DR 1.3: Evaluation of PAL prototype supporting knowledge and awareness

Rosemarijn Looije, Francesca Sacchitelli, Olivier Blanson Henkemans, Mark Neerincx, Sylvia van der Pal, Francesco Sardu

TNO, The Netherlands; Fondazione Centro San Raffaele, Italy
⟨rosemarijn.looije@tno.nl⟩

Project, project Id:	EU H2020 PAL / PHC-643783
Project start date:	March 1 2015 (48 months)
Due date of deliverable:	M24
Actual submission date:	28 February 2017
Lead partner:	Fondazione Centro San Raffaele
Revision:	final
Dissemination level:	PU

This report presents the work WP1 with a focus on T1.4: The results from the first design and evaluation cycle. This cycle contained a number of related experiments (i.e., formative evaluations of PAL prototypes). Each experiment had specific research questions concerning PAL functionality and expected outcomes (claims) concerning (1) child's knowledge, awareness, attitude (towards PAL and T1DM), self-efficacy and skills (usage and adherence), or (2), HCP's trust and acceptance, or (3) parent's attitude and knowledge

1 Executive Summary

This report presents the work WP1 with a focus on T1.4: The results from the first design and evaluation cycle. This cycle contained a number of related experiments (i.e., formative evaluations of PAL prototypes). Each experiment had specific research questions concerning PAL functionality and expected outcomes (claims) concerning (1) child's knowledge, awareness, attitude (towards PAL and T1DM), self-efficacy and skills (usage and adherence), or (2), HCP's trust and acceptance, or (3) parent's attitude and knowledge (see Figure 1). Several new PAL functions were developed that were expected to enhance these three types of variables. In total, eight experiments were conducted during the second year (the "main evaluations" were conducted from April to June 2016, and during the month of August 2016). Via the following hyperlink `https://youtu.be/DlX6IfB2dtY`, the demonstrator video of the Cycle 1 prototype can be viewed.

Following the schedule of Figure 1. The first cycle focused on the following effects with the child: knowledge, awareness, attitude, self-efficacy and skills. For professionals we looked at trust and acceptance and with the parents we focused on attitude and knowledge. The situated Cognitive Engineering method [14] was extended with a formal specification of the situated Design Rationale [12], and applied to establish a theoretically and empirically grounded development cycle. This research and development cycle entailed formative analyses and assessments (e.g. interviews with the Health Care Professionals about the usability of the control panel) and more summative tests of the interactions of the child with the System.

Overall, the new PAL functions proved to support the three high-level objectives that were derived from the Self-Determination Theory: autonomy, competence and relatedness. More specifically, task 1.4 provided seven major outcomes. First, two "long-duration" experiments showed that playing PAL activities with the PAL actor in recurring sessions is associated with increased child's diabetes knowledge, awareness, attitude (i.e. relatedness), self-efficacy (i.e., autonomy and competence) and skills to use the acceptance and PAL system. Second, diary support and self-disclosure support seem to provide further enhancements of awareness (related to the timeline), attitude, self-efficacy and skills. Third, conversational fillers proof to enhance child's attitude (i.e., relatedness) to the PAL robot. Fourth, Health Care Professionals (HCPs) are positive about the opportunities of PAL, but the system needs substantial improvement for acceptance and trust (i.e., improvement of the usability and goal structure). Fifth, parents are also positive and prove to have a good attitude towards the PAL system and seem to experience already some knowledge gain. Sixth, different classes of child's usage have been identified; enhanced personalization is needed to establish adherence over these diverse user groups. Seventh, a number of usability bottlenecks have been identified that need to be solved to improve

	Initial Phase	Intermediate Phase	Final Phase
Determinants	**Knowledge & Awareness**	Diabetes Regimen Adherence	Shared Child-Caregiver Responsibility
Children - age dependent claims	7-10 yr: + knowledge + awareness + attitude + self-efficacy + skills	7-12 yr: + regime adherence + glucose monitoring	7-14 yr: + shared responsibility + coping with anomalies - hypos/ hypers + glycemic control
Usage of the system	1 month	4 months	9 months
Caregivers - claims	*Professionals:* + trust + acceptance *Parents:* + attitude + knowledge	*Professionals:* + awareness *Parents:* + trust + skills	*Professionals:* + tailoring *Parents:* + shared responsibility
Settings	hospital, home	hospital, home, camp	hospital, home, camp, elsewhere
mHealth apps	Timeline, Quiz game	diary, quizzes, sorting game, miniApps	diary, quizzes, sorting game, miniApps

Figure 1: Cycle 1 evaluation - First Phase of the project

the PAL usage.

This deliverable is organized as follows: Section 2 describes the role of this Work Package; Section 3 describes the work done so far and the purpose of the work to be done in the future, on the basis of the reviewer's comments; Section 4 describes the methodologies used during the second year of the project to achieve the results of the Cycle 1 Experiments; Section 5 describes the ontology Engineering; Sections 6 , 7 and 8 describe the results obtained from the experiments held in this second year of the project; Section 9 describes the refinements of the PAL System; and Section 10 describes the main conclusions.

