D5.2 First Field Study
D5.5.1 First Study Evaluation

Work package	**WP5 Field Studies and Evaluation**
Task	**Task 5.2 First Field Study** **Task 5.5 Performance Evaluation**
Editor	**Linda Burgmans, Iris Geerts**
(co-)authors	**Liselore Snaphaan**
Public / confidential	**Confidential**

Project PLAYTIME

The research leading to these results has received funding from the AAL Programme of the European Union and by the Austrian BMVIT/FFG under the Agreement no 857334, ZonMw (the Netherlands) and funding from Brussels (Belgium). It reflects only the author's view and that the Union is not liable for any use that may be made of the information contained therein.

30/09/2018

PLAYTIME partner			organisation
01	JR		JOANNEUM RESEARCH Forschungsgesellschaft mbH DIGITAL – Institut für Informations- und Kommunikationstechnologien, 8010 Graz
02	FAM		FameL GmbH Steinbruchweg 20, A-8054 Seiersberg
03	SVD		Sozialverein Deutschlandsberg Kirchengasse 7, A-8543 Deutschlandsberg
04	GGZ		Geestelijke Gezondheidszorg Eindhoven en de Kempen Postbus 909, 5600 AX Eindhoven, The Netherlands
05	TIU		Stichting Katholieke Universiteit Brabant, Tilburg University PO Box 90153, 5000 LE Tilburg, The Netherlands
06	MCR		McRoberts BV. Raamweg 43, 2596 HN The Hague, The Netherlands
07	MBY		MindBytes F. Rooseveltlaan 348 B8, 9000 Ghent, Belgium
08	GEU		Ghent University Sint-Pietersnieuwstraat 25, 9000 Gent, Belgium

Contents

1 Executive Summary

The AAL funded project Playtime consisting of partners within Austria, Belgium and the Netherlands develops an integrated serious game of personalized emotion-oriented training modules to stimulate cognitive processes, to address physical activities and foster social inclusion of people living with dementia. The motivation of people living with dementia is primary triggered by the following three aspects of Playtime: (1) positive affection achieved from social engagement in playful group gatherings, (2) multimodal online training modules, including the multimodel training module Amicasa, the social-emotional game module SERES Dementia™, and an attention module with analyses of eye tracking movements, and (3) user feedback, in terms of physical performance and physical activity measured by the MoveMonitor and MoveTest (i.e. motion module).

The first field study of the Playtime project aimed to evaluate the usability, feasibly. appropriateness, and acceptability of the alpha prototype of Playtime by means of a Living Lab method. Furthermore, it evaluated the usability of the user feedback, in terms of physical performance and physical activity, for determining personalized recommendations.

In the Netherlands and Austria, 18 people with dementia and their informal caregiver were included in this study. Each participant received the alpha prototype of Playtime to try at home for a period of two weeks. The participants were also asked to join every week one group session (two sessions in total).

After the two weeks of testing, the alpha prototype was evaluated by means of a reflective semi-structured interview at study participant's home. The topic list of this interview contained several open and closed questions regarding the four evaluation areas (usability, feasibility, appropriateness and acceptability) and study participants experiences in general.

Results of the field study show that there are few issues with respect to usability that need particular attention, such as the login/password structure, language issues and the front page of the Tablet PC. The acceptability of the alpha prototype was influence by its usability: many participants want to play Playtime again but only when the major usability issues are solved. Results with respect to appropriateness show that personalizing more exercises would help to create an appropriate difficulty level for each individual in. Finally, for feasibility is was found that participants found it doable to play Playtime frequently. They advise others to play at least two times per week.

The user feedback, in terms of physical performance and physical activity, was found to be usable for determining personal recommendations: both the data of the MoveTest and MoveMonitor were relevant for detecting differences within subject and can therefore be used to customize the content of Playtime to personal needs. For example, a subject with long sitting periods, can be motivated/advised to make more with short walks

2 Introduction

Dementia is a broad category of neurocognitive disorders characterized by a long term and often gradual decrease in the ability to think and remember. Other symptoms include impaired language, personality changes, emotional problems, behavioral disturbances, and decreases in motivation (Prince, Albanese, Guerchet, & Prina, 2014). The most common forms of dementia are Alzheimer's disease (<70%), Vascular disease (<20%), Lewy Body Dementia (<5%) and frontotemporal dementia (2%) in descending order of occurrence (Prince et al., 2014). These diseases are progressive and slowly, but severely affect a person's brain, and thus affect his or her ability to live a normal live. Advancing age is the main risk factor for most forms of dementia, and with the ever increasingly aging population, the prevalence of dementia worldwide is expected to nearly double every 20 years to 65.7 million in 2030 and 115.4 million in 2050 (Prince et al., 2014). This expected increase will have profound societal challenges in the sense of costs connected to the care of dementia, the quality of life of people with dementia, and the burden on family care givers.

Currently, no disease modifying drugs for dementia are available and pharmacological treatment is limited to therapies that alleviate the symptoms. However, these treatments are not efficacious in all clients and may introduce undesirable side-effects (Galimberti & Scarpini, 2010). Non-pharmacological (or psychosocial) interventions, such as serious games, are therefore appealing alternatives or add-ons as studies suggest that (1) physical games (games that promote physical fitness) can positively affect several health areas of the players with dementia, such as balance and gait (Padala et al., 2012), and voluntary motor control (Legouverneur, Pino, Boulay, & Rigaud, 2011); (2) cognitive games (games which target cognitive improvement) can improve a number of cognitive functions, such as visuo-spatial abilities (Yamaguchi, Maki, & Takahashi, 2011); and (3) both physical and cognitive games can have a positive impact on social and emotional functions, for instance they can improve the mood and increase positive affect and sociability (Boulay et al., 2011; Yamaguchi et al., 2011) and reduce depression (Calvo et al., 2011). Very few studies investigated the effects of the use of games for social-emotional health (which encourage the players to link with their friends and/or improve their social and emotional life) in dementia, but the results are promising (Boulay et al., 2011).

Despite these results, there is to date no serious game geared towards people living with dementia that stimulates physical health, cognitive performance, and social-emotional functioning at the same time. Therefore, the AAL funded project Playtime consisting of partners within Austria, Belgium and the Netherlands develops an integrated serious game of personalized emotion-oriented training modules to stimulate cognitive processes, to address physical activities and foster social inclusion. The objective of the project Playtime is to motivate people with dementia to enter a positive feedback cycle of periodic training with sensors that enable diagnostics on a daily basis, and to receive recommendations on the basis of these data that propose more personalized and better suited exercises for improved training. The motivation is primary triggered by the following three aspects of Playtime: (1) positive affection

achieved from social engagement in playful group gatherings, (2) multimodal online training modules, including a cognitive, social-emotional and physical exercises, to offer the user playful experience at home and group gatherings, and (3) user feedback, in terms of physical performance and physical activity, that provide diagnostics to determine personalized recommendations and, in turn, optimize user experience.

