ICT-2009.7.1

KSERA Project

2010-248085

Deliverable D5.5

End Evaluation

30 November 2012
Public Document

Project acronym:	KSERA
Project full title:	Knowledgeable SErvice Robots for Aging
Work package:	5
Document number:	D5.5
Document title:	End Evaluation
Version:	1

Delivery date:	30 November 2012 (month 34)
Actual publication date:	30 November 2012
Dissemination level:	Public
Nature:	Report

Editor(s) / lead beneficiary:

Author(s): John Lemberger (MAC)

Johannes Oberzaucher (CRT)

Franz Werner (CRT)

Hadas Lewy (MAC)

Katharina Werner (CRT)

Reviewer(s): Wenjie Yan (UHH)

Glossary of terms used

Acronym/terms	Definition
ADL	Activities of Daily Living
AR	Assistive Robotics
COPD	Chronic Obstructive Pulmonary Disease
CRF	Case report form
ER	Emergency Room
FEV1	Forced expiratory volume in 1 second
FVC	Forced expiratory vital capacity
FMECA	Failure Modes Evaluation and Effects and Critically Analysis
GOLD	The Global Initiative for Chronic Obstructive Lung Disease
HCI	Human Computer Interaction
HRI	Human Robot Interaction
iADL	Instrumental Activities of Daily Living
ICF	International Classification of Functioning
KSERA	Knowledgeable Service Robots for Aging
KSERA-system	A smart home environment and a mobile platform, i.e., Nao
KSERA ubiquitous	A smart home environment
Nao	Humanoid Robot
UCD	User-Centered Design
SAR	Service Assistive Robot
Scenario	Usage narrative
Test case	Scenario based sequence for evaluation purposes
TOT	Task oriented training
Use case	Describe the functional requirements
WHO-ICF	See International Classification of Functioning
WP	Work package
User	Primary User of KSERA system: (a) persons with COPD and (b)
PT1	Prototype 1
PT2	Prototype 2
TP	Trial participant
TC	Testcase

Content

Executive summary

This document describes the End evaluation trials (procedure, results, findings and discussion) within the project KSERA. The trials were conducted in real life conditions in senior citizen homes in Austria and Israel together with a number (n=16) of older (age = 70-95) trial participants.

The test system (KSERA prototype 2) consisted of a humanoid Social Assistive Robot (SAR), together with a ubiquitous sensor system and a smart home solution (see KSERA deliverable D5.4). The trials were performed in several iterations and were observed by members of the KSERA research team.

During the trials a strong focus on the evaluation of the user experience and satisfaction, the general KSERA approach, acceptance factors related to the future intended use in homes of the KSERA users and the technical performance was set. Besides that also usability and Human-Robot-Interaction factors were evaluated.

The results verify the KSERA approach and offer the opportunity for enhancements and further design decisions for related projects. These results and lessons learned are condensed in conclusion tables at the end of the evaluation section in chapter 4.

Purpose of this deliverable

The deliverable D5.5 **End Evaluation of a SAR in a real environment** describes the end evaluation process, methods and results of the KSERA end evaluation system trials.

Suggested Readers

This deliverable is intended for the general public. It can be read by those willing to know more about the use of robotics systems and their applications in Ambient Assisted Living and shall give a deep insight of the methods and metrics used during the end evaluation of the KSERA prototype and the results gathered.

Relationship to other documents

The trial description is strongly connected to the trial plan document D5.1 and the second prototype D5.4. The metrics used during the trials are taken from the public deliverable D1.3 “SAR metrics” where they are explained in further detail.

1. Methodology

1.1. Main target of the end evaluation process

The KSERA evaluation approach addresses some crucial evaluation domains. According to the UCD the evaluation activities are iterated along the process. Technical validation tests and cooperative evaluations are included. In particular, the cooperative evaluation, involving “real” end users, is used in order to collect, study and exploit the user’s perspective on different aspects of the system, such as **system performance, human-robot interaction (HRI), usability and user acceptance.**

- The **system performance** and **HRI** measures are needed in order to collect information and feedback useful to the system design. From the users’ perspective the system performance is validated in terms of reliability, efficacy and safety – essential requirements to accept and use an innovative system.

- The **usability** measures are needed in order to guarantee the ease of use and a ”friendly” interaction of the users with the system. It results from objective measures and subjective evaluations of the efficacy and effectiveness of the system. The third important element that completes the usability evaluation is the user satisfaction (Standard ISO 9261). According to Norman (2004), people get an **emotional value** from the interaction, consisting of affective benefits such as relaxation, pleasure or fun. The evaluation can go deep into different levels of the gratification: visceral, related to the aesthetics and the external image of the system; behavioral, resulting from the correct function; and reflective, connected with experience, memories feelings.

- The **user acceptance** is a multidimensional indicator, referring to the willingness to use the system. According to the TAM – Technology Acceptance Model (Davis, 1989), some predictive measures can be captured from target users, when asked to evaluate (1) the **perceived usefulness**, in terms of practical consequences of using the system; (2) the **perceived ease of use** based on direct experience of using the system; and (3) **social factors**, connected to the individual’s internalization of the reference group’s subjective culture (Thompson, 1991).

Beside those dimensions one of the main aims of the KSERA project is the possibility to improve the quality of life of older persons in general and in particular those with COPD. The term ***quality of life*** (QoL) is used to describe the general well-being of individuals. The term is used in a wide range of contexts and mostly is adapted towards concrete research topics and related questions. Standard indicators of the quality of life include physical and mental health, but also aspects of safety, privacy and self-determination, also included in the KSERA evaluation plan.

- In the KSERA evaluation we try to answer the research question concerning an increase of the QoL by taking into account one of the main research assumptions of the KSERA project. This assumption states that a system that positively addresses several primary user needs and is accepted by the users will positively influence the **QoL** of the primary users.

These evaluation domains are related to each other and recur in the KSERA project. The evaluation strategy concerns the whole KSERA system, but particular attention is paid to the SAR, which is the main user interface that the KSERA system offers to the users. The KSERA system has been evaluated with the users in different ways, according to the level of maturity of the system (system image, low fidelity prototypes, prototype1, prototype2). Scientific resources show that the user value becomes evident while presenting the system’s attributes and mastering the user’s ability to try and use it. The user’s value assignment in the KSERA project will be observed as a cognitive process related with psychographic and contextual factors. Studying the user interaction in both **scenarios**

related to usage in daily life and in experiments and trials in real context for a longer period of time, an understanding of the desirable consequences will emerge.

1.2. KSERA evaluation aims

As a base for a holistic evaluation of the system regarding several high level hypotheses and research questions, the system initially has to be evaluated concerning its **performance and HRI issues.** This evaluation has been mainly performed during the KSERA formative evaluation and guided the development and research in the second design stage. Applicative scenarios and functionalities were investigated and studied, also in terms of user acceptance and individual impact on the end evaluation.
Concerning the evaluation of the QoL during the two user evaluation phases, our approach was to identify the users' needs related to their quality of life to put them in relation with the abilities of support and assistance given by the KSERA system. Based on this, we identified the user need domains:

- **Physical measurement and health support abilities**
- **Exercise and Training support abilities**
- **Environmental monitoring**
- **User experience and psychological impact**
- **Social inclusion**
- **Information and decision support**

These domains are the most promising features of the KSERA system for improving the QoL of people with COPD and of older people in general. These domains are also strongly connected to the scenarios described in KSERA deliverable D1.1. Based on these domains measurements concerning the question of the perceived usefulness of the KSERA approach regarding to the **Acceptance Factors** of the KSERA system and **a possible improvement in the quality of life (QoL)** of the later users were covered. KSERA deliverable D1.3 gives an overview of the evaluation domains and related indicators that were evaluated throughout the evaluation with real users.

Primary Users

The KSERA Evaluation Model for primary users is focused on the research question **how users accept and use the KSERA approach**. Beside the technical performance (but strongly related to it) the model suggests that when the trial participants are presented with this new kind of technology, a number of factors influence their decision about how and when they would use and accept the system in a future use.

To evaluate the KSERA approach on a holistic view, we evaluated it (**additionally to the system performance evaluation**) concerning the **SAR approach, acceptance and psychological impact on primary end users.**

Secondary Users

The KSERA Evaluation Model for secondary users is also focused on the research question **how users accept and use the system**. The main aim was to evaluate the user experience of several groups of secondary end users with the KSERA application (e.g. a doctor can see the physical measurement values and diagnose, an informal carer can be informed about the current primary user status, a member of a call center can perform characteristic tasks using the KSERA application, etc). The secondary end users (formal and informal carers, medical doctors, therapists, call center agents, etc.) have also been asked to evaluate the system concerning several explorative hypotheses, investigating how KSERA, as a working, flexible and adaptive system could also meet specific needs of the secondary users. The evaluation activities involving secondary users focused on the usage scenarios. The **system performance** has been evaluated in relation with the specific use-cases where the secondary user is directly involved. During the end evaluation process secondary users have been invited to evaluate the high fidelity prototype (PT2), in order to assess topics like the **SAR approach in general, psychological impact and** effects on their own and the expected acceptance and applicability of primary users.

1.3. What is the methodological base for the evaluation?

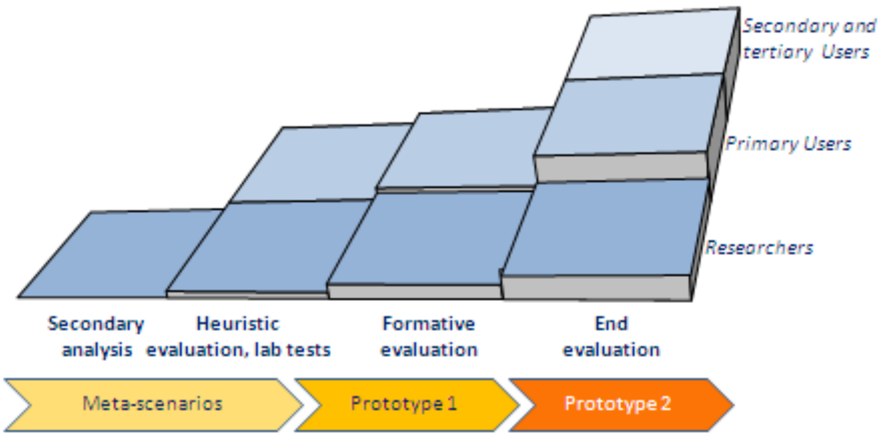
In order to take into account the above mentioned objectives we defined a set of metrics described in KSERA deliverable D1.3. These measures address the needs of all future involved user groups in order to guarantee a holistic view and results.

Beside a strong focus on the primary users and their needs, also secondary users were included in the evaluation phases. This offers a holistic picture of the needs, strengths and problems of the KSERA approach. Furthermore it provides evidence of the impact and the added value that the selected KSERA approach is offering and important information for related projects.

1.4. Overall strategy

As described before all relevant needs of **primary, secondary and tertiary** users were addressed in the KSERA evaluation and therefore also in the KSERA evaluation strategy and model. This evaluation model uses a mix of different evaluation methods and metrics to handle the different **evaluation aims.** These **aims** are represented by different macro-domains (**e.g. system performance, usability, acceptance factors, user experience, societal and psychological impact)** concerning the different user groups (see also KSERA deliverable D1.3).

As a base for a holistic evaluation of the system regarding several high level hypotheses and research questions, the system has been evaluated according to the UCD in an iterative process, where each phase provides feedback and hints for the following ones.



The KSERA evaluation strategy is based on some basic components:

- **PROTOTYPE STAGES:** evaluation activities accompanied the system development steps. At the beginning the paper mock-up and meta-scenarios have been used to collect requirements and evaluate high level interaction strategies. Then prototypes of increasing fidelity allowed to refine the system (from PT1 to PT2[1]).

[1] For the prototypes description see: KSERA Deliverable D5.2 and D5.4]

- **EVALUATION METHODS:** depending on the maturity of the system, the evaluation methods and techniques range from secondary analysis (continuous updating of the state of the art) to heuristic and laboratory validation, up to cooperative evaluations in real life context. In each evaluation phase, technological and user centred topics have been addressed.
- **USERS INVOLVEMENT:** the cooperation among researchers (engineers, computer scientists, physicians and human factors experts) and users (primary, secondary and tertiary) is the core of the evaluation. Among the users, the primary ones represent the most important target, so they have been the most involved. According to one of the KSERA hypotheses (saying that the system acceptance by the primary users influences the attitude of other users) secondary users have been involved in the End Evaluation as well. While different parts and functionalities of the system have been evaluated iteratively by the developing partners in lab trials with students and co-workers, an overall evaluation of the integrated KSERA system has taken place in real life conditions with real primary users.

1.5. Evaluation key methods

Research phase	Methods and techniques	Studied dimensions
Secondary analysis	– **Scientific literature**	– SAR State of the art
Heuristic evaluations, lab tests (low-fidelity prototype)	– **Lab tests** – **Usability checklist** – **FMECA**	– System performance – HRI and Usability – System safety
Formative evaluation (PT1)	– **Usability test** (task scenarios, questionnaires, thinking aloud), objective and subjective metrics – **Short-term trial with a sample of primary users** (in-/direct observation, questionnaires, on-site interviews), objective and subjective metrics	– System performance – User interaction and system performance – Acceptance and user value
End evaluation (PT2)	– **Short-term trial with a sample of primary users** (in-/direct observation, questionnaires, on-site interviews), objective and subjective metrics – **Long-term trial with a primary users panel** (in-/direct observation, questionnaires, on-site interviews), objective and subjective metrics – Interviews, Focus groups **with a sample of secondary and tertiary users**	– User interaction and system performance – Acceptance and user value (long term impact) – Service model, acceptance, long-term maintenance

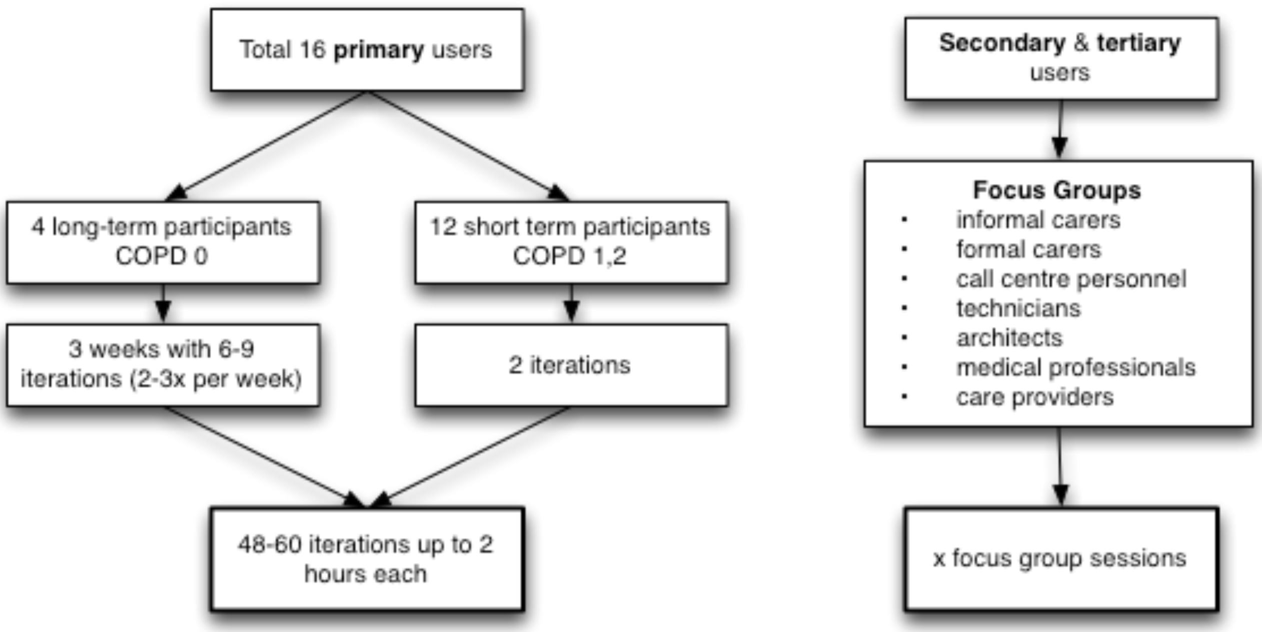
A detailed description of the methods and study designs used to study the different dimensions of the KSERA objectives can be found in the reworked deliverable D1.3 in chapter 2. In the document's annex we provide a table describing in more detail which methods have been used to evaluate the different evaluation and indicator domains.

2. Final KSERA end evaluation test setting

2.1. Description of Test Users

The criteria to take part in the end evaluation have been being at least 70 years old and cognitively healthy. Furthermore the physical abilities should permit to conduct the physical exercises. The users had to agree with the audio and video recording of the trials and signed an informed consent document.

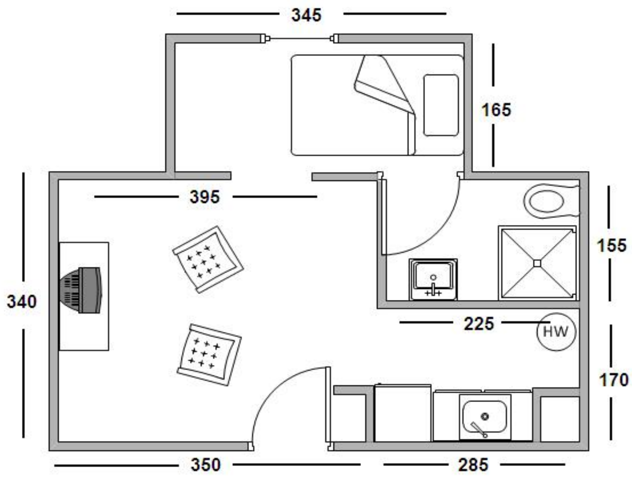
16 persons with an average age of 77 years (70 – 95) took part in the end evaluation.



2.2. Test setting of the field trials

The trials took place at one site each in Austria (Raltec) and Israel (Maccabi). Both sites used the same test methodology and test environment (with some local amendments) in technical terms. The rooms were equipped with different furniture to evaluate site dependent characteristics of the project. During the end evaluation the tests were conducted in a room mimicking a room of a real-life user apartment. This room was equipped with the KSERA prototype 2 system to become the test environment.

At the Raltec trial site the room of the „senior citizen centre Schwechat", that was also used during the formative evaluation, was equipped as a mimic environment for the tests. This mimic environment offers the opportunity to simulate a typical day of usage of the KSERA system under controlled technical conditions. A detailed description of the basic physical and environmental parameters of the test room can be found in the test plan document within the ANNEX of the



updated D5.1, July 2011.

Figure 1 Top view of the test environment at the senior citizen centre in Tel Aviv (bottom picture) and Schwechat (top picture). KSERA system components and test components are shown.

shows a topographic view of the adapted test environment for the end evaluation. The test environment is based on the environment used during the formative evaluation but amended to accommodate new functionalities of PT2 as well as experiences gathered during the formative evaluation.

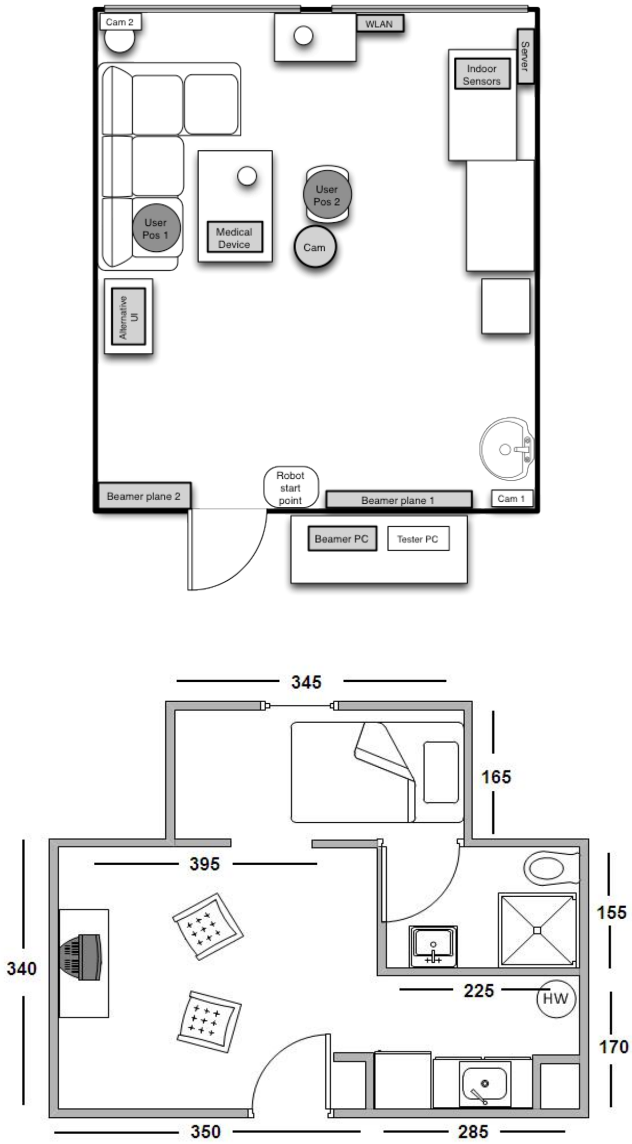


Figure 1 Top view of the test environment at the senior citizen centre in Tel Aviv (bottom picture) and Schwechat (top picture). KSERA system components and test components are shown.

