PILOT TEST GUIDELINES

Protocol and instructions

Project acronym:	POSTHCARD
AAL JP project number:	AAL-CALL-2017-045
Deliverable Id:	D4.2a
Deliverable Name:	Pilot Test Guidelines
Status:	Running document
Dissemination Level :	Public
Due date of deliverable:	M12, M20
Actual submission date:	April 2019
Author(s):	L. van den Berg, T. Dadema, H.H. Nap, A.L. Abdel Alim
Lead partner for this deliverable:	Vilans
Contributing partners:	HUG; UNIGE

Project co-funded by the AAL Joint Programme and the National Authorities and R&D programs in The Netherlands, Belgium, Switzerland and Romania, under the Grant Agreement number AAL-CALL-2017-045

Contents

Version history

Version	Authors	Date	Description
0.1	Bergh, Abdel Alim, Dadema, Nap	29-04-2019	A first version was made and will be finalized after a working mid- to high-fidelity prototype is ready.
0.5	Nap, Dadema	06-07-2019	V0.1 was updated
0.8	Nap, Dadema	10-10-2019	Version ready for final adjustments after working prototype is shared.

Project partners

Partner #	name	abbreviation	country
1	University Hospital of Geneva	**HUG**	Switzerland
2	University of Geneva	**UNIGE**	Switzerland
3	Father Equipment	**FTH**	Romania
4	ConnectedCare	**CCARE**	Netherlands
5	University of Twente	**TWE**	Netherlands
6	Alzheimer Nederland	**AZN**	Netherlands
7	Vilans	**VIL**	Netherlands
8	Conectar	**CON**	Belgium

Summary

This document describes a validation plan for the POSTHCARD, developed within the WP4 of the POSTHCARD project. The goal of the validation is to validate the knowledge acquisition due to the POSTHCARD, the basic characteristics, acceptance of the platform, their experience when playing the game and the accessibility of the user interface among the intended end-user groups (informal carers). The document includes plans for empirical testing of the Alpha and Beta version of the platform among end-users in three countries: the Netherlands, Belgium and Switzerland. Several qualitative and quantitative research methods will be combined in order to answer the desired research questions.

This is a running document and will be updated whenever the first mid-fidelity interfaces are ready for end-users to control settings and interact with the game elements, then a heuristic evaluation is planned to gain insight into the usability of the POSTHCARD interfaces.

Introduction

To assure the success of any product, it is vital to evaluate the product with a wide range of potential end-users before it is released to market. Usability tests provide a rich and large amount of input for improvements of the design of the system's User Interface (UI) and input devices. In case of a learning platform, like the POSTHCARD, it is also vital to gather an understanding about the extent to which the platform supports learning and also about the platform experience of the informal caregivers per country. Ideally, POSTHCARD should foster learning, measured by an increase in the platform scores per play session and a positive perceived learning experience. Furthermore, POSTHCARD should provide a similar or at least an above scale mid-point player experience on the end-device: laptop/desktop.

A link to the POSTHCARD platform will be provided and the player experience with will be studied with the end-users at three testing sites. The user partners will run the field trials, using the scenarios and research protocols. A main focus will be on the player experience of the end-users (in total 30 participants per testing site) with POSTHCARD and learning (advancements) in the platform. Error reports and users' satisfaction with the system and platform will be gathered and the perceived benefits of POSTHCARD will be evaluated.

Research questions and goals

The aim of the evaluation study of the POSTHCARD platform is to gather an insight in:

1. The knowledge acquisition due to playing POSTHCARD.
2. The basic characteristics of the POSTHCARD: time played for the whole platform, time played per session and challenge (task), number of sessions and challenges played per user, overall score, score for individual sessions and challenges, level (challenge) success.
3. Acceptance of the POSTHCARD platform
4. The player experience when playing the POSTHCARD platform.
5. The accessibility of the POSTHCARD user interface.

Mean Player Experience (Positive Affect, Immersion, etc.) and mean Usability (Effectiveness, Efficiency, Satisfaction) should be above scale mid-point.

Methods

The POSTHCARD pilot evaluations with older people will be conducted in Belgium, Switzerland and The Netherlands. Extensive user testing will be performed with the laptop/desktop. The main focus of the user tests will be on the:

- *knowledge acquisition (learning)*

- *user acceptance,*

- *player experience* of the POSTHCARD platform,

- *accessibility (usability)* of the POSTHCARD UI (change settings, select avatar, etc.).