The serious game Playtime may only provide benefits to people living with dementia if it is easy to use, accepted, useful, and feasible to fit into users' daily life (Meiland et al., 2017). Yet, a number of studies showed that people with dementia have problems in using many of the serious games currently available on the market. These difficulties derive from the fact that they are seldom developed considering the users' needs and context (Robert et al., 2014). Understanding the daily context of users in the development of new innovations is therefore essential to cater to their, often complex needs. Moreover, the first-hand perspective from the person with dementia itself is rarely sought in the design process of innovations (Topo, 2009). Instead, the designer or an informal caregiver generally voice product or service evaluations on their behalf. This absence of a first-hand perspective and the need to develop and evaluate innovations as an integral part of daily context of people with dementia motivates the use of the Living Lab method for Playtime. A Living Lab represents is a user-centered research methodology in which multiple stakeholders are all collaborating for creating, prototyping, validating and testing new innovations in real-life contexts (Bergvall-Kareborn & Stahlbrost, 2009; Markopoulos & Rauterberg, 2000), with the goal to develop the innovation further for a market introduction (Leminen, Westerlund, & Nyström, 2012). We rarely see people with dementia getting involved in Living Labs, nevertheless, they are perfectly capable to do so when adequately prepared (Span et al., 2013).

Given the preceding paragraphs, it is clear why the serious game Playtime needs to be carefully evaluated by the use of a Living Lab method. Therefore, this first field study aims to evaluate the alpha prototype of Playtime in order to retrieve insights on its usability, feasibility, appropriateness and acceptability in real-life environments. Furthermore, it will evaluate the usability of the user feedback, in terms of physical performance, physical activity and eye-tracking movements, for determining personalized recommendations. The results of this study are input for the development of a beta prototype of Playtime that fits with the user's needs and context.

3 Research Questions

The research questions of this study can be formulated as follows:

(1) How do people living with dementia evaluate the alpha prototype of PLAYTIME with regard to usability, feasibility, appropriateness and acceptability?

(2) Which user feedback generated by the alpha prototype of PLAYTIME is usable for determining personalized recommendations?

4 Methods and materials

4.1 Study design

For this research, an explorative in-context study was conducted to evaluate users‘ experiences with the Playtime alpha prototype and to look for improvement and personalization opportunities. The Living Labs were positioned in the Netherlands and Austria and people living with dementia were involved as co-creators rather than subjects of study (Almirall, Lee & Wareham, 2012). Each participant in both the Netherlands and Austria was asked to test Playtime during two group gatherings and at home for a period of two weeks.

4.2 Participants

A total number of eighteen persons with dementia participated in the Living Labs of the Netherlands (n=9) and Austria (n=9). All participants had a diagnosis of mild dementia with a cut-off score above 22 on the MMSE (Folstein, Folstein, & McHugh, 1975) and lived at home with an involved informal caregiver. Persons with dementia were not included if they had frontotemporal lobar degeneration (FTLD), visual and auditory processing disorders, and insufficient physical abilities to perform the movement exercises of Playtime. The physical ability of a person with dementia was assessed using three controlled tests of the MoveTest (balance, gait, chair rise) with a required score of at least one point on all tests. General characteristics of the individual study participants are included in Table 1.

The research was carried out in accordance with the latest version of the Declaration of Helsinki. All participants and their informal caregivers gave written informed consent prior to participation to the study, which was approved by a local Ethical Committee of the Netherlands (School of Social and Behavioral Sciences of Tilburg University) and Austria (Ethikkommission der Medizinischen Universität).

Table 1. Characteristics of study participants.

	The Netherlands (n=9)	**Austria (n=9)**
Female, n (%)	5 (55.5%)	5 (55.5%)
Age (years), mean ± SD	73.8 ± 8.0	76.2 ± 8.2
Diagnose, n (%)		
Alzheimer's disease	5 (55.6%)	7 (77.8%)
Mixed dementia	4 (44.4%)	3 (33.3%)
Level of education, n (%)		

Primary education	1 (11.1%)	4 (44.4%)
Secondary education	5 (55.6%)	5 (55.6%)
Higher education	3 (33.3%)	0
MMSE score, mean ± SD	24.8 ± 1.7	25.0 ± 1.5

4.3 Materials

The alpha prototype of the serious game Playtime contains an interactive mat for group gatherings and an application on a Tablet PC consisting of multimodal trainings modules. The trainings modules involve:

(1) Multimodal training module (Amicasa): this module is the core component of Playtime. It involves the person with dementia into different training modules consisting of physical and cognitive exercises (e.g. multiple choice, puzzles, spot-the-difference, memory, and knowledge-based questions). Each trainings module is based on one theme and can be played at three different levels (A, B or C).

(2) Social-emotional game module (SERES Dementia^{TM}): this module involves the person with dementia or/and the caregiver into a social-emotional context and asks for decision-making and reflection upon social imaginative scenarios. The module is applied complementary to the training units and as a serious game trains the social-emotional awareness of the user.

(3) Motion module (MoveTest and MoveMonitor): this module involves the person with dementia in several motion analysis procedures: it is capable to measure physical performance during controlled tests and physical activity of the persons with dementia within a longer period of time (1 week in daily life). Both procedures rely on a sensor unit worn around the waist with an elastic strap. With MoveTest measurements, markers can be set in de measurement data to highlight the data to be used for further analysis. Data collected with the MoveMonitor is analyzed entirely.

(4) Attention module: this module will provide a gaze interface for games that at the same time provide assessment indicators for cognitive control. Furthermore, it provides an analysis of the eye tracking movements of the person with dementia which can be used as an indicator for the degree of dementia within a specific period of time.

4.4 Procedures

In the Netherlands, 9 study participants were recruited by three main methods, namely by (1) presenting Playtime at the Alzheimer Café in Eindhoven, (2) using the Innovated Dementia focus group, and (3) using the class ‘Help me even herinneren’ of GGzE (GGZ). In Austria participants were recruited from the Sozialverein Deutschlandsberg (SVD); 6 participants were

active or new MAS-clients from the Sozialverein, 3 were recruited trough contacts of the mobile care.

In both the Netherlands and Austria, an appointment was made and two project members visited eligible study participants at home. The main purposes of this first home visit were:

(1) To obtain the informed consent

(2) To check the inclusion criteria

(3) To conduct a personal interview

(4) To introduce the main project objectives

(5) To demonstrate and explain the alpha prototype of Playtime

(6) To conduct the MoveTest and explain the MoveMonitor.

(7) To confirm the time and place for the group gatherings

(8) To schedule an appointment for the second home visit.

After the first home visit, a testing period of two weeks followed. All study participants were asked to wear the MoveMonitor 24 hours, for seven day, and to test the alpha prototype of Playtime at home and during two groups sessions. Each group session averaged 90 minutes in time and was facilitated by members of the project. During the testing period, project members were available for questions before and after every group session, and gave support by telephone. Two times it was necessary to make an appointment for further assistance in the Netherlands.