2 Embedding in project

One of the main pillars of the PAL project is that end users are really involved in the project. This involvement concerns the acquisition of user needs, opinions, preferences, feedbacks and behavioral responses during focus groups and "real" human-PAL interactions. Especially by really using the System many requirements are elicited, much more so then when talking about it. These requirements are then used to improve the System within WP2-5.

2.1 Role WP

WP1 contributes to the development of the PAL System by making sure that theory and user inputs are used to design and evaluate a System in a methodological manner. Expected effects are related to the implemented functions and specific instruments are used to measure these.

3 Tasks, objectives, results

3.1 Planned work

This year Milestone 2 had the focus. To complete this milestone, the claims as listed in the "Initial Phase" column of Figure 1 had to be evaluated (i.e., the expected effects on children and caregivers in Cycle 1). We expect for instance an increase in knowledge of the children who use the PAL System. This Milestone was reached in M15, which is a 3 month delay from the original plan (caused by bottlenecks in the technical implementation and scheduling of participants). The overall project planning has been adjusted to deal with this delay and still meet the general objectives within the overall project duration.

4 Actual work performed

In Y2 we looked at the planned work, but also tried to take the reviewer comments into account:

1. More clarity in the relation between self-management and the design and evaluation of the System
2. More clarity on the methodological tools that were applied
3. More clarity on privacy protection

Comments 1 and 2 have been addressed with a procedure to work-out and assess *research questions* with the corresponding required functionality and test measurements, the definition of a *situated Design Rationale* (see also section 9, and the (incremental) construction of the PAL *ontology* described in section 5. The paper on the method supports the understanding of relations between the self-management objectives, and the design and evaluation of the system (see [12], and section 4). Figure 3 provides the visual overview of these relations. Section 7 summarizes the research protocol: By providing more information about the use of methodological tools, this protocol has been made more explicit. In general, the questionnaires have been filled out together with the children, and we noticed from experience that this ensured the understanding of the questions. So far, filling out the questionnaires was done on paper, we will do more digital questionnaires in the future. Parents filled out the questionnaires unsupervised.

Concerning the privacy protection, we have added extra technical measures so that the servers in Italy and the Netherlands are more secure. Moreover each partner signed a Material Transfer Agreement (MTA) at the beginning of the project in which is written what the partners can exchange to whom and how, starting by making sure that all the video/data exchanged are anonymous (then nobody among the researchers could know any information related to the participants). Furthermore a list of all the researchers who can have the access to those data was created. Every partner is, however, in charge to create a copy of those anonymous data on two different hard disks and all the paper files are locked in an archive. Next to this we made sure to have one person per site responsible for the identifiers. Each participant is numbered according to first evaluation he/she participated in and then participant number. For example first evaluation during summer camp Italy in 2016, second participant is 16SUMC02, this way we can keep track of children participating in multiple evaluations. And of course the protocols are evaluated by ethical committees.

4.1 Method

The situated Cognitive Engineering (sCE) method [15] supports incremental design and evaluation. Its main strength lies in the explicit analysis of the (theoretical, empirical and technical) *foundation* of *design specifications* and the *evaluation* methods for refinement and validation. To make sure that the functions of design specification contribute to the general objectives, sCE has been extended with the situated Design Rationale (sDR) method [12]. sDR explicitly relates the sCE concepts of the foundation, design specifications and evaluation to these objectives, methods and instruments (see Figure 2) in order to reason about the design decisions made. The sCE method did not yet support this type of reasoning. There were for instance no explicit relations between a specific method and therefore objective and a function. Of course use cases take the objectives into account, but the relation was not made explicit. Furthermore, the expected effects are related explicitly to the functions and the instruments, but the interrelations between expected effects and functions were not made explicit. One function can have multiple effects, an effect can be related to different functions, multiple instruments can be used to measure the same effect, but it can also happen that one instrument measures multiple effects. These relations need to be explicated so that we can disambiguate the design and evaluation as much as possible by refining it, e.g. more specificity in instruments. Disambiguation will not always be possible, but explicating all relations makes it possible to see where there are still ambiguous relations. Knowing these ambiguities can guide further design and evaluation.

Within the PAL project, we take the theory of Self Determination Theory (SDT) [7] and use the three objectives this theory describes: autonomy, competence and relatedness to guide the methods we use (mainly from behavior change, educational and game theories). PAL functionalities and the design of them are for a major part derived from these methods. The effects are based on the expected effects for the three main objectives next to other child-PAL interaction effects (e.g. usability). These effects are correlated with each other. Finally instruments are chosen to evaluate effects as specific as possible.

In [12] the sDR method is further explained. Figure 3 shows the sDR of the first cycle evaluation. This also makes the implemented functionalities of the first cycle more explicit.