The user tests will take place at two occasions: The Alpha test in [DATE] and the Beta in [DATE]

Participants

Alpha testing [DATE]: Five participants (18+ years of age) per country will be invited for the testing of the Alpha version of the platform. Participants that participated in the previous Interviews are allowed to participate in the play sessions.

Beta testing [DATE] Twenty-five participants (18+ years of age) will be invited for the studies per country, of which 5 will participate in experimental sessions (where they will have to perform tasks for a specific time) and 20 in the ‘ open’ session (free play).

Experimenters/Interviewers

Platform tests will be prepared and organized by experimenters in each user country (The Netherlands, Belgium and Switzerland).

Vilans will provide technology support online (e.g., Skype, email, etc.). It is important for experimenters to gather knowledge about the subject-pool (e.g., what experiences do they share, are they family/friends, i.e. anything unique about the group that participates).

Ideally, two experimenters should run the evaluations (yet, one could also perform the studies). The experimenter(s) should focus on the instructions, questionnaires and tasks and the equipment. In addition, for the experimental trials (Alpha version testing [DATE] and Beta version in [DATE]) the experimenters will fill in a structured report for the evaluation of the experiments.

Platform Play & Experimental Tasks

Experimental sessions

The participants for the experimental tasks will perform platform-play tasks (to be defined when platform is ready) with the purpose of testing the usability of the interface, the player experience, user acceptance, and learning. See appendix X for details on the protocol [To be developed when platform is ready]. Tasks will be given on separate sheets of paper. Experimenters will be present during these trials and note down relevant quotes (including participant’s unique number).

“Open” sessions: These participants will play the platform freely, at their choice, probably in several sessions.

Location & Equipment

Location: The user tests will take place in a specific room dedicated for POSTHCARD in the designated institutions in the Netherlands, Switzerland, and Belgium [exact locations to be defined].

Equipment: A laptop/desktop, a mouse. Wi-Fi / Internet connection. At least one audio-recorder should be used for the interviews and focus groups. A digital camera can be used to take high-resolution images of the users and user-test setting.

Measurements

The above mentioned research questions will be answered by collecting data with different measurement instruments and with different data collection modes. We will measure the following:

Learning

In-platform scores and level advancements. Furthermore, qualitative data on perceived learning / knowledge acquisition.

Player experience, measured by the *Platform Experience Questionnaire (GEQ).*

The *player experience* of the users will be measured by the Game Experience Questionnaire (GEQ), which probes the players' feelings and thoughts while playing the game (IJsselsteijn, de Kort & Poels, 2008). The questionnaires assess game experience as scores on seven components (e.g., Immersion, Flow, Competence, Positive and Negative Affect) and should be administrated immediately after a game-session has finished. During the experimental session they will receive it after a predefined play time (e.g., 5 minutes).

User acceptance, perceived usability, preferences, and errors, gathered through interviews and questionnaires, and the diary studies.

An appropriate method to study user acceptance is by means of interviews/focus groups and diary studies. The interview/focus group moderator can elicit in-depth information of the perceptions, opinions, beliefs, and attitudes (POBAs, see, Puchta & Potter, 2004) of users about the system. Diary studies can additionally elicit non-structured information on above issues.

The **accessibility of the UI,** studied by means of the *IBM usability questionnaire,* during the interviews and by diary studies.

The *accessibility* of the UI will be studied during the interviews and by means of the IBM usability questionnaire (Lewis, 1995). The IBM usability questionnaire (Lewis, 1995) will need to be translated whenever it is not available in French or Dutch. The questionnaire contains nineteen usability items that has to be rated on a scale, running from one to seven (an example item: "I feel comfortable using this system"). The interviewees can comment on specific UI flaws, but also on their likes about certain aspects of the functionality of the system. In addition, diaries will also provide information about the accessibility of the POSTHCARD UI for informal caregivers.

Background information on participants (i.e., demographic data), measured by a *general questionnaire* (GQ) and reported by experimenters.

Background information, such as socio-demographic characteristics, cognitive abilities, experience with computer technology will be gathered by a general questionnaire that participants will fill in when giving informed consent.