After the testing period, each study participant was visited again at home by two project members to evaluate the alpha prototype of Playtime by semi-structured reflective interviews. The topic list of these interviews contained several open and closed questions regarding the four evaluation criteria (usability, feasibility, acceptance, and appropriateness). Study participants were also asked to reflect on their experiences with each trainings module in particular: the multi-module trainings modules amicasa, the social-emotional game module SERES Dementia™, the MoveMonitor and MoveTest, and the eye-tracking exercise of the attention module. After each closed question, study participants had the opportunity to elaborate on their answer. When the informal caregiver and the person with dementia provided different answers on the closed questions, the answer of the informal caregiver was taken into account as they had often a better overview of the testing period and were able to further elaborate on the question. In the Netherlands, all evaluations were audio-recorded with permission of the participants for the convenience of transcribing. In Austria, all interview responses were summarized and translated into English.

In the figure below, a summary of the study procedure is provided, including the total amount of hours people with dementia and their informal caregivers invested to participate in this study.

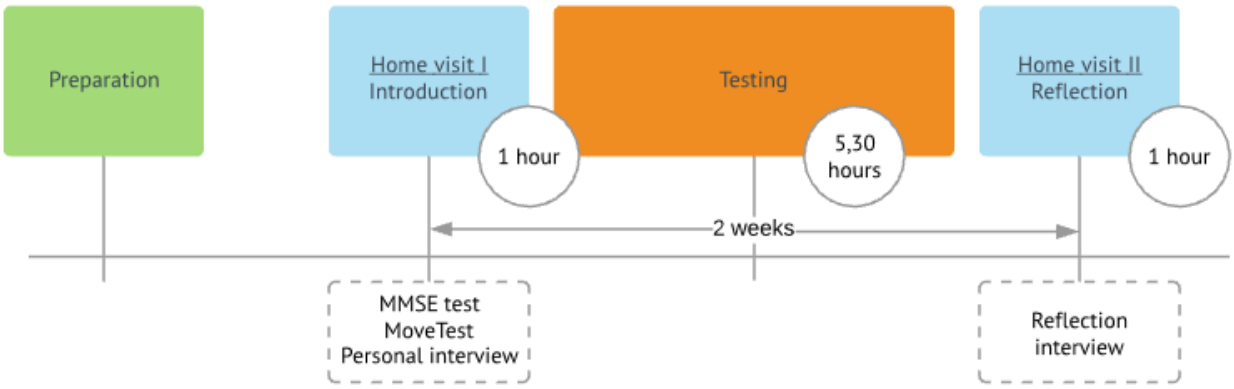

Figure 1. Summary of study procedures.

4.5 Outcome measures

4.5.1 Usability

Usability was primarily measured by the 10-item System Usability Scale (SUS) (Brooke, 1996). This questionnaire measures usability along three dimensions (i.e. effectiveness, efficiency and satisfaction) on a five-point Likert scale, ranging from strong disagreement (1) to strong agreement (5). Additionally, the usability of the manual of Playtime was measured by one self-constructed item (i.e. *"the manual was clear and provided sufficient support"*) on a five-point Likert scale, ranging from strongly disagree (1) to strong agree (5).

4.5.2 Acceptability

Acceptability was measured by a self-constructed item (e.g. "*I enjoyed playing Playtime*") on a five-point Likert scale, ranging from strongly disagree (1) to strongly agree (5). Other outcome measures of acceptability were several open-ended questions about each multimodal trainings module of Playtime (e.g. "*What do you think of the eye-tracking exercise?* "), and the frequency of use and the time spend playing on amicasa and SERES Dementia™ (e.g. "*How often and how long did you (approximately) play SERES?*"). Study participants were also asked to grade Playtime on a five-point Likert scale, ranging from (1) excellent to (5) insufficient.

4.5.3 Appropriateness

Appropriateness was measured by seven self-constructed items on a five-point Likert scale, ranging from strongly disagree (1) to strongly agree (5). Example items are: "*The text in the apps was not readable for me*", "*Playtime is suitable for people with dementia*", and "*Playing Playtime in a group had no added value for me*".

4.5.4 Feasibility

The feasibility of Playtime for everyday use at home was measured by four open-ended questions about the frequency of use and time spend playing on Amicasa and SERES Dementia^{TM}, and one open-ended question on the ideal time dedicated to playing Playtime.

4.5.5 Motion parameters

During the first field study, 2 different measurement methods were used to gain insight in the movement parameters of the subjects.

Physical performance under controlled conditions ('what you can do') was assessed using a MoveTest. Specifically, the Short Physical Performance Battery (Guralnik et al., 1994) (SPPB) was used as this provides sub-scores for a subjects' balance, gait and repeated sit-to-stand capabilities (all sub-scores range from 0 (not able to perform) to 4 (top quartile compared to peers)) and a total score for overall performance (0-12 scale, summation of the sub-scores). Next to the sub-scores of the SPPB, the MoveTest also provides durations of the separate tests.

Physical activity in daily life ('what you actually do') was assessed using a MoveMonitor. Classification algorithms classify different bouts of activity (walking, stair walking, cycling, shuffling, sitting, standing, lying and non-wearing). Of these bouts, the number and durations (mean and total) are used as outcome parameters in Playtime. Additionally, for the walking bouts, steps will be calculated (total amount and amount normalized to 24 hours).

4.6 Data analysis

4.6.1 Reflective interviews and field notes

In the Netherlands, the reflective interviews were audio-recorded and verbatim transcribed. A systematic theoretical coding analysis of the Dutch transcripts and translated Austrian summaries was carried out by applying the three steps method described by Ritchie and Lewis (2013): (1) data management, (2) descriptive accounts, and (3) explanatory accounts. Using the software package Atlas.ti (version 7.5.3), text segments were compared and contrasted, and classified into categories based on a priori codes derived from the literature. Emergent sub-codes were then developed based on patterns within each concepts and which were relevant to the literature. During the coding process, interview transcripts and translated Austrian summaries were analyzed by both the researcher of TIU and the healthcare professional of GGZ, allowing for continuous discussion and re-coding of some text fragments.

4.6.2 Questionnaires

The data obtained by the SUS and self-constructed items was analyzed by computing several descriptive statistics (e.g. means, medians, standard deviations, minimum and maximum scores) by the statistical analysis software package SPSS, version 19. Additionally, for analyzing the data obtained by SUS, a scoring template that turns the raw individual survey

ratings of all study participants into a single SUS score was used. The template considered Brooke's (1996) standard scoring method, which is manipulating statement ratings to get them a common 0-4 rating, then multiplying the sum by 2.5 to get a score that can range from 0-100. In evaluating the usability of Playtime, a SUS score above 68 was considered above average, and below 68 was considered below average (Brooke, 2013).

