demo)
Rationale:	In the Web Demo, one of the characters was clearly facing away from the user.
Priority:	Should
History:	

ID:	D2.4-S4
Requirement:	When multiple coaches are involved in the conversation it should always be clear who is talking.
Source:	Usability evaluation of the 1 st functional demonstrator (PowerPoint Demo, Web Demo)
Rationale:	Especially when the conversation switched between different characters, the PowerPoint and Web Demo's often did not make it very clear who was uttering the current sentence.
Priority:	Must
History:	

ID:	D2.4-S5
Requirement:	The connection between the speech bubble and the coach should be clear.
Source:	Usability evaluation of the 1 st functional demonstrator (Web Demo)
Rationale:	Especially when the conversation switched between different characters, the PowerPoint and Web Demo's often did not make it very clear who was uttering the current sentence.
Priority:	Must
History:	

ID:	D2.4-S6
Requirement:	Speech should be properly synchronized with the coaches' facial expressions.
Source:	Usability evaluation of the 1 st functional demonstrator (Tech Demo)
Rationale:	Lack of synchronisation seemed unnatural.
Priority:	Should
History:	

ID:	D2.4-S7
Requirement:	Facial expressions should have sufficient "contrast" in order to be recognizable.
Source:	Usability evaluation of the 1 st functional demonstrator (Tech Demo)
Rationale:	
Priority:	Should
History:	

6 Bibliography

- Benyon, D., & Macaulay, C. (2002). Scenarios and the HCI-SE design problem. *Interacting with Computers*, 397-405.
- Broekhuis, M., van Velsen, L., op den Akker, H., Øjvind Nielsen, R., & Andersson, M. (2018). *D2.2: Report on user and stakeholder needs and expectations*. The Council of Coaches Consortium.
- van Loon, J., Broekhuis, M., Konsolakis, K., Huizing, G., ter Stal, S., Katharaju, R., . . . Kotnis, S. (2018). *D9.1: H- Requirement No. 1*. The Council of Coaches Consortium.

A. Focus group protocol

	What	Time	Materials
1	Walk in	10 min	<ul style="list-style-type: none"> Name signs Coffee and tea Cookies Laptop for RRD people Beamer (large meeting room) Screen (large meeting room)
2	Introduction focus group <ul style="list-style-type: none"> - Introduce ourselves - Short explanation of what we are going to do today 	5 min	<ul style="list-style-type: none"> PowerPoint slides
3	Signing of informed consent + agreement to record audio	5 min	<ul style="list-style-type: none"> Informed consent forms Pens
4	Turning on audio recorders	0 min	<ul style="list-style-type: none"> Audio recorders (2x) Extra batteries
5	Introduction round participants <i>Output:</i> Write down names and location of participants. Also ask for gender and age.	10 min	
6	Introduce virtual coaches <i>Aim:</i> Familiarizing the participants with the existence of virtual coaches and their uses. <ul style="list-style-type: none"> - Explain what a virtual coach is - Illustrate with example scenario - Explain that there might be 2 or 3 at once. 	10 min	<ul style="list-style-type: none"> PowerPoint slides
7a	Individual interaction with Demonstrator 1 (Francois slides) <i>Aim:</i> The participants can get some hands-on experience with a dialogue between characters that have backstories and personalities, and that make some coaching comments. <i>Task:</i> Let participants fill in the table for this demonstrator.	5 min	<ul style="list-style-type: none"> 5-10 tablets with Demonstrator 1 General-character tables (on paper)
7b	Plenary discussion on Demonstrator 1 <i>Aim:</i> Getting insight into the interaction experience with the first demonstrator and getting insight into the thoughts of the participants on the backstories / humour / personalities of the coaches.	10 min	<ul style="list-style-type: none"> Note block Pens Prepared questions to guide discussion (use if needed)
8a	Individual interaction with Demonstrator 2 (Web/2D) <i>Aim:</i> The same as 7a, but with a more technically advanced demonstrator.	5 min	<ul style="list-style-type: none"> 5-10 tablets that can reach the Web/2D demonstrator online General-character tables (on paper)