In addition, the experimenters will gather knowledge about the subject-pool (e.g., what experiences do they share, are they family/friends, i.e. anything unique about the group that participates).

System data, measured by the system of the POSTHCARD.

The following data should be stored by the system: time stamps (beginning, every screen, every challenge, end, every stop and return to the platform, idle time), score, sequences, what was chosen.

Psychological, cognitive consequences of playing the platform

We hope to collect additional information on what are the psychological and cognitive consequences of playing the platform for the informal caregivers (e.g. how their everyday life changes, are they happier, does it affect their burden.). These data will be gathered through diary studies.

The above data will be measured using the following data collection methods:

1. A general questionnaire (GQ) on background information prior to the POSTHCARD (a paper-and-pencil questionnaire) (for all participants).
2. Interviews/focus groups after the POSTHCARD about POBA's (perceptions, opinions, beliefs, and attitudes). Interviews with participants in experimental sessions, focus groups with participants in open sessions).
3. A questionnaire with GEQ and IBM usability questionnaire after the POSTHCARD (for participants in experimental sessions).
4. Diary method during the whole period of playing the platform (for all participants, data recorded by the system during the platform (for all participants), a report by experimenters (for all participants).

For the open and experimental sessions, the general questionnaire will be administered on paper, prior to the platform (when giving the informed consent).

For the experimental sessions, the questionnaire that should be answered after the POSTHCARD will include the GEQ and the IBM usability questionnaire after the UI tasks. The participants will play the POSTHCARD platform for a predefined time (e.g., 5 minutes) during the experimental evaluation. It should be mentioned that the informal caregivers might need some support for filling out the questionnaires. Experimenters should be present and help if needed while answering the questionnaires.

For experimental sessions individual interviews will be administered after the POSTHCARD. For participants in the open play sessions, a group discussion (focus group) will be organized after six weeks of the trial period. A focus group is one of the few options to elicit group talk between people. However, informal caregiver will also be interviewed individually in the experimental sessions to control for a social desirability bias (i.e., to catalyze openness).

For participants (Informal caregivers) and for experimenters all instructions and questions should be written beforehand on printed cards in a form of a script/scenario (see Appendixes 1 and 2, will be added whenever working prototype is ready of platform). During the study, the instructions and questions can be read from the cards or script to make sure that everybody receives the same information.

Knowledge transfer / learning

We will also gather insights via the interviews and focus groups on the perceived learning, which will be strengthened by in-game data on advancements and the consequences for the (in)formal carers on daily life. Additional metrics will be searched for whenever a first working prototype of the platform is ready.

Procedure for the evaluations

An Alpha pilot study with 5 informal caregivers per country will take place in in the labs of a partner in the participating countries or a specific designated room, to gain insights intro improvements of the preliminary platform and game design but also to test the study procedure.

The actual user tests of the Beta version will last for 6 weeks from [DATE]. Of the 25 participants per country that will sign up for the evaluation study, five (selected randomly) will be recruited for the experimental sessions.

Timeline and overview of the procedure

Study	**Alpha**	REDESIGN	**Beta**	1.Experiment	2. Free Play
Month	Dec. 2019		March 2020	March. 2020	March 2020
Pilot phase length	1 week		6 weeks	1 week	6 weeks
Type of test	Semi-structured evaluation			Experiment	Free play
Nr. Pps.	5 per country		25 per country	5 per country	20 per country
Duration per participant	Max. 3 hours, in one session			2 * 2 hours (e.g. over 2 days, 09:00-11:00, 12:00-14:00, 15:00-17:00 hours)	Max. 6 hours, several sessions possible
Method (in the order of implementation	1. GQ prior to the platform 2. Focus group/interview after the platform 3. System data. 4. Experimenter's report.			1. GQ prior to the platform. 2. GEQ, IBM usability questionnaire after a session. 3. Score for experimental tasks. 4. Interview after the platform. 5. System data. 6. Experimenter's report.	1. GQ prior to the platform. 2. Focus group after 6 weeks. 3. Diary. 4. System data. 5. Experimenter's report.
Tasks	Platform play and user interface tasks.			Platform play and user interface tasks	Free play
Measurements	– Background information – POBA's – Perceived usability, platform experience, user needs, and errors.			– Background information – Platform experience by GEQ, perceived usability by IBM questionnaire. – Efficiency (tasks time) & Effectiveness (tasks correct). – POBA's. – Psychological and cognitive consequences. – Learning – in platform scores, perceived learning	– Background information – Perceived usability & player experience. – POBA's. – Psychological and cognitive consequences – Learning – in platform scores, perceived learning

Procedure for the experimental and 'open' sessions

Procedure for the Alpha version testing

The participants will perform a series of predefined tasks (e.g., play level, change name etc.), although semi structured and open for free play. It will be most importantly to gather an understanding about the usability and player experience before the whole game is played. At this stage it is still relatively easy to alter the interface and game play using the comments and POBAs of the participants.