4.6.3 Motion parameters

Descriptive statistics are used to describe results. A t-test is used to compare differences between groups (a p-value < 0,05 is considered to be significant) and correlation coefficients are calculated between parameters to signal relations.

5 Results

5.1 Quantitative results

5.1.1 Questionnaires

To evaluate the usability of Playtime in general, all study participants filled in the 10-items of the SUS questionnaire. By using Brooke's (1996) standard scoring method, a single SUS score of 63.9 was calculated. A score of 63.9 indicates that the usability of Playtime is below average.

Study participants also filled in nine closed questions with respect to acceptability, appropriateness and the usability of the manual of Playtime (see Table 2). When considering the minimum and maximum scores of each question, it seems that study participants answers' are widely dispersed. The reason for this wide range of scores can be found in the qualitative results of this study: The low usability of Playtime ensured that some participants were not able to play Playtime, which in turn negatively influenced the acceptability, appropriateness and the usability of the manual of Playtime. In section 5.2, the quantitative results of this study are provided.

Table 2. Descriptive statistics closed questions.

	Min	Max	Median
The manual was clear and provided sufficient support	2.00	5.00	4.00
The text in the apps was not readable for me	2.00	5.00	4.00
I could understand the language	3.00	5.00	4.00
The sound in SERES and Amicasa was good	1.00	5.00	4.00
The assignments and exercises were too difficult	1.00	4.00	2.00
Playtime stimulates my social, memory and physical capabilities	2.00	5.00	4.00
Playtime is suitable for people with dementia	3.00	5.00	4.00
Playing Playtime in a group had no added value for me	2.00	5.00	4.00
I enjoyed playing Playtime	2.00	5.00	4.00

1=strongly disagree, 2=disagree, 3=neither disagree nor agree, 4=agree, 5=strongly agree

5.1.2 Motion parameters

5.1.2.1 Acceptability

The data generated by the MoveTest and MoveMonitor give limited insight in the first question as this can be best answered by feedback from the subjects themselves in the form of questionnaires or interviews. One parameter that does give insight in the acceptability is wearing compliance of the MoveMonitor. The periods when a sensor is removed by the subject are automatically detected (non-wearing). Using this parameter the wearing compliance can be calculated as the wearing time relative to the measurement time. The wearing compliance in this study is high (95 ± 3 % in the Netherlands and 82 ± 23 % in Austria, see figure 2)

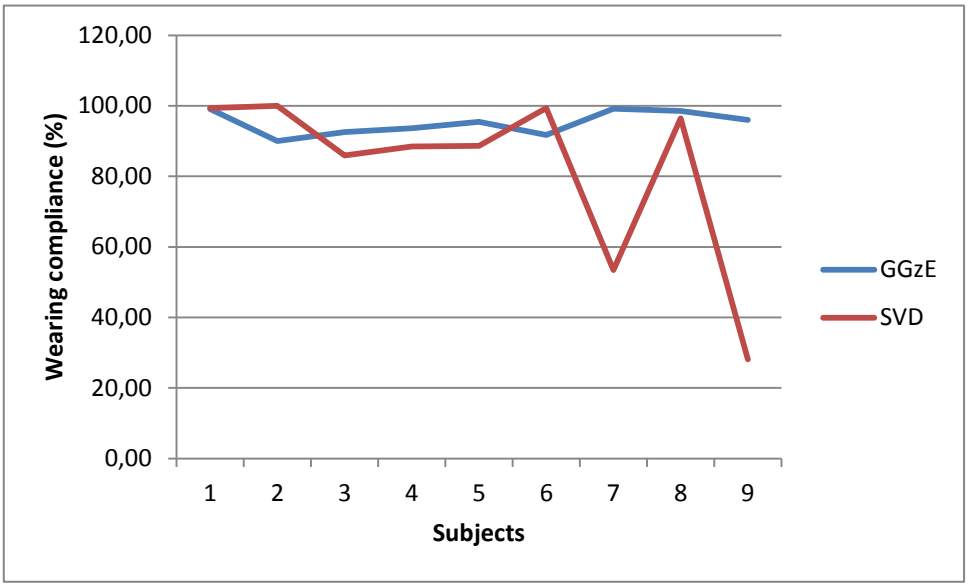

Figure 2. Wearing compliance in the Netherlands (GGZ) and Austria (SVDL).

Of the 18 subjects in total, 2 subjects wore the sensor less than requested (28% and 53%), while the other 16 subjects wore the sensor very well. Subjects were instructed to remove the sensor during aquatic activities (showering/bathing, swimming, etc.), so wearing compliance of 100% is not expected. Overall, these results show the users are willing to wear the sensor for the requested period of time (1 week).

5.1.2.2 Personalized recommendations.

The outcomes of the MoveTest and MoveMonitor can be used to personalize the content of the Playtime solution as different aspects of motion are captured in objective parameters.

The outcomes of the MoveTest can be used to personalize the content of the Playtime solution to increase physical performance. The SPPB consists of 3 separate tests: balance (3 conditions with increasing difficulty), gait and repeated sit-to-stand (chair stand). Figure 3 and table 3 show the scores of the different test and the total test. Additionally, table 1 shows the maximum possible scores. It can be seen that the maximal scores are not reached and that there are

differences between the subjects in the Netherlands (GGE) and Austria (SVDL). This shows that there is no ceiling effect and that the tests are capable of detecting relevant differences between subjects. The these (sub) scores can be used to personalize the content.

Table 3. SPPB (sub) scores and maximal possible scores.

		Maximal possible	GGZ	SVD
Sway	Total	4	2.8	3.4
	Condition 1	1	1.0	1.0
	Condition 2	1	0.9	1.0
	Condition 3	2	1.0	1.4
Gait		4	2.8	2.5
Chair stand		4	1.4	1.9
Total SPPB score		12	7.0	7.1

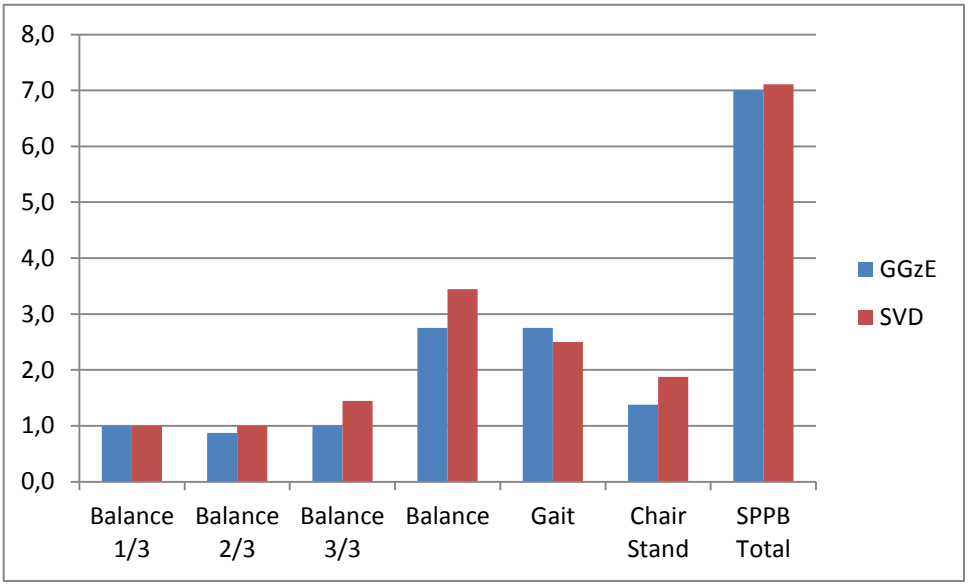

Figure 3. Mean (sub) scores of the SPPB in the Netherlands (GGZ) and Austria (SVD).