The test environments were equipped with following components:

- Medical device – a pulseoximeter „Nonin 9560“ which was used as an example of a medical measurement device
- Cam – a ceiling cam used for localization of the robot and the user and for navigation of the robot.
- Indoor sensors – indoor environmental sensors for temperature and humidity used in testcase 3
- Server – the KSERA server running the KSERA system
- WLAN – a router hosting the wireless network used to connect the wireless robot
- Beamer PC – a dedicated pc to relay the videos to the wireless video projector mounted on the robot
- Beamer plane - spaces that were kept free of furniture to allow the robot to find a free spot on a white wall for projecting videos of exercises during testcase 2

Two user positions shown mark the usual positions of the users during the trials. The users could decide themselves where they would prefer to sit.

The robot started all test cases at the spot marked “robot starting point” aside of the entrance door.

Figure 2 shows the trial site in the Living Lab Schwechat in the senior citizen home Schwechat. The picture gives an impression of the real life condition (concerning furniture, arrangement, light and sound conditions.)

Figure 2: KSERA Trial setup (Senior citizen center Schwechat)

During the test three members of the research team supported the test flow. Two researches (with inside knowledge of the technical system) were located outside the room during the test to steer the startup and stop of the testcases, control the audio and video recordings and to fill out technical performance questionnaires. The third researcher was concerned with the user involvement and stayed with the user in the room in order to be able to support the test user during the first iteration. In the second iteration only the user and the robot stayed in the room to support the impression of the future scenario of being at home with the robot.

To control the test flow the technical researchers used a technical user interface as shown in the figure below.

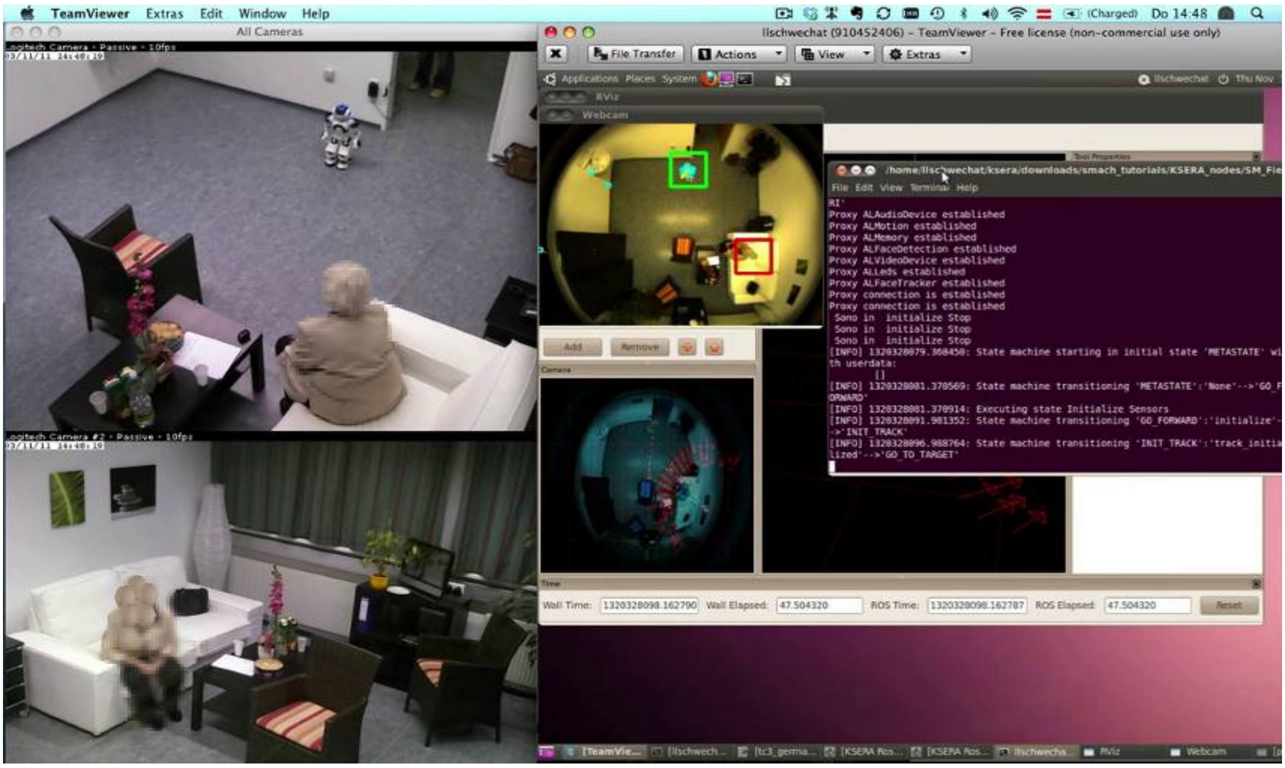


Figure 3: researcher user interface

The researcher interface gives the technical researchers the possibility to observe the scene using 3 cameras mounted on two corners and the ceiling of the room. A connection to the KSERA server enables the direct manipulation of the test flow by using shell commands. This researcher interface was screen captured during all test cases for retrospective analysis. This data and video material was also shared with the other partners in the KSERA consortium in order to realize an iterative improvement to technical problems.

The following figure shows the data flow of the KSERA system and highlights the functionality of the researcher interface.

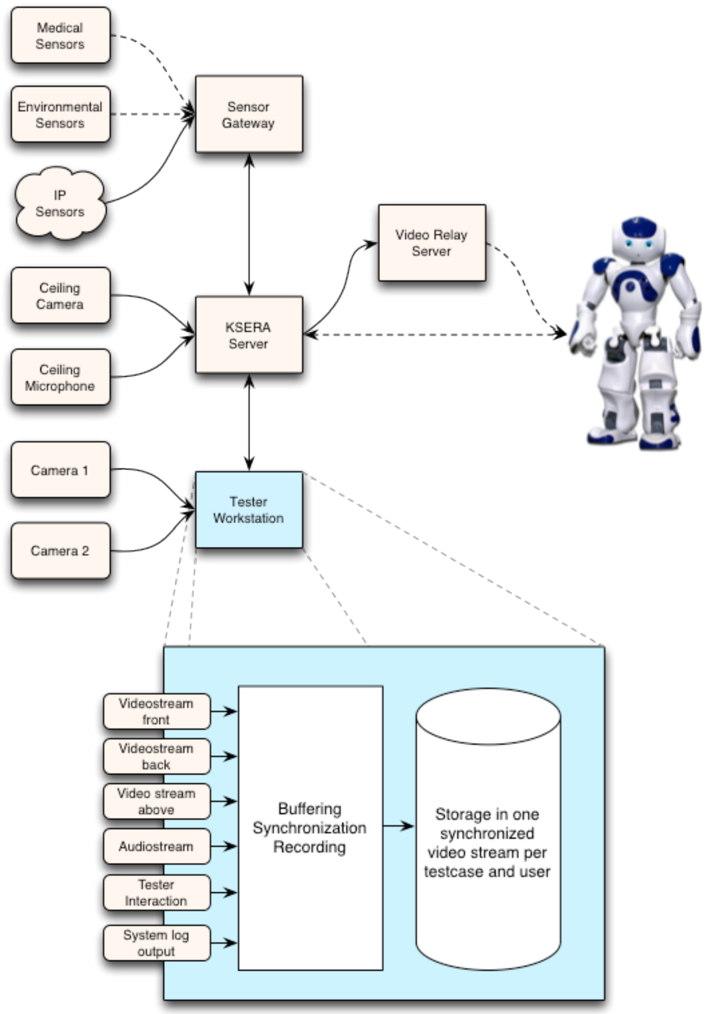


Figure 4: Data flow and synchronization of recordings within the KSERA system

All available video / audio streams, the system logging output and the researcher interaction during the test cases was time synchronized and recorded for later analysis of the trials in terms of human robot interaction and technical performance.

2.3. Test procedure during the field trials

Trial participants were invited separately into the test environment.

Three members of the research team have been present to conduct the trials, two technical experimenters and one usability experimenter.

The technical experimenters:

- controlled the system via the tester interface
- controlled the speech recognition
- started and stopped recordings, took screenshots if necessary (e.g. using kinovea)
- filled out the reliability evaluation form

The usability experimenter observed the scene directly (e.g. through an open door) but avoided distracting the system in any way (is outside the test area).

- Guided the user through the story around the scenarios (see 0)
- Noted general comments of the test user
- Helped in case the system malfunctions (e.g. Nao might crash)
- Helped the test subject in case of questions.

Each trial was conducted in the following steps (phases). The average procedure for a single test with one trial participant took about 90 minutes on average – including all of the following test phases.

Preparations prior to the test

A checklist was used to make sure that the technical system is ready for the test.

Different items (like a newspaper or equipment to prepare breakfast) were placed on the couch table to support the feeling of being in an "everyday life-situation" during and in between the test scenarios.

Reception of the user

The project was explained to the participants, his/her role during the tests was discussed and the formal agreement of the participant was given within the informed consent document.

Pre Test Phase - initial questionnaires

To rule out effects of the current mood of the participants they were asked to fill out a PANAS questionnaire [PANAS]. First Time users also filled out the WHOQoL Quesionnaire. [WHOQoL]

Explanation of the KSERA-system

Before starting with the scenarios the robot and test environment were introduced to them once more. Every visible part of the system was explained to the user:

- NAO
- Ceiling camera
- Ceiling microphone
- Research cameras / microphones
- Test setup (e.g. user in the test-environment, researcher outside the environment watching)
- Explanation how to handle the pulseoximeter
- Explanation of the optional additional local user interface (only in second iteration of Austrian trials)

Demonstration of test-cases

Iteration 1

Afterwards the user was told a story to be able to imagine how the robot could be used during a typical day of use at home. Embedded in this user story were the test scenarios that were shown to the user. After the demonstration of all scenarios the test was concluded by filling in the evaluation questionnaires.

The following test cases were conducted inside the user story: (see also annex chapter 6.1 interaction flows for a description of IF1 to IF5)

- **Getting up in the morning and asking Nao about the current weather situation – (IF4) Environmental information / warning**
  The KSERA system informs the user about the current indoor and outdoor environmental parameters, will integrate the data with the physical measurements and give instructions

accordingly. These data and recommendations are updated in the disease management application.

- **Preparing breakfast while listening to music – (IF6) Nao plays music (entertainment)** The SAR localizes the user, places itself close to the user and plays music for entertainment in front of him / her. A call of the doctor (technical experimenter sitting outside the test environment) interrupts the music. **(IF3).**
- **Doing some exercises after breakfast – (IF1) Medical measurements & physical training.** During this test scenario the user asks Nao to do some exercises together. Before starting with the exercises Nao asks the user to perform a medical measurement. The system stores the measured vital data in the disease management application and further acts depending on the vital parameter value. If the values are slightly abnormal (according to the clinical definition) an alert can be sent to a call centre and less strenuous exercises are performed. If measurements indicate immediate intervention, an alert is sent to the call center, which will react accordingly (via a provided phone call interface) - no physical exercise is performed in that case.
- **Receiving an environmental warning (IF5)** After the training finished an environmental warning is displayed by the SAR. The robot stands up, moves to the user, initiates eye contact and gives an alert orally.
- **Calling a friend to make plans for the evening – (IF2) video telephony (entertainment).** Via the socially assistive robot (SAR) a video call (using the LED beamer attached to the back of Nao) to a “simulated” friend has been established.

During the run of the test cases the technical researchers outside of the room filled out the technical reliability questionnaires. Furthermore the „thinking aloud“ approach was used and thoughts of the test user during the performance of the scenarios were noted.

Iteration 2 – free-interaction session

The second iteration of the short term trials had a similar test flow as used during the first iteration except the fact that the user already knew the system, several pre tests were not needed any more and the demonstrations of scenarios were not fully controlled by the experimenters but by the users.

The user received a handout including the information how to trigger interaction with the robot.

The technical experimenters took care of the correct execution of test cases depending on the user interaction (semi wizard-of-oz technique), the usability experimenter was watching the scene from outside the test room as well.

According to the description of TC5 activities for distraction were used.

The user gets the following instructions (text in italic was not shown on the user handout):

- Please start at least 4 functionalities
- It is possible to quit the interaction at any time by saying “Nao stop”
- In the meantime you can read a newspaper or do a short brain game (cross word or using the tablet pc on the table) – *This can be interrupted by the experimenters with an incoming call or an environmental alert*
- The list of possible functionalities and their commands:
    - *Trigger scenario 1,2,3,4 via voice command*
        - “Nao, play music” *- can be interrupted by the experimenters with an incoming call.*
        - “Nao, let’s do exercises”
        - “Nao, call the doctor”
        - “Nao, what is the weather like”
    - *Trigger actuation (only in Austria)*
        - “Nao, turn off / on the light / the fan”

For the free-interaction session the navigation of the robot has been completely undertaken by the experimenter.
After the user stated to be finished using the system (or longest 20 minutes) the usability experimenter entered the room again and the test participant answered the final questionnaires.

Iteration 3-6 – long term trials

The long term trials had a similar test flow as used during the first iteration except the fact that the user already knows the system, several pre tests were not needed any more and only one particular test case (**IF1**) has been performed.

Post test phase - final interviews

Directly after the last test run the final interviews took place regarding the following topics:

- User acceptance – general impression
- User acceptance – personal feeling
- User acceptance – satisfaction evaluation
- User acceptance – physical measurement abilities
- User acceptance – disease management
- KSERA - HRI Scale
- HRI Godspeed
- PANAS

Discharge and planning of future tests

After filling out the case report form (CRF) the user was told about the next steps of the project and further participation was planned.

Changes during the test phase

During the test the system was kept static to ensure comparable results between iterations. For technical performance reasons a Wizard of Oz technique was used in the second iteration of the trials to simulate the navigation of the robot in order to gather less performance influenced analysis results.

2.4. Focus groups together with primary users

General aims and research questions

- Discussions among patients on the experience gained during the test cases
- Stimulating the evaluation of the system with realistic scenarios and personal needs
- Collecting impressions and new requirements to make the system even more close to the users' needs.

Participants and groups

- 2 groups (each of them of minimum 5 persons, maximum 10)
- The groups have been homogeneous, including people with similar characteristics (e.g.
    - Group1: COPD patients and/or other chronic patients
    - Group2: non-patients/elderly with no specific disease)

The participants have been recruited among the field-trialists.

2.5. Focus groups together with secondary users

General aims and research questions

- Demonstrating the system features, by using meta-scenarios and demos
- To facilitate the discussions among secondary users on their relative/patients' needs and characteristic

- Discussing the possible usages of the system by the secondary users, analyzing expected benefits and risks
- Collect impressions and new requirements to make the system and the possible service models even more close to the secondary users' needs.

Participants and groups during the end evaluation

- 3 groups
  - Group 1: 12 health professionals
  - Group 2: 5 participants – representing carers (formal, informal)
  - Group 3: 3 participants – representing therapists and trainers
- The informal carers have been recruited among the informal carers of the field trialists.
- The formal carers have been recruited among the professionals that have been aware or not of the KSERA project.

3. Results – Evaluation Report

3.1. Evaluation Domains

The following chapter presents the results of the end evaluation. We will present general test statistics and the results concerning the different evaluation domains (performance, acceptance, general approach, impact). We will present the results in a holistic way, based on quantitative and qualitative data gathered from different proposed future user groups. A general conclusion is added in chapter 4.

The following abstract gives an overview on the 6 evaluated domains and related research questions.

Performance of the KSERA prototype in a Real Life situation (chapter 3.3)

The evaluation of the performance was based on the main research questions:

- How well does a technically complex, biped robotic solution – e.g. the KSERA robotic platform NAO - perform in real-life conditions, in an environment comparable to the homes of older people?
- To what extend is there an influence of the prototype's technical limitations on the HRI and acceptance performance results?

Evaluation on social abilities and social presence of the KSERA approach, related to the actual intended private future use (chapter 3.4)

On basis of the performance evaluation of the different functionalities related to the KSERA use cases, the end evaluation also aimed towards gathering data concerning the social abilities, social presence and the related **actual intended future use (ITU)** of the proposed approach in the homes of the trial participants. The ITU is defined as "the outspoken intention to use the system over a longer period in time". It was chosen as it is one of the main target parameters in several acceptance models related to usage and acceptance of technology (e.g. TAM, UTAUT, Almere, etc.) and according to these models is based on acceptance parameters like e.g. perceived usefulness (PU), trust (TRUST), perceive ease of use (PEOU), social presence (SP), perceived sociability (PS).

Within the scope of this experiment we wanted to answer the research questions:

- How are the different influencing items concerning acceptance and ITU pronounced, when using the KSERA robot approach with all known functionalities.
- What are the biggest issues for an ITU of the KSERA approach (Use is determined by (/1) intention to use and influenced by (/2) social influence and (/3) facilitating conditions).
- Is the KSERA robotic approach a good way to implement a social robot?
- Is the KSERA robotic approach a good way to implement an assistive robot?

Emotional feeling and psychological impact during free interaction with the KSERA prototype (chapter Fehler! Verweisquelle konnte nicht gefunden werden.)

The introduction of novel robot technologies in homes of vulnerable user groups and older people is changing the practice of connected care and healthcare. During the KSERA end evaluation the aim was to analyze how technological, social and psychological factors affect human robot interaction, including the major factors of usability and acceptance.

During the KSERA end evaluation we gathered video data (objective) in order to analyze emotional feelings and psychological impact factors based on quantitative data. In addition we discussed in inter- and multidisciplinary future user groups what effects could arise. The main research questions were:

- Do the primary users‘ data gathered by subjective metrics correlate with the objective measurements?
- Do the users integrate the robot into their mental model?
- What could be the burdens in a future use?

General impression of the KSERA SAR approach and the developed functionalities (chapter 3.5.1)

In order to get a holistic view of the KSERA approach the impression and ideas of different future user groups was evaluated. The needs, which were defined in WP1 were discussed under the light of the two evaluation phases, the KSERA prototype and future usage scenarios. The main research question can be defined as:

- What is the added value that the KSERA approach offers to the users in comparison with related approaches (e.g. conventional systems) concerning the user need domains?
- What are the limitations (in the KSERA SAR approach) in the different user need domains
- What should be changed, in order to really use the robot at home
- What are the “No-Gos” for a future use?

Advantages and limitations of the approach concerning the proposed usage scenarios (chapter 3.6)

The approach was evaluated regarding its advantages and limitations using a combination of state of the art solutions for ubiquitous sensing with the KSERA robotic solution. The usage scenarios, which were defined in WP1 were discussed under the light of the two evaluation phases, the KSERA prototype and future usage scenarios.

Effect of a long-term interaction on the acceptance parameters and the general impression of the approach (chapter 3.7)

Related to the evaluation of **social abilities and social presence** we briefly discuss the influence of regular interaction with the system on these descriptive parameters. In addition we present results based on the quantitative and qualitative data gathered during the trials, which are related to the general KSERA approach and the effect on a **future long term application** in (vulnerable) users’ homes.