	<i>Task:</i> Let participants fill in the table for this demonstrator.		
8b	Plenary discussion on Demonstrator 2 <i>Aim:</i> Similar to that of 7b, but since the dialogue of this demonstrator is quite similar to the dialogue in Demonstrator 1 the focus for this discussion could go more towards the slightly different manner of interaction in this demonstrator, and its look and feel.	10 min	<ul style="list-style-type: none"> ▪ Note block ▪ Pens ▪ Prepared questions to guide discussion (use if needed)
9a	Plenary interaction with Demonstrator 3 (Council of Coaches functional prototype) <i>Aim:</i> Similar to 7a and 8a, but with a side note that this is just an initial version and that it still needs much more work. <i>Task:</i> Let participants fill in the table for this demonstrator.	5 min	<ul style="list-style-type: none"> ▪ Laptop for the Council of Coaches functional prototype ▪ General-character tables (on paper)
9b	Plenary discussion on Demonstrator 3 <i>Aim:</i> Similar to that of 7b and 8b, but since the main difference is again in the interaction and look and feel, these topics could be given more focus than e.g. the backstory design/dialogues.	10 min	<ul style="list-style-type: none"> ▪ Note block ▪ Pens ▪ Prepared questions to guide discussion (use if needed)
10	Plenary discussion on all demonstrators	5 min	<ul style="list-style-type: none"> ▪ Note block
11	Task on coach characteristics and content <i>Aim:</i> Getting insight into the participants' preferences for the characteristics of a coach and the content that the coaches should provide. <i>Task:</i> Two short card-sorting tasks. Participants get a set of cards for both topics and they are asked to glue them onto a paper that has a scale ranging from 'I find this important' to 'I find this less important'. They are also allowed to write new terms on the paper.	15 min	<ul style="list-style-type: none"> ▪ Note block ▪ Sets of cards ▪ Paper with 'I find this important <- > I find this less important'-scale. ▪ Pens
12	Wrap up <ul style="list-style-type: none"> - Explain next steps taken by RRD - Ask if there are any questions/remarks - Thank participants for participating 	5 min	
13	Turning off audio recorder	0 min	
14	Walk out	10 min	

B. Demonstrator evaluation table

Deelnemer: _____ Voorbeeld: _____	Algemeen	Karakters
 + Positief		
 +/- Neutraal		
 - Negatief		

C. Card-sorting sheet

Deelnemer: _____



D. Cards used in the card-sorting task – Characteristics

Deskundigheid	Heeft humor
Ervaringsdeskundige	Maatje
Vriendelijkheid	Achtergrondverhaal <i>(kan wat over zichzelf vertellen)</i>
Karaktertrekjes <i>(bijvoorbeeld: typische uitspraken, bewegingen, stopwoordjes)</i>	Leeftijd
Geslacht	Uiterlijk

E. Cards used in the card-sorting task – Coaching topics

Recept ideeën	Tips voor bewegen
Sociale activiteit tips	Hersengymnastiek
Herinneren aan afspraken	Medicijn herinneringen
Ideeën voor uitjes	Voedingsdagboek
Kletsen over koetjes en kalfjes	Kletsen over hobby's
Praten over emoties of stress	Over het weer vertellen

Acknowledgements



The Council of Coaches project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement #769553. This result only reflects the author's view and the EU is not responsible for any use that may be made of the information it contains.

Headings and titles in this document, as well as the Council of Coaches logo use the Comfortaa font, designed by Johan Aakerlund and Cyreal and licensed under the Open Font License¹.

Additional text in this document uses the Roboto font, designed by Christian Robertson and licensed under the Apache License, Version 2.0².

The Council of Coaches logo and Blobmen graphics were *drawn freely* in Inkscape, licensed under the GNU General Public License³.

¹ Open Font License: http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=OFL_web

² Apache License, Version 2.0: <http://www.apache.org/licenses/LICENSE-2.0>

³ Inkscape License Information: <https://inkscape.org/about/license/>