The outcomes of the MoveMonitor can be used to personalize the content of the Playtime solution to increase physical activity (PA). The MoveMonitor classifies different activity bouts. These can be represented as number of periods (figure 4), mean durations of those periods (figure 5) or total durations of the separate activities (figure 6). For the walking periods, the steps taken can be calculated. For comparison between subjects, the steps are normalized to 24 hours (figure 7). It can be seen that in these basic parameters, differences between populations and subjects can be seen. This information can be used to customize the content of the Playtime application to personal needs. For example, a subject who has few, but long sitting periods, can be motivated to break up these long sitting periods with short walks.

From this data, a lot more parameters can be calculated, but care should be taken not to 'over-inform' the subject. The feedback presented to the subject should be understandable and the subject should be able to adapt his/her daily activities based on the feedback from the Playtime suite.

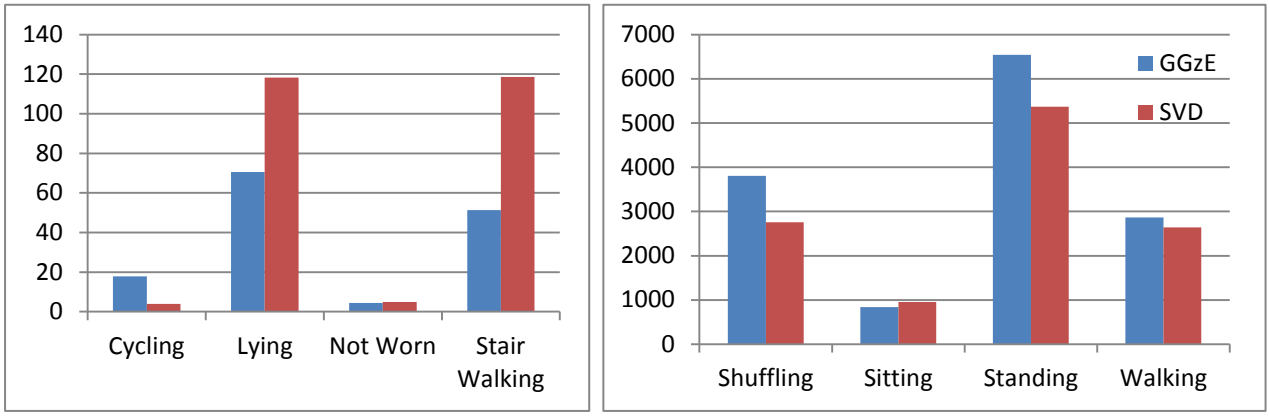

Figure 4. Mean number of periods of different activities, including periods in which the MoveMonitor was not worn in the Netherlands (GGZ, blue) and Austria (SVD, red).

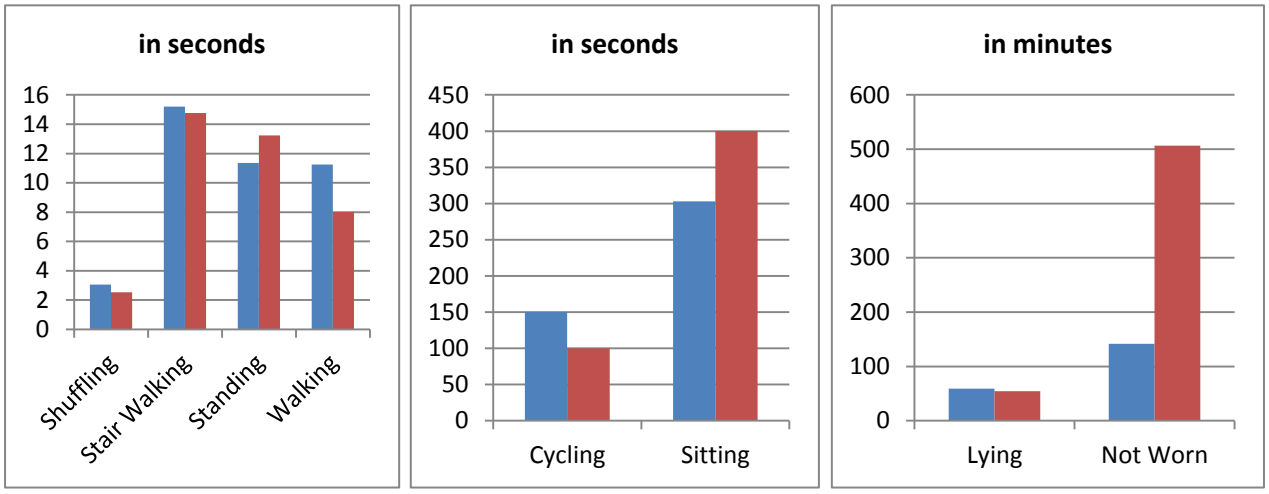

Figure 5. mean bout durations of the classified activities in the Netherlands (GGZ, blue) and Austria (SVD, red).