3.2. Test-statistics

In the trial phase together with older users (age 70 to 95) a group of 16 trial participants (n=8 in Israel, n= 8 in Austria) joined the tests. The trial included two iterations with all trial participants at every test site and 4 long-term iterations at every site with n=4 trial participants. Secondary users were included in the trial phase by presentation of the prototype, discussion sessions and focus groups as multi-disciplinary teams in Austria and Israel. As example of the gathered data the test statistic of the tests performed in the living lab Schwechat is presented:

Primary future users:

- 8 trial participants with two iterations in the senior citizen center Schwechat
  - 4 well defined test cases in the first iteration
  - A free interaction session over a period of about 20 minutes per trial participant during the second iteration
- 2 trial participants for long-term trials in the senior citizen center Schwechat
  - 1 test case during each visit of the long-term iterations
- Subjective and objective trial data
  - 8*3*2+4*2= 56 single running test cases = videos of 3 cameras (ceiling, front, back) and research interface
  - Over 600 single evaluated technical subtestcases (metrics)
  - Over 500 minutes of human robot interaction and research video material
  - Questionnaires and qualitative information sheets for every iteration
- Data gathering deploying a holistic approach
  - User discussions
  - Focus groups
  - Intra-city and age group related dissemination and discussion

Secondary future users:

- Multidisciplinary teams (carers, medical and care experts, therapists, trainers) were involved deploying methods of a user centred design approach in iterative sessions
  - Focus groups with n=5 medical and care experts were performed
  - Focus groups with n=3 trainers and therapists were performed
  - Multidisciplinary discussion with technicians, medical experts and therapists were performed.
- Data gathering deploying a holistic approach
  - User discussions
  - Focus groups

3.3. Performance of the KSERA prototype in a Real Life situation

Research Question

The following main research questions led to the evaluation of the technical performance.

- How well does a technically complex biped robotic solution such as the KSERA platform perform under real-life conditions in an environment like the homes of elderly people?
- To what extend is there an influence of the prototype's technical limitations on the HRI and acceptance performance results?

Methodology

The technical performance of the system was validated in laboratory pre trials before the execution of the respective user evaluation phases. Results regarding the technical performance during the trials were generated by analysis of an error protocol, which was filled out during each trial and afterwards using retrospective video analysis. See also the test protocol in the ANNEX (chapter 6).

Results

An evaluation of the technical performance is relevant to the topic of user acceptance of such a system since a complex robotic prototype platform was the target of evaluation and full control of the prototype was not possible for the experimenters given the real-life character of the evaluation. It can be assumed that technical problems have an influence on the acceptance results since failures of the system negatively influence the user experience during interaction with the system.

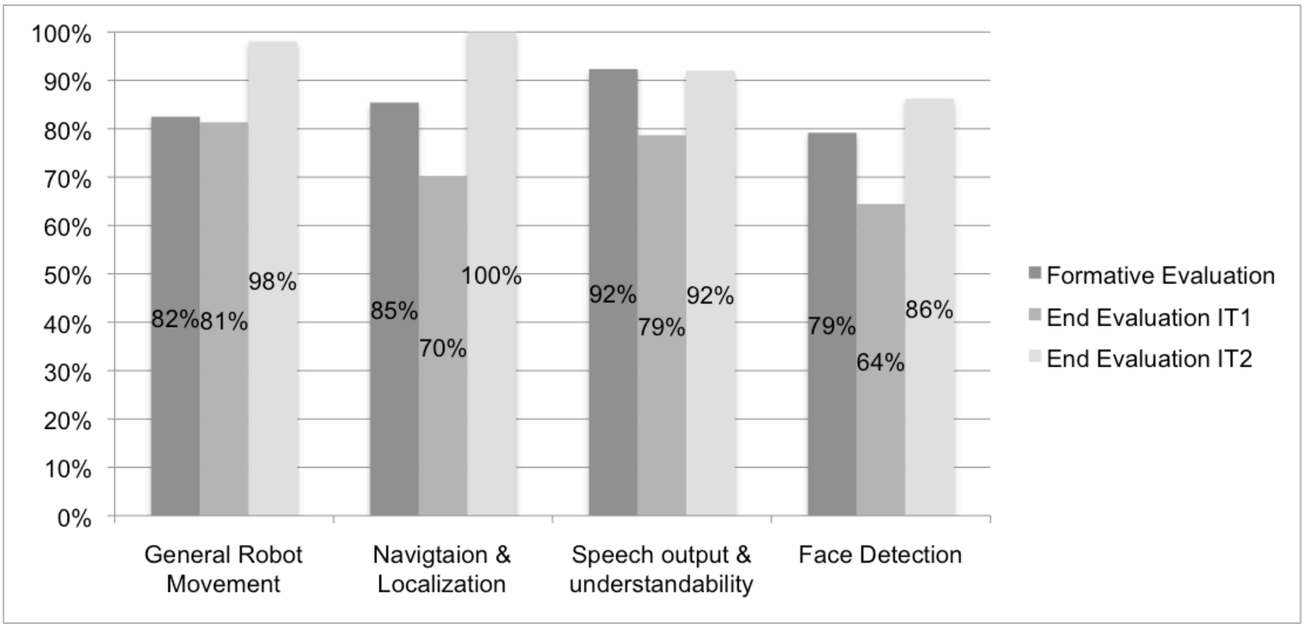


Figure 5 Summary of performance results Formative Evaluation (dark grey) vs. end evaluation iteration 1 (mid grey) and end evaluation iteration 2 (light grey)

Figure 5 shows the summary of the technical performance results for both trial phases. It presents the percentage of correct test cases over all conducted test cases. A test case was defined by one single technical functionality during one of the scenarios. One typical question for one test case would be “Did the robot say the expected sequence?”

Over all technical domains 529 questions of the formative evaluation and 306 questions of the end evaluation were evaluated. For visualization the results were clustered by the following functional domains.

“General Robot Movement” includes all movements by the robot such as walking, gesturing, moving the head towards the user and sitting down / standing up. In most cases these worked well during both phases of the trials. The slightly lower score during the second iteration might be the result of a correlation with the navigation domain, since corner cases of movements occurred when navigation malfunctioned.

"Navigation & Localization" includes user localization, localization of the robot as well as navigation of the robot towards the user and back again. Visual pattern recognition techniques were used for localization in both trial phases. The score of the second phase is significantly lower than of the first trial phase due to a change in software which led to better laboratory results but revealed less robustness under simulated real life conditions.

"Speech output & understandability" includes voice output by the robot and comprehensibility by the user. After the first evaluation phase the volume of the speech output generated was raised which is likely the reason for the slight increase in performance.

"Face detection" was done to target the robot's head towards the user in order to improve the human robot interaction. The performance highly depends on the user's face and whether or not the users wear glasses. Since mostly the same users participated in the trials, the performance results are very similar between them.

Comparison between test sites

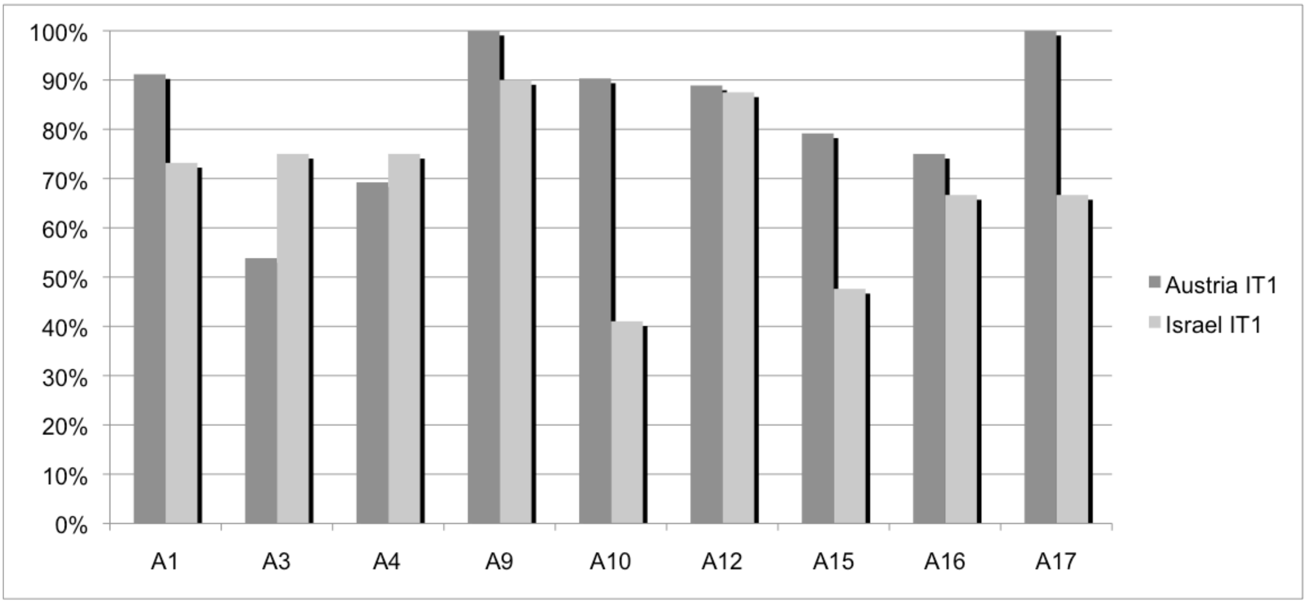


Figure 6 shows a more detailed result as comparison of the technical performance achieved at the two trial sites in Austria and Israel.

A1	Were the robot movements performed correctly
A3	Is the KSERA system calculating the correct user position?
A4	Is the KSERA system calculating the correct robot position?
A9	Is the robot "saying" the expected sequences?
A10	Are the robot's words understandable by the user?
A12	Is the robot motivating the user to perform a task?
A15	Was KSERA able to detect the user's face?
A16	Did the audio connection work as intended?
A17	Did the video connection work as intended?

Table 1 extract of performance metric questionnaire

Most performance parameters do not vary significantly between the trial sites, three parameters show a significant deviation.

Parameter A10 – voice understandability resulted in a lower score at the trial site in Israel. This is most likely due to the fact that the KSERA system was translated to German in Austria and to English (not Hebrew) in Israel since English language is commonly well known in Israel. Still voice understanding seems to be influenced by the use of a foreign language in Israel.

Parameter A15 – face detection scored lower in Israel. This is most likely due to different light conditions at the trial sites which typically result in varying results of image processing algorithms.

Parameter A17 – video connection scored lower in Israel. During the test procedure it showed that the initialization of the video projector needed a significant amount of time and expertise of the experimenters in starting up several processes in the right order. Most likely differences in training of the experimenters resulted in these variance.

Godspeed questionnaire

The Godspeed questionnaire [4] was used to evaluate the influence of technical performance on parameters of HRI and acceptance. The questionnaire was chosen because it is widely adopted and used in the evaluation of robotics and can hence be used to ensure comparability of our results with other projects. The questionnaire uses paired attributes such as “machinelike - humanlike” and a 5-point scale between these attributes to rate the system. Five clusters of attributes are built upon these attributes and are shown below.

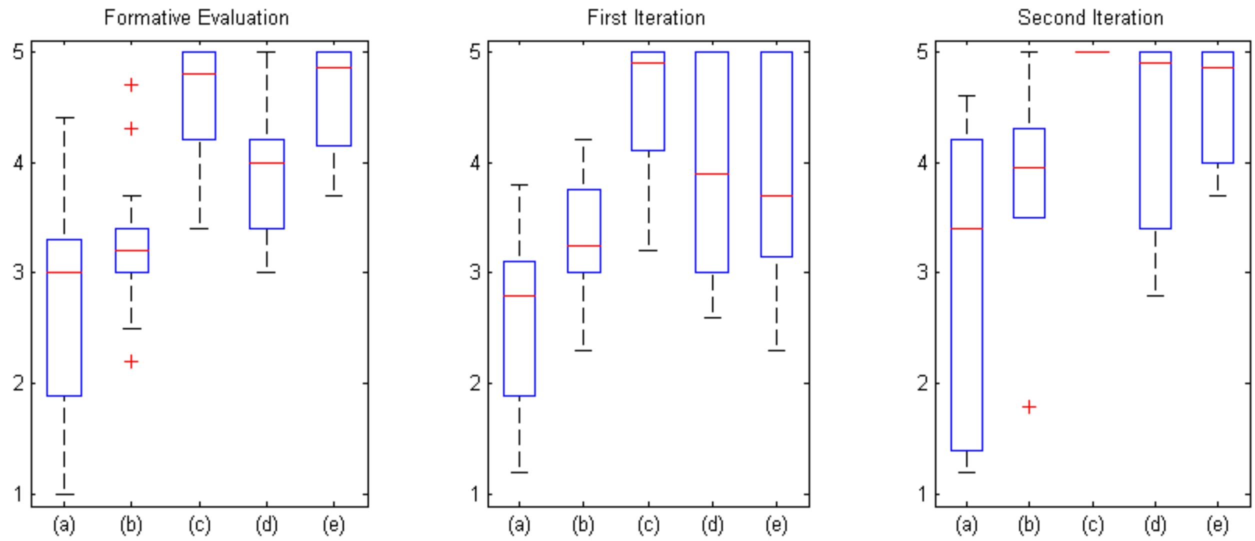


Figure 7 Results of the Godspeed questionnaire formative evaluation (left) and end evaluation (middle & right)

Figure 7 shows the Godspeed clusters for the first and second evaluation phase.

Antropomorphism (a) gives information about the compliance to human like characteristics. This cluster is affected by the visual appearance of the used robot and by the functionality shown within the implemented scenarios by the robot. Similar results were achieved for both evaluation phases and show that the users rank the robot in average between machinelike and humanlike. On the positive side the trial participants noted that the robot looks like a small child and hence quite humanlike, on the negative side the robot is clearly made of (hard) plastic and the movements are rather robot like and not as smooth as those by a human being.

Animacy (b) rates to what extend the robot appears to be a living, organic and interactive individual. The score did not change significantly between the evaluation phases. In general the score was negatively influenced by a lack of interactivity and again by the rather artificial look of the robot. To raise this score, a more interactive communication involving more natural gestures during speech output and non-functional elements such as glimpsing with eyes or tiny movements during standing would likely be beneficial.

Likeability (c) shows to what extend the robot appears to be likable, kind, pleasant and nice. This score achieved high ratings in both evaluation phases. The users were mostly appealed by the nice appearance of the robot, which resembles a small child. Although the small size of the robot prevents provision of physical help the users commented positively on this with regards to the friendly, kind and harmless look that the small size supports [3].

Perceived Intelligence (d) rates the competence, intelligence, responsibility and sensibility of the robot's appearance and actions. This factor shows a high standard deviation due to the fact that some users experienced technical malfunctions of the navigation and others did not. In erroneous cases the robot could not move correctly to the user but walked in circles or quit ending up in the wrong direction. This experience seems likely to influence the perceived intelligence of the robot. In general the users seemed to expect a lot of functionality from the system because the robot has the appearance of a human being. Some users expected that the robot understands their gestures and some talked to the robot fluently although the system could only understand "yes" or "no". This high expectation also might influence the score of perceived intelligence since expectations were not fully met.

Perceived Safety (e) rates the three attribute duplets "anxious – relaxed", "calm – agitated" and "quiescent – surprised". The cluster shows a high deviation which is likely a result of some users' expression that they think "calm" and "quiescent" should rather correlate with "safe" instead of "agitated" and "surprised". Users that felt like this mostly rated these attributes with quite low scores, others with the highest scores.

A 2-sided t-test with a confidence interval of 95% was undertaken to compare the results of the Godspeed questionnaire between the two iterations of the end evaluation. The hypotheses that the results have the same mean could not be thrown for any attribute except one. The duplet "calm - agitated" results significantly lower in the second iteration than in the first. The reason for this fact was found to be that two test participants rated this attribute significantly under average. One of these users experienced a fall of the robot to the ground, which is a possible explanation for the lower score. The other user did not experience other technical malfunctions than the rest of the group. It is hence also likely that the ratings were influenced by the question itself as described above.

The results of the second evaluation phase have generally a higher deviation since the sample size was halved.

Conclusions regarding the question of the general performance of the system in a real-life setting

High efforts were undertaken to make the system more robust after the formative evaluation phase. Given the higher complexity of the second prototype system it becomes clear that the efforts undertaken were successful in most parts since the technical performance did not lower significantly for most domains. Still the goal of higher reliability of the second prototype was not reached. This is to most extend due to general technical issues that are not sufficiently solved so far in research in general, but would be needed for a successful integration of robotic solutions into smart home environments.

General technical issues of robotic prototypes in smart homes

Navigation. This very complex AI issue could only partly be solved. Navigating a biped robot inside an unmodified living environment is a current research challenge that is not sufficiently solved so far in general.

Speech recognition. Current best performing speech recognition engines such as Apple SIRI[2], Google Voice[3] or Dragon speech[4] were not able to sufficiently allow a dialog over a distance of 2 meters with the Nao robot. Since this is a research topic that needs resources far larger that the project KSERA is able to provide, the speech recognition was simulated in all trials using Wizard of Oz techniques.

Autonomy of the robot. The used robot Nao has several issues that influence the autonomy negatively. Firstly the maximum time of operation on battery is only 45 minutes. This was sufficient to conduct the trials, but an automatic recharging mechanism is necessary to keep the robot active over longer periods such as in a field-trial. This maximum time also limits the range of operation since the robot is also limited in speed.

The operation system of the robot is provided by Aldebaran and still under development and hence also prone to errors. Technical supervision is constantly necessary to operate the robot because of possible malfunctions of the operation system.

The used motors and heat dissipation system do not allow the robot to be operated over long time periods since the robot can easily overheat.

Regarding the question of technical influence on the perception of the users

It could be shown that there is an influence of technical malfunction to acceptance parameters, but it is not completely clear to what extend. Additional trials are planned to validate the results in a more controlled test setting for comparison, e.g. using the wizard-of-oz technique to overcome the navigational problems.

---

[2] http://www.apple.com/de/ios/siri/
[3] http://www.google.com/googlevoice/about.html
[4] *www.nuance.com*

3.4. Evaluation on social abilities and social presence of the KSERA approach, related to the actual intended private future use

General description

On basis of the performance evaluation of the different functionalities related to the KSERA use cases, the end evaluation also aimed towards gathering data concerning the social abilities and social presence and the related **actual intended future use (ITU)** of the proposed approach in the homes of the trial participants, whereas the ITU is defined as "the outspoken intention to use the system over a longer period in time"

The ITU is one of the main target parameters in several acceptance models related to usage and acceptance of technology (e.g. TAM, UTAUT, Almere, etc.) and according to these models is based on acceptance parameters like e.g. perceived usefulness (PU), trust (TRUST), perceive ease of use (PEOU), social presence (SP), perceived sociability (PS).

Within the scope of this experiment we wanted to answer the research questions:

- **Q1: How are the different influencing items concerning acceptance and ITU pronounced when using the KSERA robot approach with all known functionalities.**
- **Q2: What are the biggest issues for an ITU of the KSERA approach (Use is determined by (/1) intention to use and influenced by (/2) social influence and (/3) facilitating conditions).**
- **Q3: Is the KSERA robotic approach a good way to implement a social robot?**
- **Q4: Is the KSERA robotic approach a good way to implement an assistive robot?**

Setup

The experiment included all iterations together with primary future users (see description of the test setting in chapter 0). The integrated KSERA system was used with all given functionalities with the robot as main interacting user interface.

Subjects

There were n=16 trial participants involved in the experiment (older people as primary users). Their age ranged from 70 to 95. The participants (n=8) in Austria lived independently in apartments in Schwechat. All participants had COPD Gold Level 0-1. In Israel participants with Gold Level 0-3 were part of the trials.

Procedure

In the first iteration all functionalities and possibilities of the robot were presented to the trial participants by simulating a common sequence of usage scenarios (e.g.: getting up and asking for information, breakfast, the robot remembering of taking a measurement, training after the breakfast, etc.). Before every sequence and interaction a member of the research team explained the following scenario and the usage idea. This method of interactive explanation was used in order to present the different implemented functionalities in an understandable way and to give the older trial participants the chance to implement the approach in their model of daily living. This was done in order to get a common base of understanding of the system approach – which is a base for the evaluation of ITU and related acceptance parameters. In the second iteration the trial participants had the chance to interact with the system without any proposed test flow – all implemented use cases were explained in short again and a paper with speech commands to trigger the different functionalities was handed out to the participants. The only order was to perform at least 3 interactions with the robot. In the long-term iterations a sub-set of trial participants interacted with the robot over a period of one month in iterative sessions (2-3 times a week). Therefore we chose the use case of robot supported physical training. The trial participants were invited to enter the room and interact with the robot.

During the trial sessions the scenes were filmed and observed by members of the research team – the comments given by the trial participants (thinking aloud) were protocolled. In the end of each iteration construct based questionnaires were completed and a general interview about the experiences was performed in order to get a more qualitative feedback.

Methods used

In the scope of this experiment we gathered quantitative and qualitative data using subjective and objective methods.