Procedure for the Beta version testing

At the beginning each participant will be assigned a unique participant's code that will be used to identify him/her. This code should be used/incorporated in all data collection methods and in the sessions played. In this way we will assure that data from different sources can be merged and analysed properly. The participants for the 'open sessions' will be welcomed and a plenary introduction (+/- 60 minutes) will be given by one of the experimenters to introduce the POSTHCARD platform under study (UI, controls etc.) and to provide information about the study procedure. After this, participants will be asked to read and sign an informed consent form. If necessary, the participants can be supported by the experimenters in filling in the forms. The informed consent will state what is being studied, ensures anonymous analysis, announces that audio and image recordings are going to be made, and make clear to the participants that they could withdraw their consent and cooperation at any given point in time during or after the study. Then, they will fill in the general questionnaire (GQ) (paper-and-pencil questionnaire), if they have not already done this.

In the experimental group the procedure will be similar, yet the participants will have to perform game-play tasks (see Appendix 7, to be added later when platform is ready) with the purpose of testing the usability of the interface of both the platform and the game, player experience, user acceptance, and learning. In the free play group participants will be able to play the game freely. They will be instructed to try to finish the levels of the game. They could play the game in several sessions and try to improve the scores of a previous session. In the experimental session participants will fill in the GEQ and the IBM usability questions.

The participants will be provided with the possibility to have a small break, some chitchat, and a visit to the bathroom.

The score of the participants will be recorded by the platform, just as the time of play, pauses, etc.

The experimenters will interview/involve all participants in the focus groups. For the 'open' sessions participants will participate in a large group informal session (focus group) to discuss the game play and usability of POSTHCARD. For the experimental group, the interview will take place with each participant after he/she finishes the platform. For the free play group the focus group will take place once after the 6 week pilot phase. An interview or a focus group will provide users with means to 'step outside the box' of pre-construed questionnaires. The focus during the interview/focus group will be on preferences and user acceptance of the POSTHCARD. The whole focus group will take about max. 1.5 hours for the open sessions and the interview will take approx. 15 minutes for the experimental session. (see Appendix 8 for a preliminary version of the interview questions, to be added later when platform is finished).

In the open session, participants will be also encouraged to write a diary (semi-structured) during the pilot phase period.

Each participant will receive an information sheet with the descriptions of the whole procedure (time line and short description of all the tasks) (see Appendix 2, to be added later when platform is finished).

At the end of all sessions there will be a small gift for each participant.

Data & Analysis

All data from different sources should be merged (using a unique participant's code) in order to make proper analysis and interpretation. These **data** will include:

A report from experimenters; a SPSS document or MS Excel document in which the data of the participants can be filled in. Experimenters will also need to fill in data from the GQ and their observations.

A report from experimenters on the analysis of interviews/focus groups and diaries. The interviews and diaries should rather be analyzed and translated by the end-user sites due to language barriers. Relevant reports about the usability, usefulness, acceptance of the POSTHCARD platform, can be summarized and translated to English. Relevant and remarkable citations/comments can be gathered under a theme. Relevant reports can be noted down as: "I liked the interface, especially the avatars were very nice, and they really made me feel good" (Participant 3, 76 years old, female).

A system report: a report prepared automatically by the system. For each session a unique participant's code and system data should be reported,

Data from the paper/pencil questionnaires after the POSTHCARD (GEQ, IBM usability questionnaire).

The image/video recordings.

The user partners will combine and analyze the data and report the results.

Documents/Materials to be prepared

The following materials should be prepared for the pilot test and are included in the appendix.