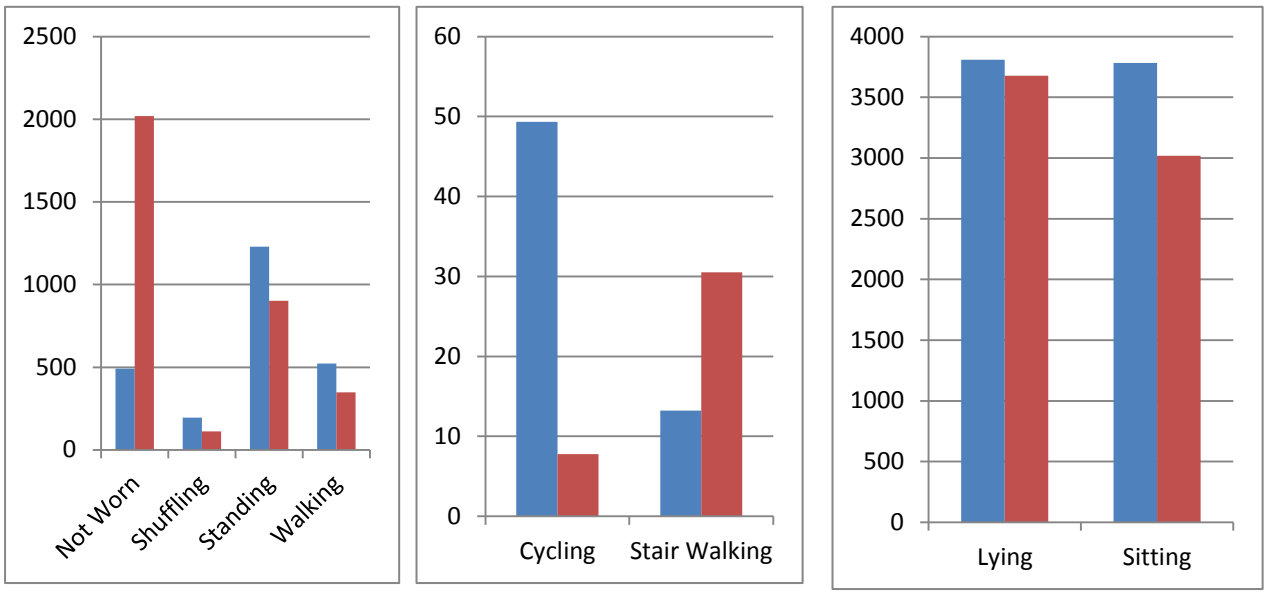

Figure 6. Total durations (in minutes) of the classified activities in the Netherlands (GGZ, blue) and Austria (SVD, red).

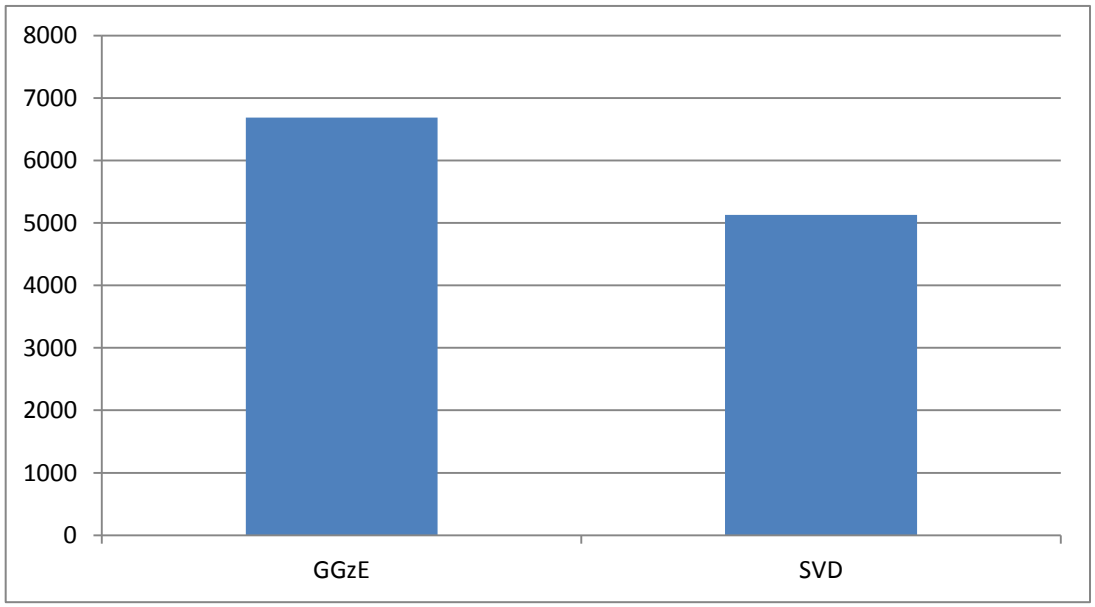

Figure 7. Average steps normalized to 24 hours for the Netherlands (GGZ) and Austria (SVD).

5.1.3 Relations to other outcome measures

The strength of a multi-modal approach like in Playtime is that multiple outcome measures can be used for feedback. Additionally, outcome measures can be combined to personalise the content of the Playtime app.

In this first field study, the MMSE of all subjects was collected and a correlation of 0,8 was found between the MMSE score and the mean sitting duration suggesting that subjects with a

lower MMSE have longer sitting periods. Prolonged sitting periods have negative health effects, so when a subject with a low MMSE score starts using Playtime, focus on breaking up sitting periods can be incorporated from the start.

Another way of using other output measures is to see whether there are differences in motion parameters between subjects with high scores vs subject with low scores on other output measures. Again, as an example, the MMSE was used. Subjects were divided in 3 groups: low MMSE (6 subjects with lowest scores), high MSSE (6 subjects with highest scores) and middle MMSE (remaining 6 subjects). Table 4 shows the significant differences ($p < 0,05$) that were found between the low MMSE and high MMSE group. Knowledge of these differences can be used to focus the motion part of the Playtime app more are less on these areas, depending on the MMSE score of the individual using Playtime.

Table 4. Significant differences between subjects with low and high MMSE scores.

Parameter		p-value
Number of periods		
	Shuffling	0,046
	Standing	0,020
	Walking	0,023
Mean durations		
	Sitting	0,023
Total durations		
	Standing	0,039
SPPB		
	Sway, 3rd condition	0,030
	Sway, overall score	0,018

5.2 Qualitative results

5.2.1 Usability

Langue was found to a big issue when evaluating the usability of Playtime. Reasons for this were:

(1) Every time the participant wanted to play Playtime, the langue had to be changed.

(2) In the Netherlands, the physical exercises in Amicasa were translated in Dutch, but study participants noted that they also would like to have a Dutch voice-over instruction.

(3) In Austria, the language within SERES was found to be an issue as it contained too much translation errors.

Study participants mentioned that they had to log in too often. Therefore, it would be easier to log in if the system would have a remember me function. It was also mentioned by participants that the login code should be easier to remember: no capitals, no underscores, no foreign language, et cetera.

Using the Tablet PC was difficult for some participants, mainly because they had not frequently used it before. The Tablet PC should therefore be simplified: updates should not appear, only the Playtime application should be on the front page, the Tablet PC needs to have a cover, and a touch screen pen need to be included

The WIFI connection was also found to be an issue by some participants, especially reconnect or connecting with the WIFI. Next to this, some participants mentioned that the audio of the Tablet PC was not clear enough.

Study participants mentioned with respect to the exercises the following:

(1) The physical exercises were often performed

(2) The puzzles and memory were too small for some participants.

(3) Some exercises can use more instruction.

(4) Study participants liked the games and were motivated to play them.

(5) Although there was enough diversity in exercises, study participants would like to be able to select, for example, only the physical exercises or only the puzzles.