Quantitative data / subjective data gathering. As main method for gathering quantitative subjective data, construct based questionnaires were used. The instruments developed to measure technology acceptance have each constructs (such as Intention to Use) represented in a questionnaire by a group of questions or statements that can be replied to on a five or seven point Likert type scale. In the described experiment Cronbach's alpha was calculated to test the construct reliability. An alpha of 0.7 and higher is seen as acceptable.

In the following we briefly describe the constructs used for subjective evaluation of acceptance towards the KSERA system and in special the robot during the end evaluation phase of the KSERA system in table 2: overview of constructs used during the evaluation of acceptance.

ANX	Anxiety	Evoking anxious or emotional reactions when it comes to using the system
ATT	Attitude	Positive or negative feelings about the appliance of the technology
FC	Faciliting conditions	Factors in the environment that facilitate use of the system
ITU	Intention to use	The intention to use the system over a longer period in time
PAD	Perceived adaptability	The perceived ability of the system to adapt to the needs of the user
PENJ	Perceived enjoyment	Feelings of joy/pleasure associated with the use of the system
PEOU	Perceived ease of use	The degree to which one believes that using the system would be free of effort
PS	Perceived sociability	The perceived ability of the system to perform sociable behavior
PU	Perceived usefulnes	The degree to which a person believes that the system would be assistive
SI	Social influence	The persons perception that people who are important to him think he should or should not use the system
SP	Social presence	The experience of sensing a social entity when interacting with the system
TRUST	Trust	The belief that the system performs with personal integrity and reliability
USE	Use/Usage	The actual use of the system over a longer period in time
LOS	Level of satisfaction	The level of satisfaction after the trial
GI	general impression	General impression of the system and approach

table 2: overview of constructs used during the evaluation of acceptance

UTAUT and TAM added constructs relevant for studies of SAR in homes of older people

The mentioned parameters are related to assumptions that base on the Technology Acceptance Model (TAM) [introduced by Davis 1989]. It is a common tool for measuring and estimating the intended future use and the related usage behavior. It bases on the measurement of perceived usefulness (PU) and perceived ease of use (PEOU). Figure 8 gives an overview of the construct interrelations. The constructs are presented in questionnaire questions, which are adapted towards the diverse research goals.

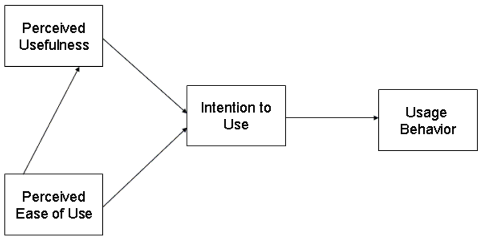


Figure 8 TAM model – Davis 1989

TAM2 and the UTAUT modell (Venkatesh 2000 and 2003) are a consistent enhancement of the TAM model and introduce also social influencing factors and moderating factors like age, gender and experience. The inter-dependicies are shown in Figure 9. The TAM as well as the UTAUT modell have been used in related projects (e.g. Ruyter et al, Heerink2010) for evaluating the acceptance of robotic solutions. Results showed that extrovert robots and robots perceived as intelligent result in a higher intention to use the solution.

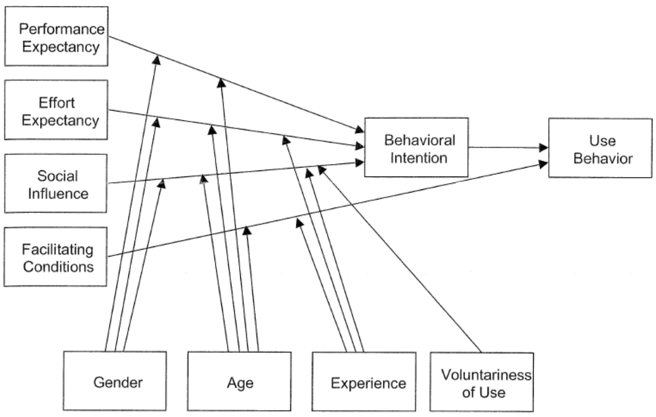


Figure 9 UTAUT model

Another related model, which was developed on basis of TAM an UTAUT is the so called ALMERE model by Heerink (Heerink2010). Its purpose is to evaluate the acceptance of a robotic solution together with older people. It combines several constructs based on TAM, TAM2 and the UTAUT model and adds relevant constructs related to the needs of the target group of older adults and vulnerable users. The evaluation concerning several acceptance relevant factors and an intended future use during the KSERA trials were based on this model in many parts. The interrelations of the constructs are shown in Figure 10.

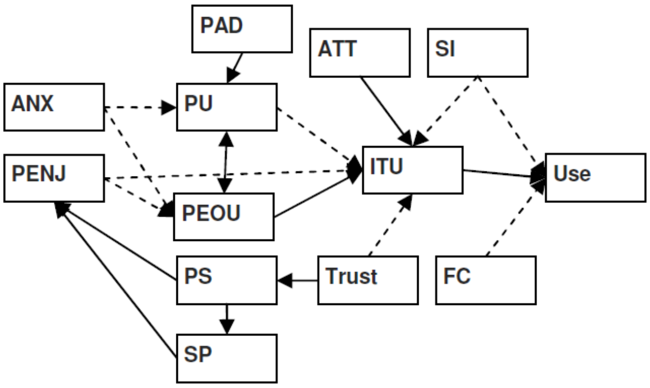


Figure 10 Almere Modell by Heerink et al [Heerink2010]

Qualitative data. Beside the construct based questionnaires also interviews concerning the general impression and possible future use scenarios of the KSERA approach were performed, recorded and transcribed. This should offer the opportunity to gather information beside the defined construct based questions. The results were discussed in focus groups together with all participants. Beside the evaluation together with primary users (older people and people with COPD) also focus groups with secondary users (e.g. medical doctors, therapists, trainers, careers, relatives, call center agents, etc.) were performed. This data gave a good possibility to validate the results from the questionnaire based evaluation and to understand results.

Results

Q1: How are the different influencing items concerning acceptance and ITU pronounced when using the robot with all known functionalities.

All test session and the questionnaires were completed by 14 participants. 2 participants quit after the first iteration. Questionnaire-scores were processed as described above with likert scale-values, while for usage times of use were calculated for each user. A detailed description of the methodology described in the above method section can be found in the public KSERA deliverable „D1.3 metrics for SAR evaluation"

In the following we present the results concerning the proposed acceptance and ITU influencing indicators. We give an overview of the quantitative construct results and of the additional qualitative, subjective and objective measurements. The quantitative results will be presented in descriptive statistics and have to be interpreted in a way that takes into account, that variables are related, not that there is a determining influence. In addition we discuss moderating factors and influences.

The following presentation will take into account the results from the performance evaluation. In order to present valid and comparable result we excluded datasets that were biased by technical problems and/or malfunction related to KSERA internal technical problems.

Table 3 gives an overview of the descriptive statistics concerning the several constructs, before presenting a working system (n=16), after the first and second iteration (n=16) and data of (n=4) trial participants, who were part of the long term trials.

n=16 - all data, <u>inclusive</u>
data including tech. malfunction

n=10 - data, <u>exclusive</u>
data including tech. malfunction

Descriptive Statistics before a working system

	min	max	**mean**	std	min	max	**mean**	std
ANX	1,00	5,00	**3,19**	1,83	4,25	5,00	**4,63**	0,63
TRUST	1,00	5,00	**4,15**	1,13	1,00	5,00	**4,10**	1,33

Descriptive Statistics Iteration 1

	min	max	**mean**	std	min	max	**mean**	std
PENJ	3,00	5,00	**4,4**	0,63	4,00	5,00	**4,79**	0,46
ANX	1,00	5,00	**3,50**	1,76	4,25	5,00	**4,73**	0,65
PAD	1,00	5,00	**3,06**	1,35	2,33	5,00	**3,67**	0,90
PS	1,25	5,00	**3,86**	1,09	3,00	5,00	**4,43**	0,70
SP	1,00	5,00	**2,73**	1,06	1,80	3,80	**2,90**	1,09
TRUST	1,00	5,00	**4,03**	1,38	1,00	5,00	**4,35**	1,33
PEOU	1,33	5,00	**3,73**	0,99	3,20	5,00	**4,06**	0,75

Descriptive Statistics Iteration 2

	min	max	**mean**	std	min	max	**mean**	std
PENJ	3,00	5,00	**4,5**	0,71	4,00	5,00	**4,83**	0,41
ANX	1,00	5,00	**3,89**	1,77	4,25	5,00	**4,91**	0,27
PAD	1,00	5,00	**3,45**	1,65	1,00	5,00	**3,33**	1,94
PS	2,50	5,00	**3,95**	0,78	3,50	4,75	**4,28**	0,54
SP	2,20	5,00	**3,29**	0,84	2,60	4,00	**3,43**	0,82
TRUST	1,00	5,00	**3,86**	1,61	1,00	5,00	**3,69**	1,79
PEOU	4,20	5,00	**4,70**	0,33	4,20	5,00	**4,70**	0,33

n=4 - all data, <u>inclusive</u>
data including tech. malfunction

n=2 - data, <u>exclusive</u>
data including tech. malfunction

Descriptive Statistics Longterm

	min	max	**mean**	std	min	max	**mean**	std
PENJ	3,00	5,00	**4,20**	1,10	3,00	5,00	**4,20**	1,10
ANX	2,00	2,50	**2,25**	0,35	5,00	5,00	**5,00**	0,00
PAD	1,00	2,00	**1,50**	0,71	4,00	4,67	**4,33**	0,47
PS	1,75	3,00	**2,38**	0,88	4,75	4,75	**4,75**	0,00
SP	1,00	1,40	**1,20**	0,28	3,60	3,60	**3,60**	0,00
TRUST	2,00	3,00	**2,50**	0,71	5,00	5,00	**5,00**	0,00
PEOU	5,00	5,00	**5,00**	0,00	4,80	4,80	**4,80**	0,00

Table 3: Descriptive analysis of the most relevant constructs for evaluation of the acceptance of the approach and an intended future use

MOST relevant constructs (PENJ, SP, PS, TRUST, PAD)
In the following we give a more detailed overview about the constructs that are additional to the classical UTAUT model and are relevant for studies concerning the acceptance of the use of SARs in homes of older people.

- PENJ: Perceived enjoyment

Literature reports that perceived enjoyment (PENJ) with a technical system positively affects the subjective intentions towards a future use of the system (Heerink2010). An element of pleasure, fun or entertainment when interacting with a technical system and in this respect also with a robot may influence the acceptance of the system. This is even more important in the case of a socially assistive robot. Studies (Zhang 2006) point out that perceived enjoyment has a direct effect on ease of use and intention of use. This fact is also most important for a use of a SAR together with vulnerable target groups, who are intended to use the robot solution as a helper in daily life.

During the end evaluation of the KSERA prototype the PENJ was evaluated by asking construct related questions after the different test-iterations (1st iteration, 2nd iteration, long-term iterations). The overall result shows that the PENJ was rated with high levels (mean>6.0 on a 7-likert scale) in all iterations. During the long-term test iterations, which were conducted together with n=3 persons the PENJ value stayed constant at a level>6.5 on a 7-likert scale.

These quantitative results were also represented in the qualitative comments presented by the trial participants.

> *“I think he’s funny.”*
>
> *“I just like him.”*
>
> *“Usually I don’t like technical stuff, but I like him.”*
>
> *“I’m surprised what he is able to do.”*
>
> *“The first thing I would do in the morning is saying ‘good morning’ to the robot before giving any commands.”*
>
> *“I got the feeling that he understands me, when he is looking at me.”*
>
> *“To be honest – I was really looking forward to the test today.”*
>
> *“There would be an inhibition threshold to talk to him when somebody else is in the room.”*
>
> *“It would be nice, if from time to time he asks me how I am currently feeling.”*
>
> *“I like it when he’s looking for my face.”*

- **SP: Social presence**

The social presence can be described as the feeling of being in the company of a social entity: the perceptual illusion of non-mediation (Lombard1997). It does not seem to be unusual for humans to engage with technology as if it were a social entity (Shibata2003). If the technology is shaped like an animal or even humanoid this effect gets even stronger. This social presence and the expected social abilities of the robot influence the described PENJ. If the factors for social presence raise, also the PENJ values raise.

During the end evaluation the values for the SP varied between 2.9 <mean<3.6 mean on a 5-point likert scale in the included data-sets. It could be recognized that there was difference between the first iteration (mean=2.9, std=1.09), where the abilities of the robot were influenced by a higher rate of technical problems and the second iteration (mean=3.43, std=0.82), which was a free interaction session of 20 minutes, where only the robot and the trial participant interacted with each other and technical problems were avoided and worked around by the research team using "wizard of oz" methodology. During the long term iterations, where the trial participants already knew the system very well, the levels for SP stayed on a level of mean>3.5.

"I can imagine that he will be perceived like a pet or a stuffed animal."

"I think it is good for people who are living alone."

"I like it when he comes to me and (I like) the lights in his eyes."

"He is a nice conversational partner – he does not talk back, but listens to me."

"The different colors in his eyes look funny."

- **PS: Perceived sociability**

Perceived sociability as a construct in the evaluation as a part of an acceptance model is especially important during the evaluation of acceptance of a SAR solution. A SAR that supports vulnerable target groups needs social abilities to function as assistive device. **Related studies found that the construct is a determining influence on Social Presence and Perceived Enjoyment. On the other hand the construct TRUST** has influence on this construct.

During the KSERA end evaluation the PS of the robot was rated on a 5-point likert scale from mean=4.43, std=0.7 (1st iteration), mean=4.28, std=0.5 (2nd iteration) to mean=4.75 during the long term iterations. The lower rates of technical performance during the 1st iteration had no effect on the value of the PS compared to the 2nd iteration (the free interaction session).

"It is nice always having someone around to talk to."

- **TRUST: Trust**

The construct TRUST is a very important measure in order to get feedback about the trust the trial participants have in the robot and the whole solution – which is most important in the case the robot works as a helper in daily life and gives advice to the user. It gives also feedback about the future compliance of the user towards the robot's advices and is therefore essential for a the human-robot-interaction. Trust is claimed to have a direct influence on Intention to Use (ITU) and is related to social abilities or social behavior: a robot or screen agent with more social abilities is supposed to gain more trust by its users.

During the KSERA end evaluation the trial participants showed levels for TRUST mean>3.69. The long term trial participants evaluated the TRUST even with a level of mean=5.0. These results show, that the KSERA approach of combining the KSERA robotic helper with eHealth applications in daily life is trusted, which is very relevant for a future use in homes of the users and an application for disease management support of vulnerable people (e.g. COPD or older people in general).

"I would trust the advice of the robot as long as it seems useful and reasonable."

"I would trust him, but before I follow his advice, I would think about it, if it makes sense."

"I would trust him. If I wouldn't, it would not make sense at all."

- **PAD: Perceived adaptability**

As KSERA is dealing with a very vulnerable group of target users (people with COPD and older people in general) the future adaptability of the solution is a very important feature, as the conditions and the abilities (e.g. mobility, sensorial abilities, effects of medication) of the users change over time. Ambient assistive technology would need to adapt to these changes in conditions in order to provide appropriate support. Robots and especially humanoid robots offer a great chance for adaptability of the technical solution. Related projects introduce adaptability in the design guidelines for robotic products that support the ecology of aging (Forlizzi2007). If people perceive the system to be adaptive towards their needs, they will find it more useful and will accept the system more easily in their own living circumstances.

The quantitative result show, that the KSERA system was perceived as an adaptive system, which could be adapted in the daily life of the trial participants and could be adaptive to their needs and limitations. Although the results show high levels of PAD, some of the trial participants recognized problems in adaptability related to their limitations, e.g. sound level, speed of the robot. However they were confident in the fact, that the KSERA system and related future systems could overcome these issues easily and that the approach of using a humanoid robot with its multimodal communication possibilities and the mobility would be the solution of choice for them.

"It would be nice if he would be able to adapt to how I am currently feeling and give me some advice like drinking more or laying down for a while."

"I can imagine that he can adapt to my personal needs, if he is programmed like that."

Overview on quantitative results concerning Q1:

Figure 11 and Figure 12 give an overview about the quantitative results related to the ITU influencing factors as boxplots. Figure 11 shows descriptive boxplots of all datasets – including those where technical problems were evident during the evaluation iterations. Figure 12 shows data excluding the datasets with technical problems. The effect on the constructs is evident (significance test t-test)

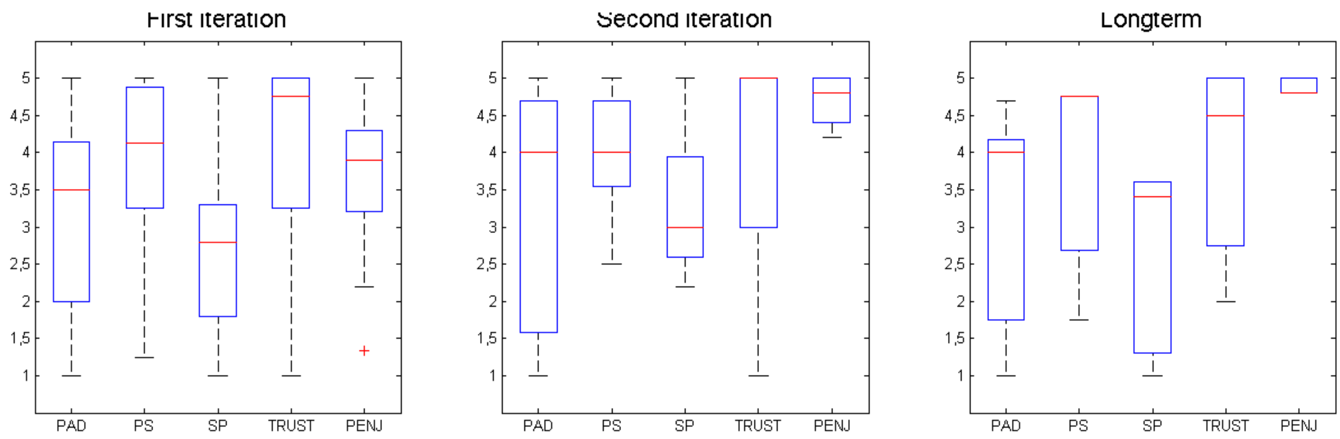


Figure 11: Overview on the results concerning Q1 inclusive malfunction data

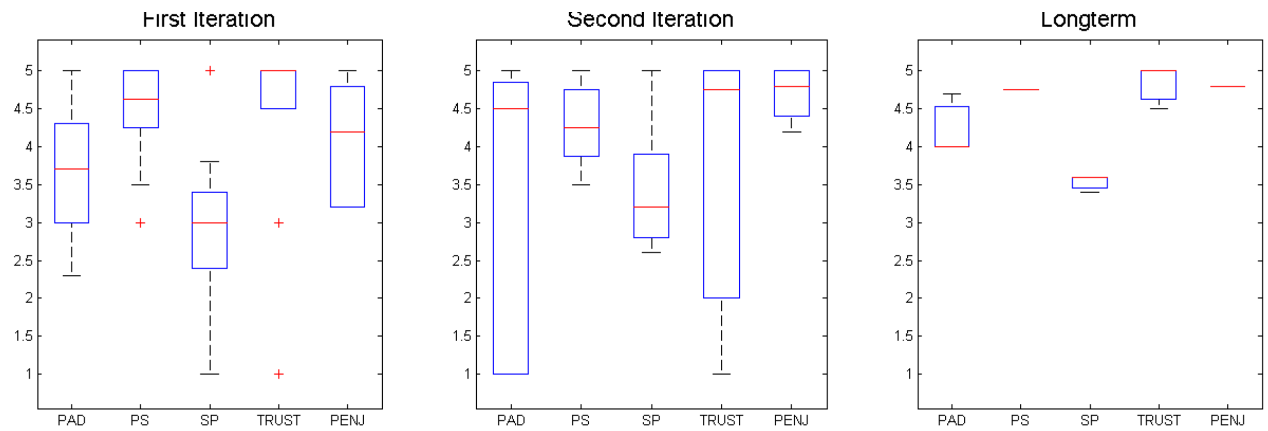


Figure 12: Overview on the results concerning Q1 exclusive malfunction data

Figure 13 shows the interconnections and relations of the parameters defined by the Almere Model for the evaluation of social presence and social abilities related to the proposed ITU. Figure 14 again gives an overview of the course of the parameters after the different periods of the end evaluation.