- A script – scenario with the procedure for experimenters.
- A script – scenario with the procedure for participants.
    - 3 versions (Alpha version testing, Beta experimental test, Beta free play test).
- Instructions for the ACT for participants – User manual.
- An excel spreadsheet for each country to fill in the data on participants by experimenters.
- An informed consent form
- A general questionnaire (GQ) – self-administered paper & pencil form.
- Description of tasks for the experiment.
- Procedure, questions for the interviews/focus groups. 2 versions (experimental test, free play).
- GEQ for experimental sessions (paper&pencil form).
- IBM usability questionnaire for experimental sessions (paper&pencil form).
- A form of a report for data from the system.
- An excel from to enter data from GQ, GEQ, and IBM usability questionnaires
- A diary ‘questionnaire’.

References

Charness, N., & Holley, P. (2004). The new media and older adults: Usable and useful. *American Behavioral Scientist, 48*, 416-433.

De Schutter B, Abeele V Vanden (2008). Meaningful Play in Elderly Life. *Proceedings of ICA 2008, Communication for social impact*. Montreal: Quebec

De Schutter, B., Brown, J.A., & Nap, H.H. (2015). Digital games in the lives of older adults. In D. Prendergast & C. Garattini (Eds.), Ageing and the Digital Life Course. Brooklyn, NY: Berghahn Books.

De Schutter, B., Gerling, K., Brown, J., Collins McLaughlin, C., Mosberg Iversen, S., Allaire, J., Rice, M., & Nap, H.H. (2015). Gerontoludic for fun!. Gerontechnology Vol 13, No 4.

De Schutter, B., Eynon Black, D., & Nap, H.H. (2015). Teaching Older Adults to Play Minecraft (p. 4). Presented at the Foundations of Digital Games, Pacific Grove, CA.IJsselsteijn, W.A., de Kort, Y.A.W. & Poels, K. (2008). The Platform Experience Questionnaire: Development of a self-report measure to assess the psychological impact of digital platforms.

IJsselsteijn, W.A., Nap, H.H., de Kort, Y.A.W., and Poels, K. (2007). Digital Platform Design for Elderly Users, Proceedings of Futureplay 2007 (Toronto, Canada, 14-18 November 2007), pp. 17-22.

ISO 9241-11 (1998). Ergonomic requirements for office work with visual display terminals (VDTs)Part 11: Guidance on usability. Geneva: International Organization for Standardization.

Lewis, J.R. (1995). IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *International Journal of Human-Computer Interaction, 7*, 57-78.

Morrell, R.W. (Ed.). (2002). *Older Adults, Health Information and the World Wide Web*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Nap, H.H., & IJsselsteijn, W.A. (2016). Design for senior gamers. Gerontechnology Vol 15, Supplement.

Nap, H.H., IJsselsteijn, W.A., & de Kort, Y.A.W. (2009). Age Differences in Associations with Digital Gaming. *DiGRA 2009 Proceedings* (London, 1-4 September 2009).

Nap, H.H., de Kort, Y.A.W., & IJsselsteijn, W.A. (2009). Senior Platformrs: Preferences, Motivations & Needs. *Gerontechnology, 8(4)*, 247-262.

Puchta, C., & Potter, J. (2004). Focus Group Practice. London: SAGE Publications.

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction* (3rd ed.) . Reading, MA: Addison-Wesley Publishing.

Xie, B. (2002). Older adults, computers and the Internet: Future directions. *Gerontechnology, 2*, 289-305.

Appendix

Script – scenario with procedure for experimenters

Script – scenario with procedure for participants

Platform instructions – Users Manual

Excel spreadsheet to fill in data by experimenters

Excel file named “Document 4”.

Informed consent form

2 Word files named “Document 5”: one for open sessions and one for experimental sessions.

General questionnaire

Word file named “Document 6”. In includes:

General questions,

Health and perceived control

Interaction with technology

Experience with computer platforms

Tasks for experimental sessions

Interview/focus group questions

Word file named “Document 8”.

Platform experience questionnaire

Word file named “Document 9”.

Ibm usability questionnaire

Word file named “Document 10”.

For with system data

Excel file named “Document 11”.

Form to enter data from questionnaires

Excel file named “Document 12”. It includes:

Instructions for evaluators,

Form to fill in data from GQ,

Form to fill in data from GEQ & IBM usability questionnaire.

Diary questionnaire

Word file named “Document 13”.