Finally, study participant mentioned that the different applications would be easier to find if they could be implemented into one application. This could also solve the confusing different application names: SERES, Amicasa, Articulate.

5.2.2 Acceptability

The question '*Would you like to test Playtime again?*' was positively answered by all study participants. They mentioned that is was great to stay active by playing games, especially the physical exercises. However, almost all participants mentioned that the usability needs to be improved to gain more acceptability.

Some study participants noted that they would like to select a difficulty level for the physical exercises as these were now to ease. Also the difficulty level of the puzzles and games was too ease for some participants. Therefore, these participants would like to see a more challenging difficulty level in Amicasa.

Both the study participants and their informal caregivers mentioned that Playtime as nice to play together. Some of them even played every day together.

The group sessions were evaluated differently: while some participants stated that the group sessions were nice to do had an added value, noted others that they perceived the group sessions as less appropriate due to three reasons:

(1) A lot of games were individually orientated and not suitable for the whole group.

(2) The whole group had to play the same difficulty level. For some participants this was too difficult or too easy.

(3) In the Netherlands, some study participants did not like the group on beforehand, but joined for participation in this study.

5.2.3 Appropriateness

The exercises were found to be appropriate for most study participants. However, for some participants, even the most difficult level was still too easy.

Study participants mentioned that there was too much noise in Amicasa to indicate that a level was completed. With respect to text it was mentioned that the Amicasa application could use more text (or other forms to instruct the exercises). The font of some text could also be bigger. One participant mentioned that the contrast of the text was insufficient. Participants mentioned with respect to the SERES application that it contained to a lot of text, which was difficult to remember.

Study participants provided the following feedback to improve the appropriateness of Playtime:

(1) Use of pictograms or audio instead of lots of text.

(2) Some of the exercises in Amicasa should bigger: the puzzles, the memory game and fill in the letter or word.

(3) Some of the graphics are somewhat childish.

(4) Make it possible to select a difficulty level for the physical exercises. Some study participants found it a little bit silly that they had to perform exercises in a chair, while they walked every day 30 minutes.

(5) Make it possible to select a topic in SERES. Now, you have to play topic that does not to your situation, such as driving a car.

5.3 Feasibility

Every study participant was asked to give an indication of how much they had used SERES Dementia™ and Amicasa. The following table shows their frequency the frequency of use per week for every study participant.

Table 4. Frequency of use per week.

Frequency of use (per week)	SERES Dementia™ (n=18)	Amicasa (n=18)
Not	6	1
1 time	4	1
2 to 3 times	4	8
3 to 4 times	4	7

Daily*	1

* Only with respect to the movement exercises (in Amicasa)

Every study participant was also asked to grade Playtime, 1 being excellent, 5 being inadequate. The following table shows the scores provided to the alpha prototype of Playtime in general.

Table 5. Scores provided to Playtime in general.

Score	n=18
1= Excellent	2
2 =Good	7
3 =Sufficient	5
4 =Mediocre	2
5 =Insufficient	2

5.3.1 Multimodal training module (Amicasa)

Starting the game was very difficult for a lot of the participants. In the Netherlands, study participants had to manually change the language every time they wanted to play. The logon code was found too difficult by many of the participants, due to the big and small letter, punctuation marks, and the difficult word. Therefore, participants recommended several time a remember me function to log in.

The lay out of the game was sometimes perceived as somewhat childish. The contrast in the game (the white lettering on the pink background) could for some participants be more clearer. It was also mentioned that the size of the exercises in Amicasa should be bigger to be able to better see the exercises. Exercises that were found to small are the puzzles, the memory game and fill in the letter or word. In the puzzle there is a faded background of the picture. For some participants this was confusing, while others considered it as very helpful. Next to this, the question mark of the exercise fill in the letter was found to be confusing for some participants; it distracted them from making the exercise as they though that a question was hidden behind the mark and clicked on it. Some participants also indicated that the missed a return button. By the math exercises, for example, there was no way to get back to the previous exercise when participants accidently clicked on the continue button.

The exercises where mostly fun to do. However, study participants mentioned that there should be a lot more themes to avoid that themes need to be played repeatedly. It was also advised to create an extra, more difficult, level because the most difficult level was still too easy for some study participants. On the other hand, some exercises of the easiest difficulty level were mentioned to be too difficult. Some study participants explained this by the provided explanatory information: for some exercises more explanation was needed in order to understand the

purpose of the exercise. Next to this, the touch screen of the Tablet PC did not work properly for every study participant. A pen was advised, but did not solve the tapping problem. Although the physical exercises were performed by most of the participants and were considered as useful, it was stated that different difficulty levels could make the exercises more suitable to stay active.

Not all exercises of Amicasa were suitable for the group sessions as these were individually orientated. Members of the group session had, for instance, to wait while one member was puzzling. Study participants therefore suggested a separate group version with multi-player exercises. In addition, study participants mentioned that they would like to select a difficulty level for each player of the group. Now, you had to select one level, which was too easy for some members.

5.3.2 Social-emotional game module (SERES Dementia™)

For SERES the Articulate application was used. Many participants found that very confusing. Some study participants could not play SERES as they could not find the application on the Tablet PC. Other participants mentioned that the Articulate application did not load the SERES game properly, or in such a slow pace that participants quit the game. In Austria, the translation contained many mistakes.

Most of the participants that played SERES noted that it was nice to do. Interestingly, one study participant mentioned that SERES had great potential, and another participant found it very useful, especially when played together. Study participants also provided feedback for improvement. It was it would be great to minimize the text or even get an audio version, as it was a lot to remember. Also was the text too small for some participants, which made it difficult to read. Some of the topics were not suitable for every person. Therefore, selecting a topic instead of gender was suggested for gameplay both at home and during group sessions. .

For the group session in particularly, the SERES application on a Tablet PC was very small. We were able to solve this by using the computer version on a television screen. While reading the story, many participants needed a summary before they could discuss what the best way of coping with the situation was. Participants often stated that a summary would be nice, as there was a lot of text to remember. Talking about these subjects with people in similar situations was found to be pleasant by most of the group participants.

5.3.3 Motion module (MoveTest and MoveMonitor)

The MoveTest was easy to perform for the participants, in the Netherlands the test gave a failed message one time, in Austria two tests failed.

For the MoveMonitor participants were asked to wear the belt of the MoveMonitor for one week. Most study participants found that is was no problem to wear the belt this period, however a few participants thought it was too long. Other remarks regarding the belt of the MoveMonitor were:

(1) The belt itself could be somewhat longer

(2) The Velcro sting

(3) The belt shifts sometimes

(4) Wearing the belt at night is not comfortable.

One suggested improvement was to make it waterproof, so you do not have to remember to remove the belt before taking a shower.