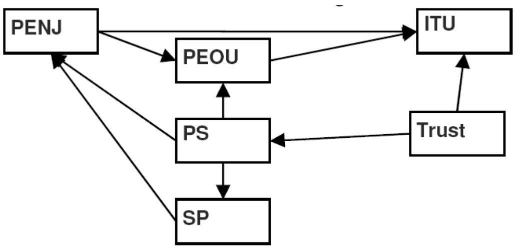


Figure 13: Model for studies on social presence and social abilities

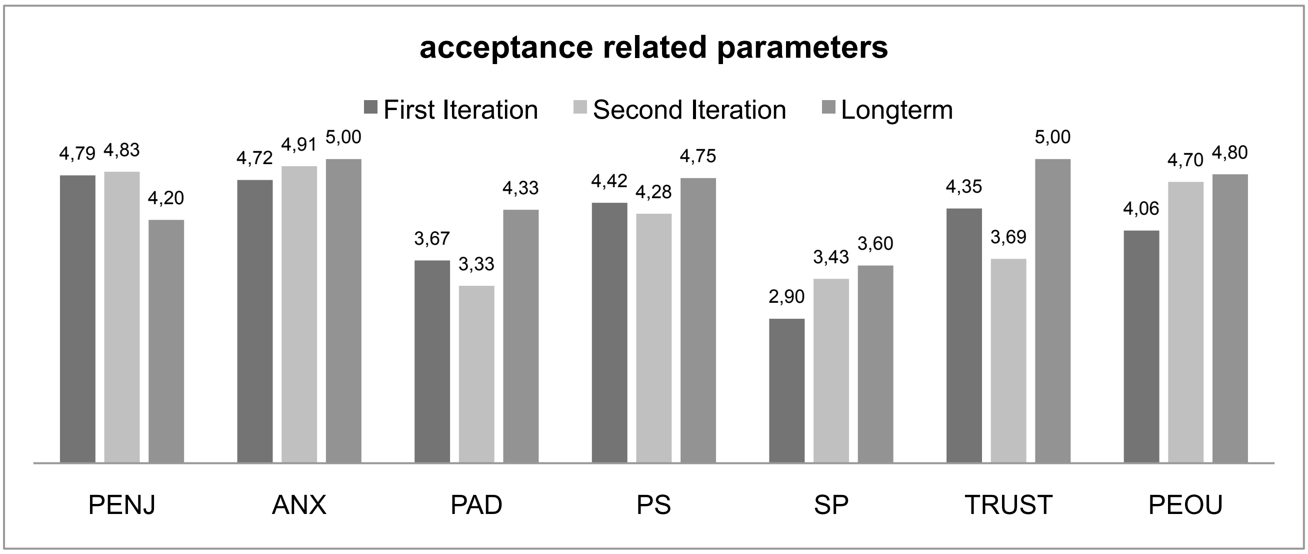


Figure 14: Overview of the course of parameters during the end evaluation phases (n=8) excluding dataset with technical problems

Based on the results presented in the previous chapter we will discuss three related research question in the following.

Q2: What are the influencing factors for an ITU of the KSERA approach (Use is determined by (/1) intention to use and influenced by (/2) social influence and (/3) facilitating conditions.

The quantitative results showed, that the influencing factors for the intended future use (PENJ, PEOU, TRUST) were situated in areas with values of mean>4.0, which can be described as high levels. The intended periodic future use showed a level of mean=3.6. Qualitative feedback from the trial participants showed, that they saw an advantage of the SAR approach in many use cases e.g. compared to state of the art solutions (e.g. feedback from a computer system). They stated that the KSERA approach is easier to use and self explanatory. Another ITU influencing factor – the TRUST towards a technical system – was also confirmed by the trial participants. This trust was described by some trial participants and they stated, that it was different than trusting e.g. “only personal computer”.

One trial participant said that she was very afraid of using a computer, due to her history in her last working years. However she had absolutely no problem using and interacting with the robot. She said, that it’s not a technical system for her – even if she was aware, that it is a machine.

During the KSERA trials the perceived enjoyment PENJ was not only directed towards entertainment, but also with e.g. doing exercises together with the robot – being able to perform as shown by the robot, getting information about the own status. Although the levels for the PENJ were also very high, especially this factor is very dependent from several moderating and facilitating factors as well as adaptability towards their subjective situation and use case. Therefore no general statement about the PENJ related to the KSERA approach can be set.

Q3: Is the KSERA robotic approach a good way to implement a social robot?

One of the main aims in KSERA was to implement social abilities in the robot. The aforementioned evaluation factors also give feedback about the proposed approach. Both the factors for social presence (SP) and perceived sociability (PS) give a feedback on how accepted the KSERA approach was to implement characteristics in the robot, that make the robot a social robot – a robot that the people can integrate in their social and mental model – a robot that the trial participants can understand and interact with. In the KSERA approach the social robot included

- communication by question and answering, hints for several use case purposes
- expression of emotions by facial expressions (e.g. LED colors, sounds or speech, gestures)
- usage of natural interaction cues (face recognition, face direction recognition, gestures)
- exhibition of a defined personality ("the friendly helper in daily disease management")

The KSERA approach of using a humanoid robot with the mentioned social abilities was appreciated by the trial participants. The factors for PS and SP are located in value-areas comparable to related studies (e.g. Heerink2010, Forlizzi2007, Zhang2006, Lombard1997). The combination of a social robot together with the proposed KSERA abilities of disease management, entertainment, communication and home control resulted in the impression of a social and intelligent helper.

Nevertheless the trial participants also recognized the deficits of the system (e.g. speed, robustness, performance, adaptively – see also performance evaluation). However this had only minor effects on the mentioned impression concerning the sociability. Therefore the KSERA approach can be seen as promising and was accepted by the trial participants.

In future projects the dialog system should be optimized and extended. The communication skills were described as "adequate", however the trial participants noted the limited adaptability in the conversation. The combination of verbal and non-verbal interaction cues and the multimodal conversation strategies facilitated the engagement and transfer of information.

Q4: Is the KSERA robotic approach a good way to implement an assistive robot?

KSERA used a SAR approach instead of an often used service robot approach[5]. Beside the social role – which was described in the previous passage a SAR also has to fulfill an assistive role as well. This assistance can be assistance in daily life activities, assistance with mental or physical deficits, assistance with communication tasks or assistance in carrying things. In KSERA the main focus was on

- assistance for disease management tasks (motivation for physical parameter measurements, information provision, motivation for physical training),
- communication (in case of emergency towards a call center, for contacting friends or relatives) and
- daily life activities (home automation support).

[5] http://www.service-robots.org/

The described parameters of PEOU, TRUST, SP and PS provide a good feedback concerning the acceptance of the approach as an assistive robot (whereas assistance is related to the described domains). In addition to the already mentioned parameters of TRUST, SP and PS, the perceived ease of use (PEOU), which is influenced by several usability factors (e.g. effectiveness, efficiency, satisfaction, learnability, memorability, error rates) is identified to play a crucial role. In the KSERA end evaluation the PEOU was rated with mean levels of mean=4.06, std=0.7 during the first iteration, mean=4.70, std=0.33 after the second iteration and mean 4.80 during the long-term iterations (all on a 5-point likert scale).

Nevertheless these results are based on a working system without major malfunctions (as it will be part of future user homes). The PEOU is influenced by malfunctions and low technical performance parameters of a large extend. Therefore the robustness of the solution was described as most important. Related to this the solution of using a biped robot for assistive tasks - given the state of the art robustness – cannot be confirmed in general. A discussion on this topic can also be found in chapters 3.5.1, where the general approach and added value is discussed in a more holistic way.

3.5. Emotional feeling and psychological impact during the free interaction with the KSERA prototype

General description

The introduction of novel robot technologies in homes of vulnerable user groups and older people is changing the practice of connected care and healthcare. During the KSERA end evaluation the aim was to analyze how technological, social and psychological factors affect human robot interaction, including the major factors of usability, and acceptance.

During the KSERA end evaluation we gathered video data (objective) through observation in order to analyze emotional feelings and psychological impact factors based on quantitative data. In addition we discussed in inter- and multidisciplinary future user groups what effects could arise. The main research questions were:

Q1: Do the primary users‘ data gathered by subjective metrics correlate with the objective measurements?

Q2: Do the users integrate the robot into their mental model?

Methods used

During the evaluation the primary users were interviewed in the course of the evaluation sessions. The secondary users were interviewed and met in focus groups, where the topics were discussed in interdisciplinary discussions.

Results

Q1: Do the primary users‘ data gathered by subjective metrics correlate with the objective measurements?

In order to analyze the objective datasets we transcribed the video data (n=8 trial participants) and introduced a grid, that offered the possibility to analyze the interaction concerning several emotion domains (e.g. uncomfortable, amused, interested). Figure 15 to 22 show interactions and emotions by the trial participants during a free interaction session with the robot. For the end evaluation and review concerning emotional feeling of the trial participants all videos were analyzed by local researchers, who could also understand the language and have the same cultural background.

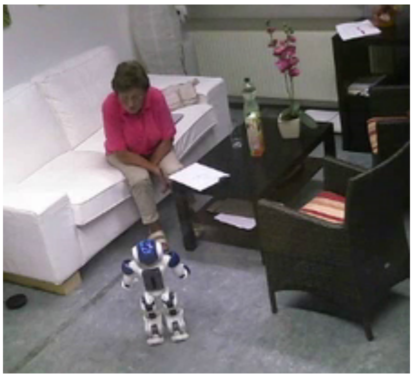


Figure 15: Free interaction with the robot

Figure 16: uncomfortable

Figure 17: ashamed

Figure 18: amused

Figure 19: jittery

Figure 20: interested

Figure 21: impatient

The analysis showed, that the subjective measures could be verified by the objective analysis of the interaction between the trial participants and the robot. A more detailed and quantitative presentation of grid data was not performed during the KSERA trial analysis

Q2: Do the users integrate the robot into their mental model? Is the KSERA robotic solution perceived as a social robot? What are factors that influence the perception based on expectations?

Based on the quantitative results multidisciplinary discussions were performed in order to analyze the proposed research questions. In this context one of the main statements was, that primary as well as secondary future users said, that use cases and scenarios shown by the research members and also the way the use cases were implemented were easy to understand and useable for them. In the following we will briefly discuss the influences on the integration of the robotic idea in daily life.

Personal idea of a robot. One of the main issues was the already existing personal idea of a robot, which was strongly influenced by representation of robotic solutions in media over the last decades. Due to this there have been certain expectations in many acceptance and usage influencing domains (communication, navigation, adaptability, degrees of freedom in movement, context awareness, etc.). State of the art solutions in robotics (even combined with external sensors and inference like in the KSERA prototype) cannot sufficiently fulfill these expectations. During the trials the participants stated, that these high expectations were often due to the humanoid shape of the robot. They know, that there are robots integrated in the homes of people already (e.g. vacuum cleaning robots), however those machines were perceived in a different way and were not seen as the solution, that should be used for uses cases tested in KSERA.

Aim of a robotic solution. In addition to the personal idea of a robot the purpose and aim of a robot was discussed. Again this idea was influenced by the personal idea of a robot that can support people by implementing a service robot approach. Although the idea of introducing a robot, whose main aim is to assist socially all trial participants stated, that service robot abilities – at least to some extend - will be needed to introduce a robot in the users homes in future.

Humanoid Shape and Abilities. The KSERA robot platform NAO with its humanoid shape was perceived as a sympathetic platform, that does not originate negative emotions e.g. fear. Due to the humanoid shape and capabilities it was very easy for the trial participants to understand the multimodal communication strategies. No explanation was needed to understand the advices, hints or commands.

In this context the size of the robot was discussed. On the one hand a small robot like NAO (height=60 cm) was seen as beneficial, as it would e.g. not be able to harm a person and therefore not generate negative emotions. On the other hand the trial participants also stated the disadvantages in the small size of the robot concerning the limited functionalities and limitations concerning the area of vision (e.g. if the trial participants would lie in bed). One important dimension of physical space is vertical height. Social dominance hierarchies typically imply vertical metaphors of having a low position as being more obedient than having a high position or being tall. Based upon the psychological implications of this for social relations and nonverbal behavior [Hall05], perceptions of power [Schubert89], and expressing authority [Argyle88], it is plausible that this vertical dimension would influence the psychological experience of thrown voice locations, too.

Dominance and role of the robot. Related to the last passage the role of the robot was discussed. In KSERA we implemented a socially active and friendly robot. This was appreciated by all trial participants. During the multidisciplinary sessions with motivation trainers and therapist also characteristics like being a dominant trainer or a rigorous teacher were discussed. These roles of the robot were not implemented in KSERA, but seemed to be beneficial in special cases in order to support extrinsic motivation strategies in future projects. It was also stated by the experts, that these implementations will have a severe effect on the acceptance of the robot at the users' homes and on the compliance concerning the proposed usage scenarios.

Area of privacy. In the course of the KSERA end evaluation we discussed the area of privacy, which the robot was allowed to enter. In both trial sites (Austria and Israel) the implemented area of privacy (r>1m) was accepted and appreciated. A more detailed study on the area of privacy was performed during the project in a lab study at the TU Eindhoven (Torta2012).

In summary the KSERA approach was appreciated and implemented in the mental model. It was perceived as useable in a future use at home. The limitations of the state of the art were obvious to the trial participants, however they could understand the future use based on the scenarios and use cases.

3.5.1. General impression of the KSERA SAR approach and the developed functionalities

General description

In order to get a holistic view of the KSERA approach the impression and ideas of different future user groups was evaluated. The needs, which were defined in WP1 were discussed under the light of the two evaluation phases, the KSERA prototype and future usage scenarios. The main research question can be defined as:

- **Q1: What is the added value that the KSERA approach offers to the users in comparison with related approaches (e.g. conventional systems) concerning the user need domains?**
- **Q2: What are the limitations (in the KSERA SAR approach) in the different user need domains**
- **Q3: What should be changed, in order to really use the robot at home**

Methods used
During the evaluation the primary users were interviewed in the course of the evaluation sessions. The secondary users were interviewed and met in focus groups, where the topics were discussed in interdisciplinary discussions. The perceptions and attitudes of professional in the health fields towards the KSERA system as well as identification of advantages and disadvantages of marketing and assimilation of the use of robots were examined through a focus group held in November and Decemeber 2012 and in January 2013.

Trial participants
The groups of primary users were the same as described above.

One group of secondary future users consisted of health care employees who work in multidisciplinary centers of Maccabi Health Care Services (called "MOMA") that provides a solution to chronically ill patients using advanced technologies for remote monitoring such as the smart phone, video conferences and tele-medicine. The group included representatives of the following professions: geriatricians, nurses, social workers, dietitians, and clinical pharmacologists, a total of 12 participants (including research staff, headed by the Chief Researcher, Dr. Joseph Bregman). One group included care professionals in Austria in the senior citizen center Schwechat and formal carers and health experts from related institutions. One group included therapist related to the geriatric field and COPD training.

Results
Discussion group participants held, in general, an extremely positive attitude towards the system, viewing it as an important added value beyond what they currently utilize in the online center. In addition to the perceived benefits of the system, they also proposed numerous and diverse improvements to increase human robot interaction and therapeutic benefits – based on an understanding of the potential and recognition of the many possibilities embodied by the system.

Q1: What is the added value that the KSERA approach offers to the users in comparison with related approaches (e.g. conventional systems) concerning the user need domains?

The following points show the most significant perceived advantages perceived by **secondary users**, which were stated several times during the discussion and focus groups

- **Increased response to treatment**
- **Alert in case of emergency situations**

  for caregivers and family members

- **Solution for loneliness**
- **Replacement for the human caregiver – for patients for whom a human caregiver seems not to be the right solution**
- **Supplementary and adjuvant solution to the attending physician's treatment**

  The system can reduce the number of visits via parameters that are monitored when the doctors speak with the patients through video conference and telephone.

- **Lightening the burden and increasing confidence of the family members / primary caregiver**

Participants in the group discussion discussed the benefits of the system for family members. They believe that family members should be involved and receive current information about the elderly patient's condition. They also believe that the information should first be delivered to the family member.

- **The issue of novelty and renewal that might positively affect assimilation of the system also arose.**

One of the main target groups of KSERA are vulnerable user groups like **people with COPD and older people** with age related limitations in general. The specific user need domains (defined in KSERA WP1) and related use cases were the basis for consideration concerning different clusters of interest (mobile UI, availability, novelty, social abilities, etc.) related to the added value of the proposed approach.

Mobile User interface

The idea of a multi purpose mobile user interface was appreciated by all primary future users and was described as an advantage for their daily life. Due to the characteristics of the different limitations of the trial participants (e.g. mobility, sensory, chronic diseases, etc.) the advantages of the subjective future scenarios differed. The paradigm change of a user interface that comes to the user – no matter if it is triggered by the system or by the user – was appreciated.

The trial participants were confident that a robot, who reminds them of performing different tasks – e.g. measurement of physical parameters is an advantage compared to technical state of the art solutions e.g. a ringtone as it is more comparable to being reminded by a human being.

The trial participants stated a real advantage concerning the possibility of a mobile user interface for communication purposes in case of emergency but also for daily life communication. This change towards a mobile multimodal (video, sound) communication was appreciated especially for emergency situations, when the user can't trigger a communication channel by his/her own any more. But also for daily life the mobile video communication was described as a meaningful tool – in case the performance of the robotic solution (in terms of e.g. speed, robustness, adaptability) would be enhanced. During the end evaluation this approach was favored towards state of the art solutions (e.g. a tablet solution).

Secondary users (e.g. medical doctors, carers, therapists) also saw an advantage especially in case of emergency. The possibility of the user interface to enter the area of interest of the user and give a multimodal feedback on the relevant situation was stated as an advantage towards state of the art solutions (e.g. sound communication via a wrist watch) Also the possibility of a mobile motivator, that comes to the area of interest (e.g. the place where the person is sitting during the day) and tries to motivate the users was appreciated.

Single Multipurpose Interface

The possibility of using one central user interface for several activities and purposes in daily life was appreciated very much by the trial participants. In this context also the mentioned (see chapter 3.4) humanoid implementation and the social abilities (e.g. concerning communication, expression of emotional states, etc) were seen as an advantage compared to solutions like a tablet, TV or single DM tools. In this context also the integration of several functionalities (e.g. measurement of physical parameters, environmental information, house automation, novel forms of communication) were judged as advantageous.

The possibility of having only one intelligent interface, that can be interacted with in a very intuitive way (concerning multimodal input and output channels) was seen as an advantage. This intuitive way also overcame the prejudice concerning technical solutions and personal computers in special.

Also the secondary users quoted, that one single interface is an advantage in their point of view. Especially if it is context sensitive and aware, as well as connected to common interfaces that are already used by primary users in daily life.

They quoted, that the possibilities of implementing several functions and features in the robotic solution, which can be adapted towards the subjective needs (e.g. special physical training, reminders and hints related tot the subjective situation and entertainment components) would be a benefit and support also for their work. E.g. they can use one interface to give instructions to the primary users remotely and use the same interface for e.g. showing physical exercises.

Availability and Regularity

In this respect secondary users quoted, that the 24-7 availability and the possibility of regular use of such a solution in combination with the possibility of monitoring of several health related parameters would be a great benefit for the treatment. Current solutions only give them the possibility to show and offer information regularly but not sufficiently frequent and often not in an intuitive way (e.g. physical therapist can only offer written descriptions of training movements)

Q2: What are the limitations (in the KSERA SAR approach) in the different user need domains

In the following we will discuss the limitations of the KSERA socially assistive robot (SAR) approach independently from the performance of the existing KSERA prototype – based on user needs, use cases and future scenarios. The limitations of the existing integrated prototype and the effects on a real use at home will be discussed in the next question Q3.

Performance limitations for a usage as mobile user interface

As described in the previous parts of the document the idea of a mobile user interface was appreciated very much by primary as well as secondary users. The approach of using a very small, humanoid robot (NAO) in the role of an assistive robotic helper however showed limitations in the realization of the proposed use cases.

Speed. One of the most obvious limitations of the KSERA approach was the insufficient speed of the robot. The trial participants stated that in case of emergencies, where a fast interpretation of the situation would be necessary, the speed would be insufficient if the place of emergency was not in the same room, where the robot is located. It would take several minutes or even worse if the robot would be located in a different floor (only elevator solutions would be possible – as the robot cannot climb stairs). Also in case of the communication scenarios the speed would be a very limiting factor – as the robot has to approach the user before an incoming call can be established. If the robot is too far away from the user it takes an exceeding amount of time to fulfill that task. The caller would most probably end the call before it would be established. The video communication via a projector needs some distinct pre conditions (e.g. distance to wall, angle to wall, light conditions) therefore it takes by now too much time to establish a video call.

Another limiting factor concerning a future use as a mobile user interface stated by tertiary users (technicians) was that the KSERA SAR approach needs external sensor data (e.g. a ceiling cam in order to localize the robot and person. It would be very difficult to install these camera systems in every room of the primary users' homes.

Robustness. The robustness of the solution is one of the most influencing factors concerning several acceptance related factors. E.g. the PEOU is influenced by malfunctions and low technical performance parameters of a large extend. Therefore the robustness of the solution was described as one of the most important issues. Related to this the solution of using a biped robot for assistive tasks - given the state of the art in robot robustness – cannot be confirmed in general. The robustness was also related to ethical dimensions (e.g. safety, security, privacy). Only if this robustness will increase the trial participants stated the system to be useable in older people‘s homes.

Size and Strength. The KSERA SAR approach uses a very small robotic solution (size<60 cm). This fact limits the usage e.g. for communication or expression by the robot for different user needs domains, if the primary users are very limited in their area of possible interaction (e.g. people with higher level of COPD (GOLD 4) or older people who are frail and have to stay in bed. In this case the robotic solution can be used as interface (e.g. next to the bed), but not as mobile interface any more. In strong connection to the size is the related strength of the robot. Although per se a socially assistive robot should not be used for carrying things, in many use cases a robot that can e.g. bring a measurement device to the vulnerable user would be a big benefit – as immobile persons often cannot reach the devices by themselves.

Endurance and power consumption. In order to use the robotic solution as a mobile user interface the endurance of the batteries and the related time of usage has to increase (actual status: 30 minutes in working mode).

Limitations in the multimodal in- and output channels

The limitations concerning the communication abilities of the robot are mostly related to the input channels of the proposed robotic solution and the general state of the art. Whereas the proposed solution of using a multimodal robot with a humanoid shape, can be used gainfully to communicate towards the user (talk, gestures, facial expressions, etc.), the abilities concerning the input channels are limited.

Speech recognition. The speech recognition rates have to achieve a high rate of performance (>95%) in order to guarantee a fluent and meaningful communication with the robot. This is even more important for vulnerable user groups, who have severe limitations in talking (e.g. reduction of breath due to COPD). A reliable speech recognition as one of the main input channels is also important in terms of safety, privacy and security. An example: If the system misinterprets the inputs and opens an unwanted video communication in a situation, where the user does not want to be filmed, the user will lose the trust in the system and will most probably also not use it in case of an upcoming emergency, which has effects on the safety of the person.

The KSERA speech recognition is below the described values, which is related to the used speech recognition solution and on the other hand on the use case. If the robot is in the room, not located directly in front of the user the recognition rates are lower. Also the distance to the user is important. The rates are very dependent on the loudness and clearness of the spoken words. This is another crucial problem, as many users of the proposed vulnerable target group can speak only silently.

Face recognition. The face recognition of the proposed approach is limited by the size of the robot. If the user is e.g. lying in bed, there is no possibility for the robot to recognize the face due to the fact, that the robot cannot see the face with its cameras.

Single user approach. Another limiting factor is the used “single user” approach. The developed setting and algorithms are limited to one user and one robot in the room. This scenario will not be realistic in a use at older or vulnerable primary users at home who are not living alone.

Q3: What should be changed, in order to really use the robot at home - Future Improvements and Developments for the System

According to the primary users the following changes / enhancements would be useful:

- Nao should move faster
- Nao should be able to check the correctness of the exercise performance and give appropriate feedback
- The voice level of Nao’s voice should be adapted to surrounding noise (e.g. louder if the tv is on)

- Nao should react to the current personal state of the user and give hints according to that (e.g. „you should have a glass of water" or „maybe you should lay down for a while.")
- Nao could also motivate to do other things like reading a book
- Nao could entertain in different ways (not just playing music) – reading out or telling some stories or jokes, play some games, …

The following statements are qualitative results based on the outcome of the focus group discussions with secondary users in Israel.

- "Transmission to the center of non-response to therapy – indicating failure to take the medication but also on correct taking of medication„
- It should have a system of questions that the robot will ask such as when it notices low blood pressure, it will ask "Did you take the medication?" or "Did you eat something today?"
- "The ability to receive and process information from the patient's vocal / emotion in the patient's voice, such as shouting, etc. Under these circumstances, the robot will ask the patient whether he wishes to contact family or the caregiver."
- "The robot should play a greater role in therapy. For example, the role should not be limited to instructions to perform physical activities but also in terms of adapting the site where the patient is about to perform physical activity or check whether the environment is sufficiently safe (rugs, darkness, lighting) in order to prevent falls, a major issue among the elderly population."
- "A network of robots" – the ability of the robots to communicate with each other so that they are part of the group of robot caregivers, possibly alleviating loneliness
- "The ability of various caregivers to connect with the robot – when the various caregivers arrive such as an emergency response physician), they should have the option of connecting or the ability to record through the robot. Even the ability to transcribe the visit."
- "Inclusion of other features that help break the loneliness such as a karaoke system or book reading."
- "Communication with medical services, such as help in scheduling appointments, since this is an elderly population that must schedule appointments and based on the discussion group participants' experience, this is an existing need."
- "The addition of parameters for monitoring such as weighing (to make sure that the patient has not lost weight) "
- "The ability to program the robot from a distance, to add features or programs without having to physically touch the robot, such as updating exercises, is perceived as an advantage."

Other Issues:

- Importance of linking the system with other caregiver information systems – not a "stand alone".
- Coping with situations involving an HMO decision to end treatment through the robot, the possible problem of dependence on the system.
- The willingness and ability to pay for the robot/service – must be revisited and examined.

3.6. Advantages and limitations of the approach concerning the proposed usage scenarios

General description

The approach was evaluated regarding advantages and limitations using a combination of state of the art solutions for ubiquitous sensing with the KSERA robotic solution. The usage scenarios, which were defined in WP1 were discussed under the light of the two evaluation phases, the KSERA prototype and future usage scenarios.

Methods used

During the evaluation the primary users were interviewed in the course of the evaluation sessions. The secondary users were interviewed and met in focus groups, where the topics were discussed in interdisciplinary discussions.

Results

In the following a description of the results concerning advantages and limitation of the approach using a combination of state of the art solutions for ubiquitous sensing with the KSERA robotic solution is given. These advantages and limitations are discussed concerning the usage scenarios for a relevant disease management defined in WP1. Disease management as a systematic approach to the long-term management of patients with chronic illnesses plays a crucial role in daily life of people with COPD and older people in general. Most of the conditions that are related to a disease management approach are chronic and progressive with episodes (in COPD exacerbations) and require acute intervention. Both the chronic nature and the exacerbations are controllable. Such control in most cases relies on the patient. Thus, patient self-management in collaboration with the healthcare team, carers or medical experts should reduce or even better prevent the need for acute interventions and hospitalizations. While standard medical care focuses on acute interventions, disease management attempts to limit or prevent the need for such intervention. In KSERA the following usage scenarios were defined:

- **Monitoring of physical parameters**
- **Supported Communication,**
- **Environmental Information,**
- **Medical alert and emergency support**
- **Therapy support**

Monitoring of physical parameters:

One of the main bottlenecks in regular self-performed physical parameter monitoring at home is the motivation of the users/patients to perform the measurements and document them. In KSERA a combination of a socially assistive robot together with ubiquitous physical parameter sensors (e.g. a pulse-oximeter) were used to realize monitoring usage scenario (see description in chapter 0).

Advantages. During the evaluation the feedback on this usage scenario was perceived very positively by the primary users. The aspect of having a tool that reminds the user of performing the measurements regularly was appreciated. In this respect also the multimodal interface for explanation purposes (e.g. how to measure) was seen as advantage - also for user groups with cognitive decline. The direct contact with the user interface (the robot) was perceived in a positive way – the robot was perceived as a motivator. This expressed feeling correlates with the quantitative measurements (e.g. KSERA motivation score, PEOU). The possibility of a multimodal feedback (e.g. speech, movements, expression of emotion) for motivating the users, but also for giving them feedback was appreciated.

Similar feedback was given by secondary users. They also appreciated the possibility of having a tool for actively motivating the primary users to perform physical parameter measurements.

Limitations. One of the limitations, that were stated by the trial participants was the non-ability of the robot to bring the measurement device to the person. This was perceived as a disadvantage especially for immobile older people and persons with COPD level>GOLD 3. Another perceived limitation was the missing awareness of the person, who is the actual user of the system. It would have been perceived as great benefit, if the robot could recognize the person and would be therefore more individualized. In a future use primary and secondary users would appreciate if the system could give feedback on exact measurement values, but also in a more abstract way (e.g. traffic light metaphor).

IT supported communication:

Communication supported by the KSERA system was presented to the trial participants for daily life communication scenarios (e.g. with a friend or relative) and medical support scenario (e.g. with a medical expert or call center). (a description of the use case can be found in chapter 0). The attitudes and preferences of family or friends who communicated with them remotely via the device were assessed through the described metrics.

Advantages. Overall experiences were positive. Responses from our trial participants indicated that in general they appreciated the potential of this technology to enhance their physical health and well-being, social connectedness, and ability to live independently at home. The approach was described as helpful and meaningful especially for future users, who are physically impaired and/or immobile. The possibility of having a mobile interface that approaches towards the area of interest instead of the human user being the mobile agent was rated as a very positive feature. The control of the communication (e.g. video communication) via human-robot dialog was easy to understand and handle for the group of trial participants. The approach of having a beamer was seen as a positive feature by the primary user trial participants. However the trial participants also stated that a state-of-the-art solution (e.g. a tablet using Skype) was comparable to the beamer solution.

Future remote users, who were friends, relatives or medical experts, were more likely to test the mobility features and had several suggestions for additional useful applications.

Limitations. The proposed approach for IT supported communication was limited by performance influencing factors. The size of the robot and the capability of carrying weight by the robot limited the also the abilities and performance of the projector solution (e.g. a very small and light weighted projector with lumen limitations had to be used). The robot's localization and navigation skills and the related speed towards the area of interest and beaming was one of the obvious limitations and was criticized by all trial participants. Only if this limitation can be handled the approach of having a mobile projector for video communication for daily life tasks and (remote) emergency usage was seen as a meaningful approach. Also the limitations in the areas in older people's homes, where the video picture can be projected, was seen as a critical point. Limitations were also identified concerning different user groups.

The used approach was also considered as problematic concerning privacy aspects, as a mobile camera – as used by the robot – can limit the privacy of the future users. Therefore strategies for feedback concerning video recording, possibilities for video interruption and secure data transmission have to be implemented.

Environmental information

During the KSERA trials the possibility for gathering, interpreting, abstracting and presenting environmental information to the primary users was demonstrated using the described scenarios. The approach was also discussed with medical experts and carers.

Advantages. The approach of having enriched information (e.g. environmental information together with physical parameters information) presented by the central user interface (the robot) was rated as a positive feature. The approach of having one multimodal user interface was described as logic way for presentation of the information. The presentation of parameters in an abstracted way was described as very positive.

Limitations. In order to achieve a meaningful use of the information enriched by environmental information, the system has to be more adaptive to the users' needs and abilities. The inference based on the environmental parameters combined with the physical parameters (as incremental part of the system's context awareness) has to be more advanced in order to be used over a longer time at users' homes. During the KSERA project the environmental information was partly abstracted (e.g. based on rules the PM10 / PM2,5 values were classified in different classes – related to possible outdoor activities). However this abstraction is up to now limited by insufficient context awareness and is also a crucial point concerning safety aspects.

Medical alert and emergency support

The approach of using a mobile interface combined with sensors and a context aware based inference system as a tool for medical alert was presented to trial participants representing primary users and secondary users. The system could be used as remotely navigated video communication system and/or be triggered by abnormal situations.

Advantages. The approach of having a mobile video communication system in case of medical alert or critical situations was appreciated by all trial participant groups. The most advantageous aspect was the possibility of having a solution that can approach towards the user, who is in a critical situation and can open a communication channel (sound, video) in order to get a feedback concerning the user's condition and to give direct support. The autonomous mode of the robot as well as a remotely steered version was appreciated by medical experts. A combination of a robot, which is aware of the position of the primary user and approaches to it and a remotely steered version, which can be steered in the right position directly at the person was described as most beneficial.

Limitations. Comparable to the limitations in the usage as communication tool for daily life the speed, size and robustness of the solution were criticized by the experts.

Physical Therapy Support

During the KSERA trials the robot was used as a robot personal trainer, with the purpose to motivate the primary users to do their exercises regularly.

Advantages. The approach using a humanoid robot was described as beneficial by the group of primary users and medical and therapy experts. Compared to other robot solutions and graphical presented avatars the different training movements can be presented in a more natural way and can easily be understood and followed. A humanoid robot and especially a socially assistive robot seemed to be a very good solution for this use case of IT supported physical training in older people's homes. The degrees of freedom and the presence in the area of interest and training offered the possibility of a motivated training by the users. The combination of physical training based on physical parameter measurements and environmental data was described as enrichment.

Limitations. During the KSERA project there was no solution implemented in order to analyze the movements of the person doing excercises. This analysis was described as a very important feature by primary and secondary user groups and definitely limits the solution, if not implemented. Another feature stated by medical experts was the possibility to record training data and abstracted parameters (e.g. activity protocols)

3.7. Effect of a long-term interaction on the acceptance parameters and the general impression of the approach

General description

Chapter 3.4 describes the general acceptance influencing parameters measured with quantitative and qualitative metrics. In the following chapter we briefly discuss the influence of regular interaction with the system on these descriptive parameters. In addition we present some results based on the quantitative and qualitative data gathered during the trials, which are related to the general KSERA approach and the effect on a **future long term application** in (vulnerable) users' homes.

Setup

The experiment included 6 iterations. The integrated KSERA system was used with all given functionalities with the robot as main interacting user interface.

Subjects

There were 4 participants involved in the experiment. Their age ranged from 70 to 91, all were female. The participants (n=2) in Austria lived independently in apartments in Schwechat. All participants had COPD Gold Level 0-1. The trial participants in Israel (n=2) had COPD level 0.

Procedure

In the first iteration all functionalities and possibilities of the robot were presented to the trial participants by simulating a common sequence of usage scenarios (e.g.: getting up and asking for information, breakfast, the robot remembering of taking a measure, training after the breakfast, etc.). Before every sequence and interaction a member of the research team explained the following scenario and the usage idea. This method of interactive explanation was used in order to present the different implemented functionalities in an understandable way and to give the older trial participants the chance to implement the approach in their model of daily living. This was done in order to get a common base of understanding the system approach – which is base for the evaluation of ITU and related acceptance parameters.

In the second iteration the trial participants had the chance to interact with the system without any proposed test flow – all implemented use cases were explained in short again and a paper with speech commands to trigger the different functionalities was handed out to the participants. The only order was to perform at least 3 interactions with the robot.

In the long-term iterations a sub-set of trial participants interacted with the robot over a period of one month in iterative sessions (2-3 times a week). Therefore we chose the use case of robot supported physical training. The trial participants were invited to enter the room and interact with the robot.

During the trial sessions the scenes were filmed and observed by members of the research team – the comments given by the trial participants (thinking aloud) were protocolled. In the end of each iteration construct based questionnaires were completed and a general interview about the experiences was performed in order to get a more qualitative feedback.

Important: As described in chapter 3.3 and 3.4 the technical performance of the system was a significant influencing factor on the acceptance parameters. In order to achieve comparability and to present meaningful results concerning the effect of the long-term interaction, we excluded trial results which were influenced significantly by the technical system. As criterion we defined, that all test cases during the trial had to work. Only if minor problems occurred (e.g. a longer search time for the face or not a direct navigation to the trial participant) the test case was defined as successful. In the other cases the data was excluded.

Methods used

In the scope of this experiment we gathered quantitative and qualitative data using subjective and objective methods.

Results

Quantitative results concerning the ITU influencing parameters

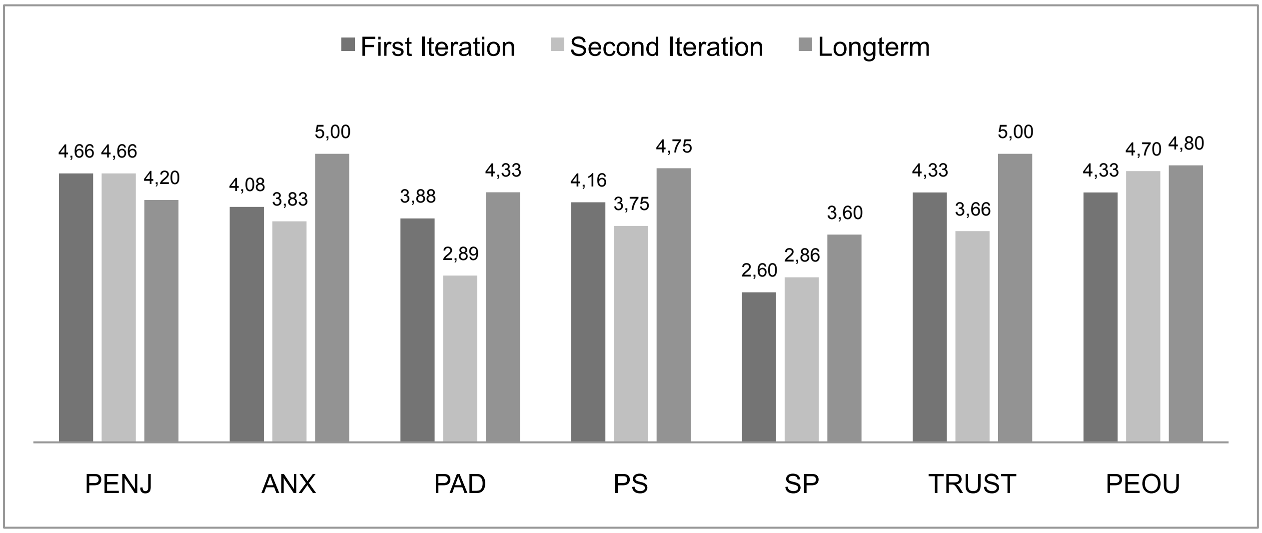


Figure 22: Results of acceptance parameters – long-term trial participants

Figure 22 gives an overview of the acceptance parameters related to the Almere evaluation model (see also chapter 3.4) of the long-term trial participants (n=3). It shows that the influencing factors (except for the factor for perceived enjoyment) have a positive trend. The negative trend of the PENJ was expected, as the trial participants were part of the same usage scenario (medical measurement and physical training) during 6 iterations. Factors like TRUST and PEOU were expected to rise (preconditioned a working system, without severe and obvious technical problems). The factor for social presence SP and perceived sociability PS both showed a positive trend, which approved the approach of including several strategies to develop a social robot (HRI and communication strategies, verbal and non-verbal cues, emotion representation, etc. – see also chapter 3.4). In the following we will present some results based on the quantitative and qualitative data gathered during the trials, which are related to the general KSERA approach and the effect on a future long term application in (vulnerable) users' homes.

General impression of the approach

Multimodal Interaction.

During the trials it was evident, that during the first iterations the effect of getting to know the new system and its abilities was in the focus of interest of the trial participants. The possibilities to **communicate** with the system were limited and therefore also the interaction (HRI) was limited. One way to encourage longer interactions with the robot may be to make the verbal communication more interactive. In comparison to the first trial during the KSERA formative evaluation we have implemented a more complex dialogue system. Nevertheless the speech and word recognition turned again out to be the crucial point. As described in chapter EX1 the performance and related recognition rates were low and dependent on distance, location, surroundings and subjective parameters (e.g. trial participants' talking ability). A dialogue system based on speech recognition therefore has first to overcome the actual state of the art limitations. Those multimodal interaction abilities are presented in parameters e.g. SP, PS. TRUST, PEOU.

Related to this also the presentation of **emotions** is an important feature if the SAR is to act human-like, it should respond to events in an emotional manner. For example the robot could become visibly happy when greeting a person who interacts with it—or annoyed if that person is not. During the KSERA project we implemented emotional cues in the prototype. The purpose of these cues was to support the interaction in a multimodal way without being too importunate. The combination of gestures, verbal communication and emotional cues was described as discreet and self explaining. Therefore the approach of using a humanoid robot, which can present gestures, emotions and verbal feedback is promising as it offers a possibility for a self-explaining user interface, that can easily be implemented in the future users mental models – especially of older users.