During the first field study, each participant received a report with their own results collected by the MoveTest and MoveMonitor. Many participants mentioned that they would like to receive some advice together with the results in order to be able to actually improve their daily activities. One of the participants said that it now does not seem to have a clear goal within the Playtime suite if you do not know how to achieve improvement. The reports on itself were clear for most. When asked how to improve the report, there was mentioned that the text could be what larger, and there should be no English words in the results.

5.3.4 Attention module

Evaluating the attention module of Playtime was difficult for most participants, because the this module had completed during the first house visit, several weeks after the evaluation. Feedback that was provided by study participants is that the goal of the attention exercise within the game was not clear. Participants also found it unclear that the exercise was finished, they were missing a confirmation that the exercise had been completed successfully. The touch screen for the game was challenging for a lot of the study participants, it did not work properly. The screen with the mask, were the participants had to place their face on a frame, was challenging for most participants. Finally, it was mentioned by study participants that they would play the attention module more often when:

(1) There are several different other attention games

(2) The game is offered in a more playful manner

(3) It is helpful in adjusting or selecting the correct difficulty level of the other games

6 Conclusion and Outlook

6.1 Conclusion

Regarding to the domain **usability;** there are a few topics that needs attention. First the login/password structure with a combination of letters, capital letters, digits, special signs and the length of the login was too difficult. Furthermore, too many times Playtime ask to insert this difficult password. Second usability problem is the language issues. In the Netherland the login, and Amicasa has to be fully translated. Participants were confused because some feedback or errors were in German and some in Dutch. Preferable is to use a system in the same language of the participant. Also the translation of the SERES game into the German language needs to be improved. Third attention point is that some participants did not have much experience of using a tablet so the front page should be easy as possible. For example, all standard apps such as gallery, messages, play store et cetera could be clustered in one map and/or moved to another page. In the front page only the Playtime app should be seen. This possibly leads to less confusion for the participant. The different names for the Playtime app was also confusing, Amicasa/Playtime SERES in the Articulate application was not preferable, some participants did not find the exercise program because of these names being confusing.

Acceptability was influenced by the usability results, many participants wanted to play again but only if the mayor issues are solved as described with the usability section. SERES should not be opened in Articulate because the time that is needed to load SERES is not acceptable for the most participants. Participants do like most puzzles and games, although the usability issues has to take into account. Participants are motivated to play the Playtime suite because it can help them to stay active.

With respect to **appropriateness**, personalizing more exercises would help to create an appropriate difficulty level for each individual in which increases the acceptability. This is mention many times, within math exercises of Amicasa, the physical exercises, and SERES. The amount of text seems to decrease the acceptability for SERES and also the font of the letters (too small) at Amicasa. The sound was appropriate for most participants. Some exercises could use some more explanation, or visualization to understand the assignment better.

For **feasibility** participants mentioned that the frequent of playing playtime is doable. They advise others to play at least two times per week. Some suggest every day or every other day.

As expected, the maximal MoveTest scores for the people with dementia are not reached. The tests are capable of detecting relevant differences within subjects. The outcomes of the MoveMonitor can be used to personalize the content of the Playtime solution to increase physical activity (PA). Both the MoveTest and the MoveMonitor can therefore be used to customize the content of the Playtime application to personal needs. For example, a subject with long sitting periods, can be motivated/advised to make more with short walks. The feedback presented to the subject should be understandable and the subject should be able to adapt his/her daily activities based on the feedback from the Playtime suite.

As the data collected in the first field study is limited, care should be taken to make conclusions, but the differences found already indicate that a multi-modal approach can give more insights than a single-modal approach. The findings of this field test can be used as input for the main field study.

6.2 Outlook

To come to a marked ready product several aspects of the usability, acceptance and appropriateness op Playtime should be taken into account and be improved. In the beginning of 2019, a follow-up main field study of 10 weeks will be performed. The aim of this main field study is to evaluate again the usability, feasibility, appropriateness and acceptability of Playtime, next to users' users' gaming experience and motivation to continue gameplay.

7 References

Almirall, E., Lee, M., & Wareham, J. (2012). Mapping living labs in the landscape of innovation methodologies. *Technology innovation management review*, *2*(9), 12.

Bergvall-Kareborn, B., & Stahlbrost, A. (2009). Living Lab: an open and citizen-centric approach for innovation. *International Journal of Innovation and Regional Development*, 1(4), 356-370.

Boulay, M., Benveniste, S., Boespflug, S., Jouvelot, P., & Rigaud, A. S. (2011). A pilot usability study of MINWii, a music therapy game for demented patients. *Technology and health care*, 19(4), 233-246.

Brooke, J. (2013). SUS: a retrospective. *Journal of usability studies*, 8(2), 29-40.

Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

Calvo, B. F., Castillo, I. C., Serna, A., de Lucena, V. M., & Campos, F. R. (2010). El efecto del formato de intervención individual o grupal en la estimulación cognitivade pacientes con enfermedad de Alzheimer. *Revista de psicopatología y psicología clínica*, 15(2), 115-123

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12(3), 189-198.

Galimberti, D., & Scarpini, E. (2010). Inflammation and oxidative damage in Alzheimer's disease: friend or foe?. *Frontiers in bioscience (Scholar edition),* 3, 252-266.

Legouverneur, G., Pino, M., Boulay, M., & Rigaud, A. S. (2011). Wii sports, a usability study with MCI and Alzheimer's patients. *Alzheimer's & dementia: the journal of the Alzheimer's Association,* 7(4), S500-S501.

Leminen, S., Westerlund, M., & Nyström, A. G. (2012). Living Labs as open-innovation networks. *Technology Innovation Management Review*, *2*(9).

Markopoulos, P., & Rauterberg, G. W. M. (2000). LivingLab: A white paper. *IPO Annual Progress Report*, 35, 53-65.

Meiland, F. J., Bouman, A. I., Sävenstedt, S., Bentvelzen, S., Davies, R. J., Mulvenna, M. D., ... & Dröes, R. M. (2012). Usability of a new electronic assistive device for community-dwelling persons with mild dementia. *Aging & mental health*, 16(5), 584-591.

Padala, K. P., Padala, P. R., Malloy, T. R., Geske, J. A., Dubbert, P. M., Dennis, R. A., ... & Sullivan, D. H. (2012). Wii-fit for improving gait and balance in an assisted living facility: a pilot study. *Journal of aging research*.

Prince, M., Albanese, E., Guerchet, M., & Prina, M. (2014). World Alzheimer Report 2014: Dementia and risk reduction: An analysis of protective and modifiable risk factors. *London: Alzheimer Disease International.*

Span, M., Hettinga, M., Vernooij-Dassen, M., Eefsting, J., & Smits, C. (2013). Involving people with dementia in the development of supportive IT applications: a systematic review. *Ageing research reviews,* 12(2), 535-551.

Yamaguchi, H., Maki, Y., & Takahashi, K. (2011). Rehabilitation for dementia using enjoyable video-sports games. *International psychogeriatrics*, 23(4), 674-676.

.