The possibility of having **several feedback channels** in parallel (e.g. sound, vision) was perceived as promising, as on the one hand it can support the communication and interaction with users with sensory deficits (e.g. sight, hearing impairments) and on the other hand be helpful in situations, where the users' focus is not set to the robot. In such situations this multimodal interaction is a very familiar interaction strategy and easy to understand by the users.

Another important feature for long term interaction of a human with a robot is the **person identification and personalization**. Only if the robot can recognize where the person is and who the person is, a reasonable interaction and trust in the system will be achieved in a future use in people's homes. During the KSERA project a localization of the trial participants in combination with face recognition was implemented. This direct search for the user's face and the directed communication to the users was perceived as natural and human-like. However often disproportional long times for finding the users' faces occurred (t>30 s). This was perceived as too long. In many cases the users did not know what was happening.

In future projects face recognition with person identification is recommended. It would offer a personalization of the robot and the system. Also the head pose estimation has to be advanced in order to achieve more context awareness, which will have a direct effect on the system abilities and the impact of the system.

To establish long-term relationships, the robot has not only to identify but also "get to know" people who frequently work with it. This cognitive ability was not implemented during the KSERA project. However this ability is a key element of human-human interactions. If the robot can learn about a person's interests (such as storylines that they prefer, information they like, exercises they prefer), personalizing the interactions will perhaps make the experience more enjoyable (also over a longer period of time).

One of the main statements during the discussions and interviews with the trial participants was, that long-term human-robot social interaction can be greatly improved through an understanding of human-human social interaction. During the KSERA project it was only partly possible to implement human-human communication strategies (e.g. multimodal communication). Especially the robot-perception abilities were limited. In the following some recommendations are presented. These recommendations are related to the KSERA use cases, but can be seen as common rules for enhancing the social abilities of the robot and will support a motivated long-term use of the solution.

Introduction and Greeting. Greeting should be used to make the robot engaging, to shape expectations for the ease of interacting with it. A personalized greeting would be even better. In this relation person identification would be a great benefit.

Communication and Dialog. A diversified dialog based on a robust dialogue system would guarantee a meaningful communication and long term trust. The robot would be perceived as socially present and adaptive towards several situations. If this would be achieved, the robot would not be recognized as a “boring” system and therefore be used more often and frequently, which would be a great benefit for systems aiming at daily life support and disease management. Beside this the purpose of the action and interaction always has to be clear for the users. Misunderstanding can lead to not using the solution at all.

Communication Pause and Departure. The robot should have the ability/ mechanisms to end up the interaction in an appropriate way (based on the local/cultural social norms). The robot should recognize, when the focus of the person is set to other activities or the person leaves the interaction area. This would make the robot a discreet companion.

Limitations concerning the analysis results

Certain technical limitations exist in the prototype 2 (PT2) system that had to be circumvented during the trials to allow meaningful results concerning the gathered user experience data. The following points changed the results of the technical performance evaluation in a positive way and hence limit the significance of certain parts of the technical evaluation.

- Light conditions were set to be as static as possible. Artificial lights were used, external lights were omitted by closing the curtains and using window blends. → This Influences the localization, navigation and face detection results positively
- The user was kept active during the navigational part either by giving small tasks, such as „please can you take the pulsoximeter“ or „take a look in the newspaper“ in order to allow the technical system to locate the user by the user’s movements. This was necessary to conduct the trial but influenced the technical performance of the navigation component in a positive way when compared to a real-life situation → influences the user localization, navigation positively
- User and robot positions were carefully chosen to avoid obstacles and allow for a simple path from the robot’s starting position to the desired end location → influences the navigation positively
- Key input was used instead of word recognition → this did NOT influence the recognition rate of the word detection since this measurement was done before the key input and the value was taken from the system logs

4. Summary and Conclusions

In the following we present a condensed view of the conclusions. In addition we present a table 4, comparing facts /state of the art, results of the KSERA trials and hints for related projects.

4.1. Performance

High efforts were undertaken to make the system more robust after the formative evaluation phase. Given the higher complexity of the second prototype system it becomes clear that the efforts undertaken were successful in most parts since the technical performance did not lower significantly for most domains. Still the goal of higher reliability of the second prototype was not reached. This is to most extend due to general technical issues that are not sufficiently solved so far in research in general, but would be needed for a successful integration of robotic solutions into smart home environments.

General technical issues of robotic prototypes in smart homes

Navigation. This very complex AI issue could only partly be solved. Navigating a biped robot inside an unmodified living environment is a current research challenge that is not sufficiently solved so far in general.

Speech recognition. Current best performing speech recognition engines such as Apples SIRI, Google Voice or Dragon speech where not able to sufficiently allow a dialog over a distance of 2 meters with the Nao robot. Since this is a research topic that needs resources far larger that the project KSERA is able to provide, the speech recognition was simulated in all trials using Wizard of Oz techniques.

Autonomy of the robot. The used robot Nao has several issues that influence the autonomy negatively. Firstly the maximum time of operation on battery is only 45 minutes. This was sufficient to conduct the trials, but an automatic recharging mechanism is necessary to keep the robot active over longer periods such as in a field-trial. This maximum time also limits the range of operation since the robot is also limited in speed.

The operation system of the robot is provided by Aldebaran and still under development and hence also prone to errors. Technical supervision is constantly necessary to operate the robot because of possible malfunctions of the operation system.

The used motors and heat dissipation system do not allow the robot to be operated over long time periods since the robot can easily overheat.

Regarding the question of technical influence on the perception of the users. It could be shown that there is an influence of technical malfunction to acceptance parameters, but it is not completely clear to what extend. Additional trials are planned to validate the results in a more controlled test setting for comparison, e.g. using the wizard-of-oz technique to overcome the navigational problems.

Possibilities and Limitations of the prototype. The actual KSERA prototype was described as a good tool for performing first trials and for evaluating the use cases in lab and living lab situations. In order to test the use cases at the older people's homes a more robust prototype has to be implemented. The humanoid, biped approach was appreciated for use case related to mimicking or human-like expressions. On the other hand the experts stated that the biped approach state of the art is not robust enough up to now, to use it in real life conditions without permanent external observation by researchers.

4.2. Acceptance

The KSERA approach was evaluated with high levels in several constructs that give feedback regarding an intended future use (ITU). The results and construct interrelations can be seen in Figure 23. The acceptance parameters however are very dependent on the performance and robustness of the prototype. The described results therefore are based on trials, where the prototype worked well. Data sets with malfunction were excluded.

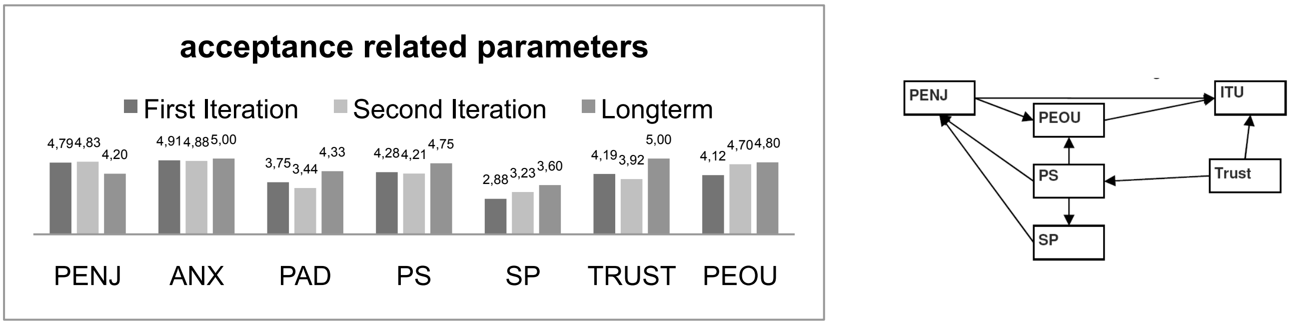


Figure 23: Acceptance constructs and Almere Model for evaluation of socially assistive robots

In addition to the quantitative results, trial participants (primary and secondary users) gave feedback concerning the different acceptance domains. All perceived the robot as easy to use and enjoyable system. They also trusted the system, as it seemed to be an intelligent system with inference abilities. The system also was perceived as socially present and sociable.

4.3. The KSERA approach as a socially assistive robot

One of the main aims in KSERA was to implement social abilities in the robot. The aforementioned evaluation factors also give feedback about the proposed approach. Both the factors for social presence SP and perceived sociability PS give a feedback on how accepted the KSERA approach was to implement characteristics in the robot, that make the robot a social robot – a robot that the people can integrate in their social and mental model – a robot that the trial participants can understand and interact with. In the KSERA approach the social robot included

- communication by questioning and answering, e.g. providing hints for several use case purposes
- expression of emotions by facial expressions (e.g. LED colors, sounds or speech, gestures)
- usage of natural interaction cues (face recognition, face direction recognition, gestures)
- exhibition of a defined personality (“the friendly helper in daily disease management”)

The KSERA approach of using a humanoid robot with the mentioned social abilities was appreciated by the trial participants. The factors for PS and SP are located in value-areas comparable to related studies e.g. (Heerink2010, Zhang2003). The combination of a social robot together with the proposed KSERA abilities of disease management, entertainment, communication and home control resulted in the impression of a social and intelligent helper.

Nevertheless the trial participants also recognized the deficits of the system (e.g. speed, robustness, performance, adaptability – see also performance evaluation). However this had only minor effects on the mentioned impression concerning the sociability. Therefore the KSERA approach can be seen as promising and was accepted by the trial participants.

In future projects the dialog system should be optimized and extended. The communication skills were described as “adequate”, however the trial participants noted the limited adaptability in the conversation. The combination of verbal and non-verbal interaction cues and the multimodal conversation strategies facilitated the engagement and transfer of information.

KSERA used a SAR approach instead of an often used service robot approach. Beside the social role – which was described in the previous passage a SAR also has to fulfill an assistive role. This assistance can be assistance in daily life activities, assistance with mental or physical deficits, assistance with communication tasks and assistance in carrying things. In KSERA the main focus was on

- assistance for disease management tasks (motivation for physical parameter measurements, information provision, motivation for physical training),
- communication (in case of emergency towards a call center, for contacting friends or relatives) and
- daily life activities (home automation support).

The described parameters of PEOU, TRUST, SP and PS provide a good feedback concerning the acceptance of the approach as an assistive robot (whereas assistance is related to the described domains). In addition to the already mentioned parameters of TRUST, SP and PS, the perceived ease of use (PEOU), which is influenced by several usability factors (e.g. effectiveness, efficiency, satisfaction, learnability, memorability, error rates) is identified to play a crucial role. In the KSERA end evaluation the PEOU was rated with mean levels of mean=4.06, std=0.7 during the first iteration, mean=4.70, std=0.33 after the second iteration and mean 4.80 during the long-term iterations (all on a 5-point likert scale).

Nevertheless these results are based on a working system without major malfunctions (as it will be part of future user homes). The PEOU is influenced by malfunctions and low technical performance parameters of a large extend. Therefore the robustness of the solution was described as most important. Related to this the solution of using a biped robot for assistive tasks - given the state of the art of robot robustness – cannot be confirmed in general. A discussion on this topic can also be found in chapters 3.5.1, where the general approach and added value is discussed in a more holistic way.

4.4. Added value and usage domains

Disease Management. The combination of a social robot and the proposed KSERA abilities of disease management (e.g. motivating the user to measure physical parameters, gathering and presentation of environmental data, supported physical training) were appreciated by all user groups. The presented use cases were seen as beneficial by secondary users related to healthcare.

Communication and Social Interaction. The possibility of introducing a mobile system for video communication for social and consulting as well as emergency purposes was appreciated by primary and secondary users. Beside the video communication using the mobile projector also the advantage of different approaches (e.g. static monitor, which is triggered by the system) was emphasized by the trial participants. The robot should however be the central guide and motivator to use the other input/output channels. The mobile user interface was seen as an advantage for vulnerable groups e.g. people with COPD or older people with decreased mobility.

Daily life support. In the course of the KSERA end evaluation the possibilities for support in daily life were evaluated. The added value of introducing a mobile interface, that can be steered in a very natural way (e.g. via speech input) was seen as a big advantage for vulnerable users. The idea of using one interface for many, different tasks was appreciated. Nevertheless the added value is dependent on the robustness and performance, as well as a smooth integration of several modules (speech recognition, navigation, localization, etc.).

The approach of using a socially assistive robot for daily life support was on the one hand seen as a fruitful approach, however a lot of support use cases proposed by the trial participants can only be realized using a service robot (e.g. carrying certain equipment).

The combination of a robotic interface and a smart home was seen as a real advantage for user groups like in KSERA, as limitations in both concepts could be lessened. For example several user interfaces in smart homes can be replaced by one mobile multipurpose interface. The robotic solution can be extended by an enhanced context awareness enriched by data of smart home sensors, external sensors and virtual sensors.

4.5. Condensed result table

The following table gives an overview of lessons learned during the formative evaluation, including a brief discussion on the different topics and outlook on the future prototypes in the KSERA project. Some of the points are KSERA specific, but the included content is summarized in order to give input for related projects. The results from the QoL and HRI evaluation study is therefore combined with common known facts and the overall KSERA performance. The table has a strong focus on the SAR and related performance parameters.

Evaluation Domain	Facts concerning the general robotic approach	KSERA evaluation results	General discussion and hints based on experience of the end evaluation
Appearance of the SAR	• *Related studies (e.g. Duffy 2003, MacDorman2006, Waigner2006) show, that physical (and in particular humanoid) robots are preferred over virtual agents and text descriptions* • *The appearance should match the task and role of the robot in daily life (Waigner2006)* • *The appearance is not as much important as the function*	• *The results of the end evaluation showed, that the humanoid robot was seen as potential tool for assistive tasks (in terms of assisting as a socially assistive robot and not a service robot), which was trusted to give reasonable feedback using different interaction channels. The appearance was evaluated very positively.* • *The humanoid robotic approach (e.g. for exercise instructions) was seen very positively, also in comparison with text or avatar based systems. This is also related to the humanoid appearance and direct identification with the embodiment and implementation in the trial participants' mental model of the different use cases and use scenarios.* • *Related to the humanoid appearance of the robot, the trial participants also had the expectancy, that the robot had human abilities (e.g. understanding, communication abilities, humanoid movement – e.g. concerning speed, etc.)*	• *The main capabilities of a SAR – including the most important motivating and communicating skills – were analyzed and evaluated in several lab studies (e.g. Torta2010, DeWitt2012) and implemented in the second KSERA prototype. The multimodal capabilities of the KSERA robot solution NAO were extensively used and appreciated by the different trial participant groups.* • *Deploying a humanoid robot approach also raises high expectations towards the whole solution. The users expect human abilities by the robot concerning different domains of interaction and capability (multimodal expression, context awareness, intelligence, navigation and movement, etc.). However state of the art solutions do not cover all those expectations concerning robustness and adaptability.* • *The robotic solution used in the KSERA project did not meet the expectations, most of the secondary, but also primary users had. Although the approach of using a socially assistive robot was clear for the trial participants, they expected a more robust solution.*
Interaction and communication	• *Robots should provide humanlike communication and interaction skills – using multimodal communication possibilities (Kobayashi2008, Leite2008, Li2006).* • *The robotic (and overall) systems should also use additional channels for communication (screens, beamers, etc.)*	• *The multimodal expression abilities (talking, gestures, movements, eye-movements,…) were appreciated by the trial participants and also expected, due to the humanoid appearance of the robot. The limited number and lack of performance of the input channels (e.g. word recognition, face detection, etc.) was obvious to the participants and also recognized as a lack of the system.*	• *The context awareness and performance concerning the different input channel abilities (in particular efficiency and effectiveness of the word recognition) have to be enhanced compared to the KSERA performance in order to realize longer, more realistic and adaptable communication flows and interaction cues. This will lead to a system that will be perceived as more humanoid, which has clear effects on a possible future use in daily life.*

		• *The advantage of different approaches (e.g. static monitor, which is triggered by the system) was emphasized by the trial participants. The robot should however be the central guide and motivator to use the other input/output channels*	• *The abilities of a humanoid robot being a helpful motivator should be emphasized for future prototypes. This implements enhanced demonstration of (non-functional) humanlike gestures (e.g. hand-movements) and mimic-like expressions (e.g. simulated glimpsing, etc.)* • *If adequate and robust working capabilities are not possible to implement, the actions and action flow of the system should be designed in a way to match the situation (e.g. only one answer is possible, or questions will be repeated using the same question statement explained with different sentences) or even more restrict the context of use (e.g. have a clear flow of warnings/ hints without the need of user input in critical situations)* • *Express more what will happen – the robot has to be more communicative – It will be of great importance to implement clear feedback loops concerning the status and next steps of interaction – again these can be expressed by voice output, gestures or mimics. It has to be clear to the user, what the robot is performing and what is the robot's intention.* • *Provide cues for long-term (or longer scenario-) tests.* • *The adaptability and personalization has to be enhanced compared to the KSERA performance. It was emphasized by the trial participants, that the robot has to recognize with whom it is interacting.*
Movement and Navigation	• *Robots should provide humanlike navigation strategies and movement patterns.* • *Robots should be fast enough to fulfill their purpose in time*	• *The possibility of the robot to move towards the user was appreciated and also wished to be performed during emergency situations* • *The KSERA system had to deal with a couple of limitations in the real life situation (lights, shades, movements of second user, etc.). This sometimes had strong effects on*	• *The stability concerning user and robot position detection has to be enhanced in order to have a stable base for robot navigation towards different places (in front of the user, in front of a wall, in front of different objects).* • *Monomodal sensor approaches should be replaced / enlarged by multimodal approaches, in order to meet real life requirements (e.g. different*

		the navigation performance – which were also obvious to the user. During navigation tasks with only one user in the room the performance reached average values. • *The speed of the robot and the time for the different movement distances varied too much (however e.g. the step speed is mostly depended on the robotic platform) and the speed was evaluated as too slow.*	*sensor information is needed in different rooms or parts of rooms e.g. if a ceiling camera picture is not available or if sonar sensor information is misleading due to obstacles, corners or wall and floor characteristics.* • *Speed and movement naturalness could be enhanced in order to enhance the score of animacy and humanoid perception. Those parameters will have a direct correlation to the overall acceptance of the system (Duffy 2003).* • *Cues for longer scenarios (error resistant and more adaptive cues) have to be implemented in order to meet real life requirements.* • *The SAR has to perform movements and navigation in a humanlike way (concerning direction, speed and movement performance). This performance will have a direct effect on the perceived system abilities and perceived animacy and intelligence of the system.*
Support and Social Skills	• *Robots should provide humanlike skills in order to give multimodal support (give oral hints, walk to areas of interest, show movements, etc.) (Andre2004a, Andre2004b, Duffy 2003, Li2006)* • *Video communication possibilities are an enhancement in older users' daily life and offer a good opportunity in closing the digital divide.*	• *The KSERA system approach of giving multimodal hints (e.g. for monitoring special physiological parameters) or warnings was very appreciated during the trials. Also the reminding possibilities.* • *The support hints were too complex and long for the users, they would have liked shorter answers and hints, what they should do – knowing the facts (e.g. air pollution – is it possible to go for a walk?). The second prototype already included a more abstract presentation of the diverse measurements. However the trial participants expected the system to be more adaptive towards the different users and day to day situations. The implementation of a medical expert system implemented in the KSERA disease management system in combination with the feedback of e.g. call*	• *Appropriate conversational, multimodal strategies have to be emphasized.* • *Adaption towards the users' skills have to be implemented and will be an enhancement in the perception of the system's abilities.* • *Video communication should be part of the future prototypes and related approaches. Different strategies (e.g. mobile solutions using the beamer pack on the robot, using static touch screen solutions as used in smart homes and alternatives) can be used, deploying the SAR's abilities as a mobile interface and motivator, tutor, etc.*

		center agents via the KSERA video communication was appreciated and was seen as an adequate addition to the system based feedback. • *The time of reaction of the system was adequate for the tested use cases.* • *The video communication possibility was appreciated by both primary and secondary users.*	
Motivation Skills	• *A humanoid robot is a powerful tool for performing motivational tasks – deploying humanlike strategies (Kobayashi2008, Kozima2003, Leite2008, Li2006, Sidner2004).* • *Multimodal approaches show a very good performance in motivating older people for different ADLs (e.g as reminder, tutor, trainer, etc.) (Press2004, Johnson2004, Leite2008).*	• *The KSERA abilities for motivating to perform different tasks (physical parameter monitoring for self disease management, training exercise performance, critical dealing with situations) showed very good evaluation results.* • *The trial participants interacted and cooperated with the robotic system and the defined aims (e.g. task oriented training, measurements, etc.) were performed by the users.* • *The trial participants gave a very positive feedback concerning the perceived intelligence and politeness, the sympathy towards their robotic companion and the trust towards the system.*	• *Above described factors concerning HRI and acceptance of the solution could enhance the motivational performance in future prototypes.* • *The direct activation by a mobile agent and especially a humanoid SAR enhances and extends common known motivation strategies (verbal, avatar-based, film, etc.) (Press2004, Johnson2004, Looije2006)*
High Level Skills **(trust, empathy)**	• *Robots (in particular humanoid, social agents) should trigger the right user behaviors by raising trust and empathy* • *Special dialog strategies lead to optimized task performance flows (Kobayashi2008, Leite2008)*	• *The robot was seen as a sympathetic motivator and helper, whom the trial participants would trust in daily life.* • *During this test phase different dialog strategies were not yet implemented*	• *Above described factors of interaction could enhance the score for high level skills in future prototypes. A special focus should be set on longer interaction and communication cues (Lim2009, Looije2009, Looije2006, Rumpler2009), in order to improve the perceived animacy intelligence and anthropomorphism.*
Personalisation and Adaption	• *Robots will be very personal , long term companions in future, therefore strategies for maintaining relationships have to be implemented*	• *KSERA does not provide such strategies by now*	

table 4: Condensed table of results in the different evaluation domains

5. Literature references

[Andre2004a] E. André, L. Dybkjær, W. Minker, and P. Heisterkamp (eds.), Affective Dialogue Systems, Tutorial and Research

Workshop (ADS 2004), KlosterIrsee, Germany, June 14-16, 2004, Proceedings, Lecture Notes in Computer Science, Volume 3068, Springer Berlin Heidelbery, 2004.

[Andre2004b] E. André, M. Rehm, W. Minker, and D. Bühler, Endowing Spoken Language Dialogue Systems with Emotional

Intelligence, in (André et al. 2004), pp. 178-187, 2004.

[Argyle1988] Argyle, M. Bodily Communication. Methuen, New York, NY, USA, 1988.

[Bartneck2003] C. Bartneck, Interacting with an Embodied Emotional Character, in J. Forlizzi, and B. Hanington (eds.), Proceedings of the 2003 international conference on Designing pleasurable products and interfaces, ACM Press, New York, NY, pp. 55-60, 2003.

[Bartneck2006] C. Bartneck, Measuring the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety

of robots, 2006

[Brickmore2005] T. W. Bickmore and R. W. Picard, Establishing and Maintaining Long-Term Human-Computer Relationships, ACM

Transactions on Computer Human Interaction, 59(1):21-30, 2005.

[Castellano2009] G. Castellano, I. Leite, A. Pereira, C. Martinho, A. Paiva, and P. W. McOwan, It's All in the Game: Towards an Affect

Sensitive and Context Aware Game Companion, in (Cohn et al., 2009) pp. 29-36, 2009.

[Davis 1989] Davis FD, Bagozzi RP & Warshaw PR (1989) User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science, 35 (8): 982-1003

[DeWitt2012] Dirk de Wit, Maurice Spiegels, Larn van de Vries, Alexandros Mileounis and Michail Pavlakis, Games and playful interaction to support Senior-robot interaction, HTI Design Track A, TU Eindhoven 2012

[Duffy 2003] B. R. Duffy, Anthropomorphism and the social robot, Robotics and Autonomous Systems, 42(3-4):177-190, 2003.

[Forlizzi2007] Forlizzi J (2007) How robotic products become social products:an ethnographic study of cleaning in the home. In: ACMSIGCHI/SIGART human-robot interaction, pp 129–136

[Hall2005] Hall, J.A., Coats, E.J. and BeLeau, L.S. Nonverbalbehavior and the vertical dimension of social relations. Psychological Bulletin, 131, 6 (2005), 898-924.

[Heerink2010] Marcel Heerink, Ben Kröse, Vanessa Evers, BobWielinga,Assessing Acceptance of Assistive Social Agent Technology by Older Adults: the Almere Model, Int J Soc Robot (2010) 2: 361–375

[Press2004] Press, Menlo Park, CA, pp. 36-43, 2004. J. Grolleman, B. van Dijk, A. Nijholt, and A. van Emst, Break the Habit!

Designing an e-Therapy Intervention Using a Virtual Coach in Aid of Smoking Cessation, in IJseelsteijn W. et al. (eds.)

[Johnson2004] D. Johnson, J. Gardner, and J. Wiles, Experience as a moderator of the media equation: the impact of flattery and

praise. International Journal of Human-Computer Studies 61(3):237-258, 2004.

[Koay2007] K. L. Koay., D. S. Syrdal, M. L. Walters, and K. Dautenhahn, Living with robots: Investigating the Habituation Effect in Participant's preferences during a longitudinal human-robot interaction study. Proceedings of the 16th IEEE International Conference on Robot and Human InteractiveCommunication (RO-MAN 2007), IEEE Press, Washington, D.C., pp. 564-569, 2007.

[Kobayashi2008] K. Kobayashi, K. Funakoshi, S. Yamada, M. Nakano, Y. Kitamura, and H. Tsujino, Smoothing humanrobot speech interaction with blinking-light expressions, in Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008), IEEE Press, Washington, D.C., pp. 47-52, 2008.

[Kozima2003] H. Kozima, C. Nakagawa, and H. Yano, Attention coupling as a prerequisite for social interaction, in Proc of the 12th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2003), IEEE Press, Washington, D.C., pp. 109-114, 2003.

[Leite2008] Leite, A. Pereira, C. Martinho, and A. Paiva, Are emotional robots more fun to play with? in Proc. of the17th IEEE International Sympoium on Robot and Human Interactive Communication (RO-MAN 2008), IEEE Press, Washington, D.C., pp. 77-82, 2008.

[Li2006] S. Li, B. Wrede, and G. Sagerer, A dialog system for comparative user studies on robot verbal behavior, in Proc. of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006), IEEE Press, Washington, D.C., pp. 129-134, 2006.

[Lim2009] M. Y. Lim, R. Aylett, W. C. Ho, S. Enz, and P. Vargas, A Socially-Aware Memory for Companion Agents, in (Ruttkay et al. 2009), pp. 20-26, 2009.

[Lombard1997] Lombard M, Ditton TB (1997) At the heart of it all: the concept of presence. J Comput-Mediat Commun 3(2). Available online: http://www.ascusc.org/jcmc/vol3/issue2/lombard.html

[Looije2006] R. Looije, F. Cnossen, and M. A. Neerincx, Incorporating guidelines for health assistance into a socially intelligent robot, in Proc. of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006), IEEE Press, Washington, D.C., pp. 515-520, 2006.

[Looije2009] R. Looije, M. A. Neerincx, and F. Cnossen, Persuasive robotic assistant for health self-management of FP7-231868 SERA deliverable 3.1 Version 4/05/2010 page 21 of 22 older adults: Design and evaluation of social behaviors, International Journal of Human-Computer Studies, vol. in Press, Corrected Proof, 2009.

[MacDorman2006] K. F. MacDorman, Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: an exploration of the uncanny valley, in Proceedings of the ICCS/CogSci-2006 Long Symposium:Toward Social Mechanisms of Android Science, pp. 26–29, 2006.

[Rumpler2009] M. Rumpler, EXSTASIS – An Extended Status Model for Social Interactions, in (Ruttkay et al., 2009), pp. 49-55, 2009.

[Schubert2005] Schubert, T.W. Your highness: Vertical positions as perceptual symbols of power. Journal of Personality and Social Psychology, 89, 1 (2005), 1-21. 17. Sterne, J. The Audible Past: Cultural Origins

[Shibata2003] Shibata T, Wada K, Tanie K (2003) Statistical analysis and comparisonof questionnaire results of subjective evaluations of sealrobot in Japan and UK. In: Proceedings IROS 2003

[Sidner2004] C. L. Sidner, C. D. Kidd, C. Lee, and N. Lesh, Where to look: a study of human-robot engagement, in Proceedings of

the 9th international conference on Intelligent user interfaces (IUI'04), ACM Press, FP7-231868 SERA deliverable 3.1 Version 4/05/2010 page 22 of 22 New York, NY, USA, pp. 78-84, 2004.

[Torta2012] *Elena Torta, Johannes Oberzaucher, Franz Werner, Raymond H. Cuijpers, James F. Juola, Attitudes Towards Socially Assistive Robots in Intelligent Homes: Results From Laboratory Studies and Field Trials*

[Venkatesr2000] *Venkatesh V & Davis FD (2000) A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science, 46 (2): 186-204*

[Venkatesr2006] *Venkatesh V, Ramesh V (2006) Web and Wireless Site Usability: Understanding Differences and Modelling Use. MIS Quarterly, 30 (1): 181-206*

[Venkatesr2003] *Venkatesh V, Morris MG, Davis GB & Davis FD (2003) User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly, 27 (3): 425-478*

[Waigner2006] *J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, The role of physical embodiment in humanrobot*

interaction, in Proc. of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006), IEEE Press, Washington, D.C., pp. 117-122, 2006.

[Walters2008] *M. L. Walters, D. S. Syrdal, K. Dautenhahn, R. te Boekhorst, K. L. Koay, Avoiding the uncanny valley: robot*

appearance, personality and consistency of behaviour in an attention-seeking home scenario for a robot companion. Autonomous Robots 24(2):159-178, 2008.

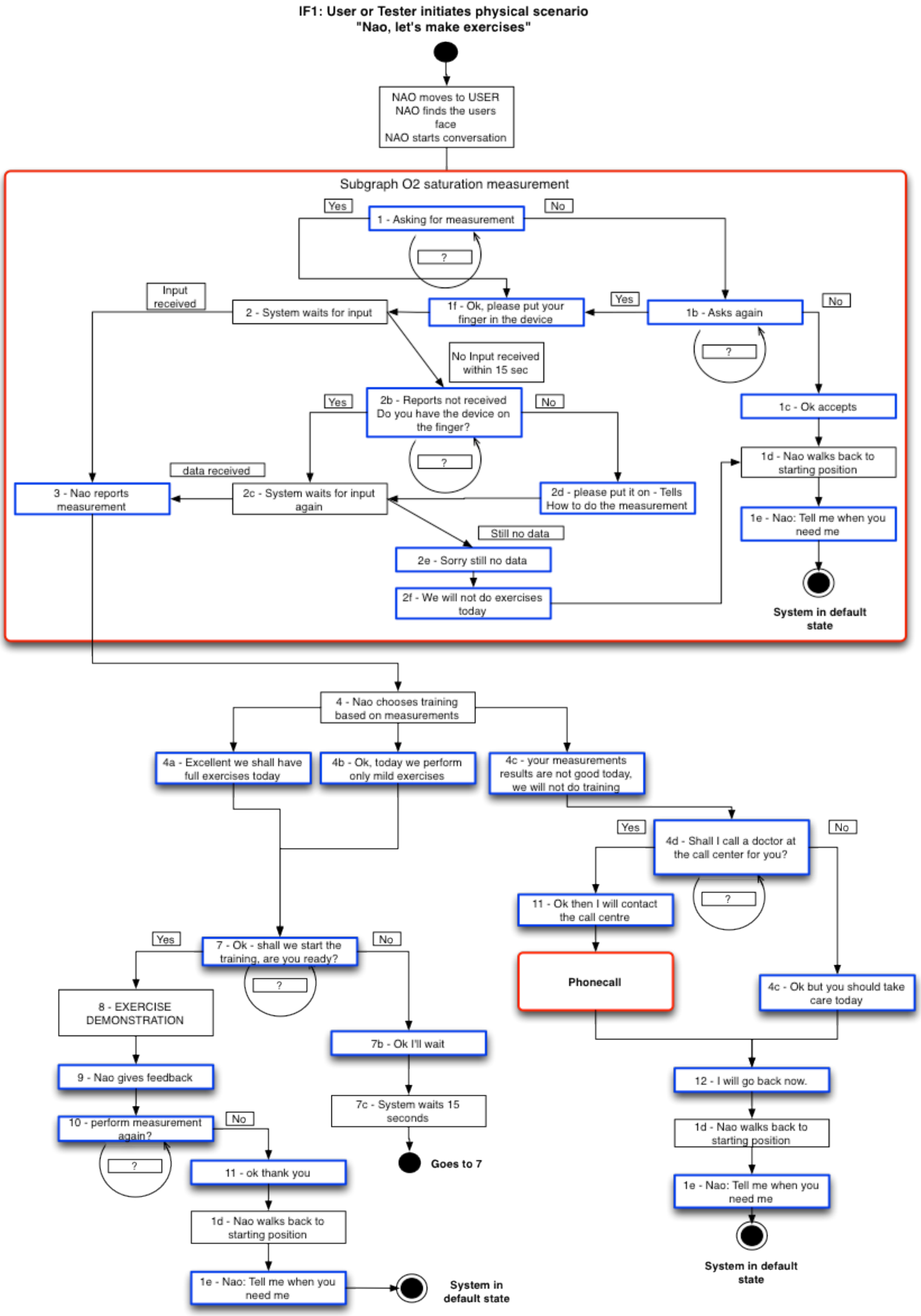
[WHOQoL1997] *WHOQOL-PROGRAMME ON MENTAL HEALTH, DIVISION OF MENTAL HEALTH AND PREVENTION OF SUBSTANCE ABUSE WORLD HEALTH ORGANIZATION, 1997*

[Zhang2006] *Sun H, Zhang P (2006) Causal relationships between perceived enjoyment and perceived ease of use: an alternative approach. J Assoc Inf Syst 7(9)*

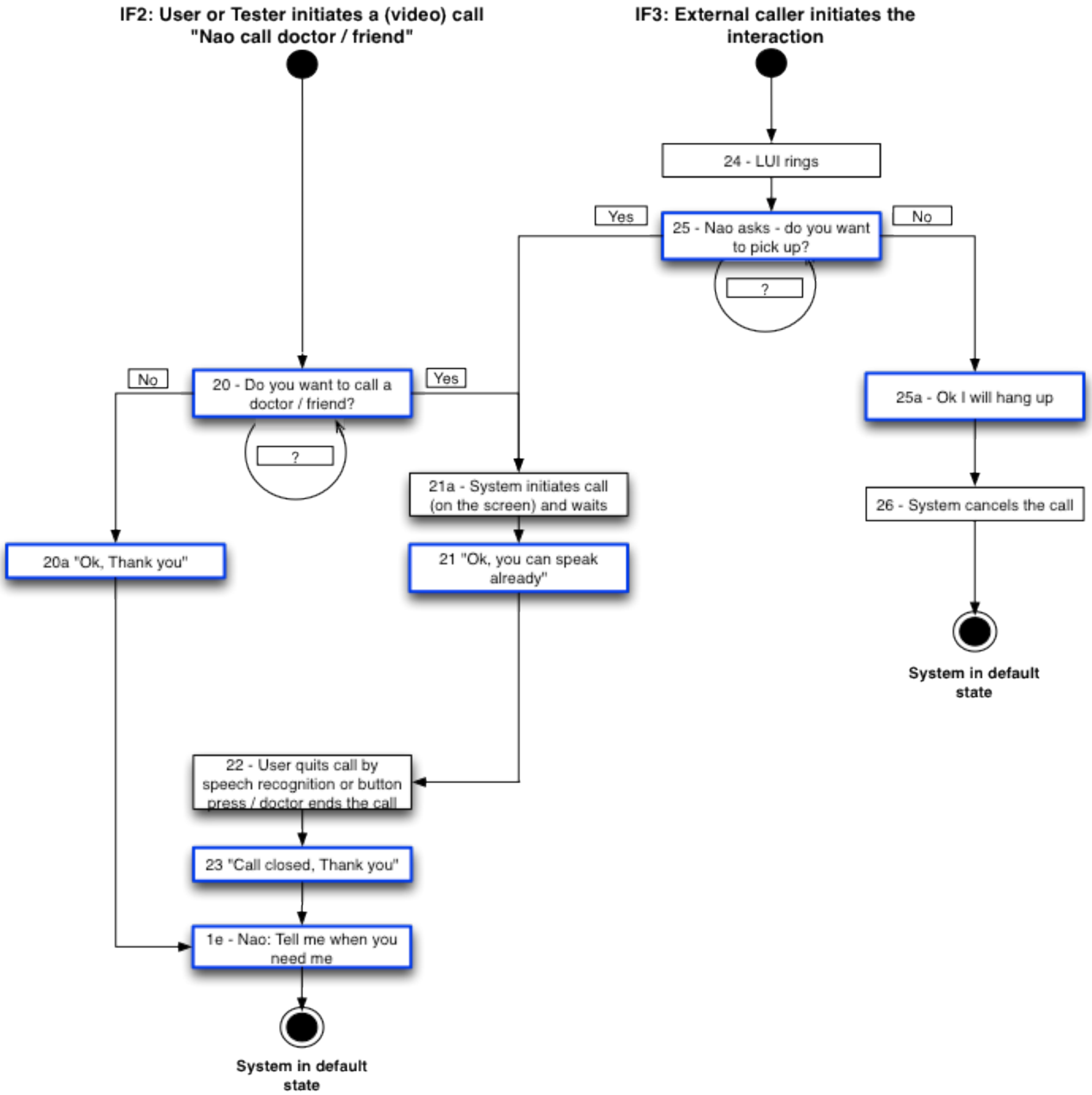
6. Annex

6.1. Used Interaction Flows

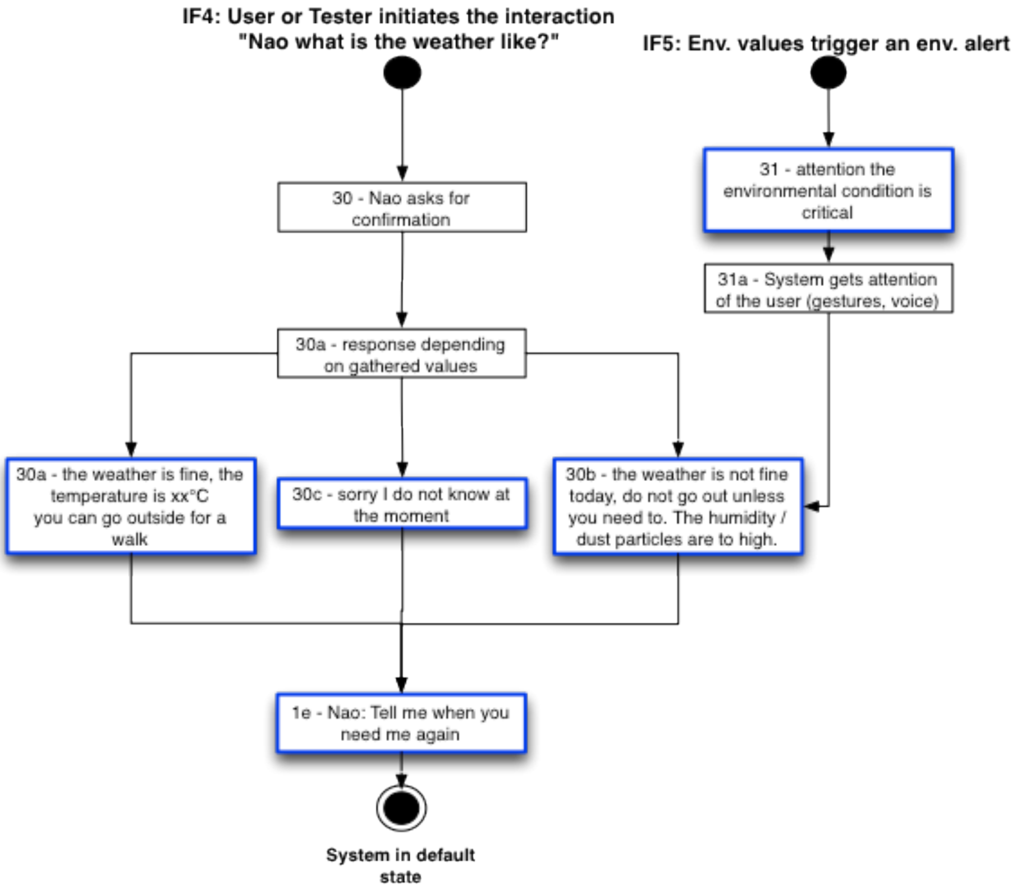
1. Interaction flow 1 – test case 1



Interaction flows 2/3 – test case 2



Interaction flows 4/5 – test case 3



Interaction flow 6 – test case 4