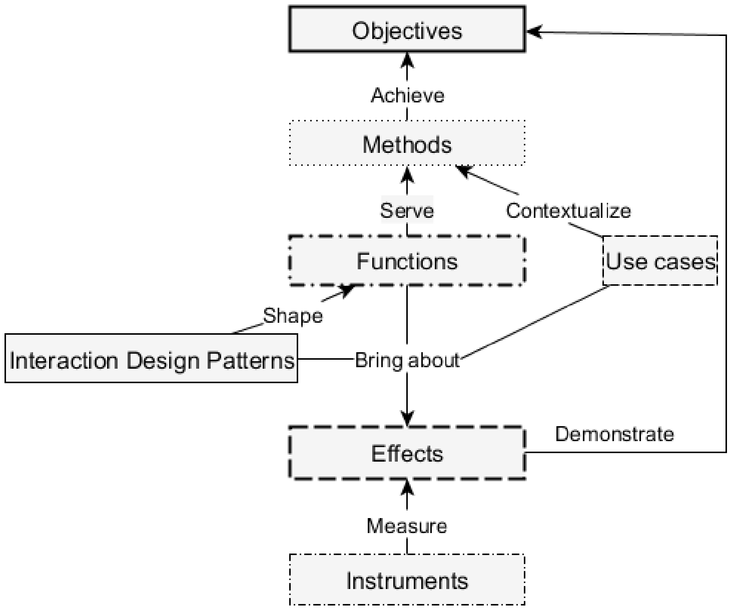


Figure 2: Generic concept map of the situated Design Rationale (sDR).

4.2 Evaluations

The first PAL development cycle contains a number of related experiments, which provide a coherent set of formative evaluations of PAL prototypes. Each experiment has specific research questions concerning PAL functionality and expected outcomes (claims). These research questions were derived from the description of work: (1) child's knowledge, awareness, attitude (towards PAL and T1DM), self-efficacy and skills (usage and adherence), (2), HCP's trust and acceptance, and (3) parent's attitude and knowledge (see Figure 1). Several new PAL functions were developed that were expected to enhance these three types of variables. Figure 3 provides an overview of these functions for the PAL system, focusing on the PAL modules for the children (i.e., the control and information panel for the caregivers are not part of this Figure).

In total, eight experiments were conducted during the second year. Table 1 provides an overview of these experiments. For the main research questions, the Table gives an indication of the general outcomes. As indicated in the Table, the following sections summarize the background and results.

5 Ontology Engineering

Engineering the PAL ontology is executed in a systematic, iterative, and incremental development process. Because PAL covers a large domain of interest, we have developed separate ontology models as high-level building blocks for smaller, more specific areas of interest (frames). We have subsequently modeled each frame by either developing a new ontology or by selecting a relevant, existing model from (global) libraries that has a scope similar to the frame of interest. Available ontologies and approaches are therefore assessed on relevance and, possibly, adapted and integrated for our purposes. In addition, relevant theories and models of the concerning scientific research fields have been identified and formalized for adoption in the ontology.

The frames we have identified and modeled so far are, among others: (1) human-machine roles and actors involved in self-management, (2) tasks and goals (including self-management tasks and associated goals), and results and the setting they take place in, (3) diabetes self-management activities and games, (4) issues related to medical examinations (e.g., lab values), and (5) dialogue management through a combination of dialogue acts and shallow semantic frames. A more elaborate PAL ontology may include interaction and behavior models of robot and avatar, a model for privacy of information of self-management activities, a model to cover the agreements and social contracts between patient and PAL actor (avatar or robot), and a model for emotion and sentiment that covers the emotional responses of both PAL actor and child to interaction as well as the general state of mind of the child.

		Child					HCP		Parent	
		Knowledge	Awareness	Attitude	Self-efficacy	Skills	Trust	Acceptance	Attitude	Knowledge
Pre-cycle 1	Conversational fillers [20] 6.1			+Question: Relatedness						
	Meander Medical Centre 2015 [13] 6.2	+Quiz	+Interview	+ Question: Relatedness	+Question: competence, autonomy	+/- Usage		∼Interview	∼Interview	
	Diary support and self-disclosure [11] 6.3		+Timeline	+ Question: Relatedness	+Question: competence, autonomy	+Usage				
Cycle 1	Cycle 1 Evaluation [1] 7	+/-Quiz	+Interview	+Question: Relatedness, QualOfLife	+Question: competence, autonomy	+ Usage	+Usage	+Usage	+Interview	+Anecdote
Post-cycle 1	HCP Control panel [10] 8.1							+/- Usability Test		
	Self-disclosure & bonding [6, 5] 8.2			+Usage, Question: Relatedness						
	Camp yr2,Italy 8.3					+/- Scenario walk-through		+/- Interview		
	Camp yr2,NL 8.4		+/- Cartoon and acting elicitation	+Question: Relatedness, Teach style, Feedback					+ Question: Feedback + Interview	

Table 1: Year 2 experiments

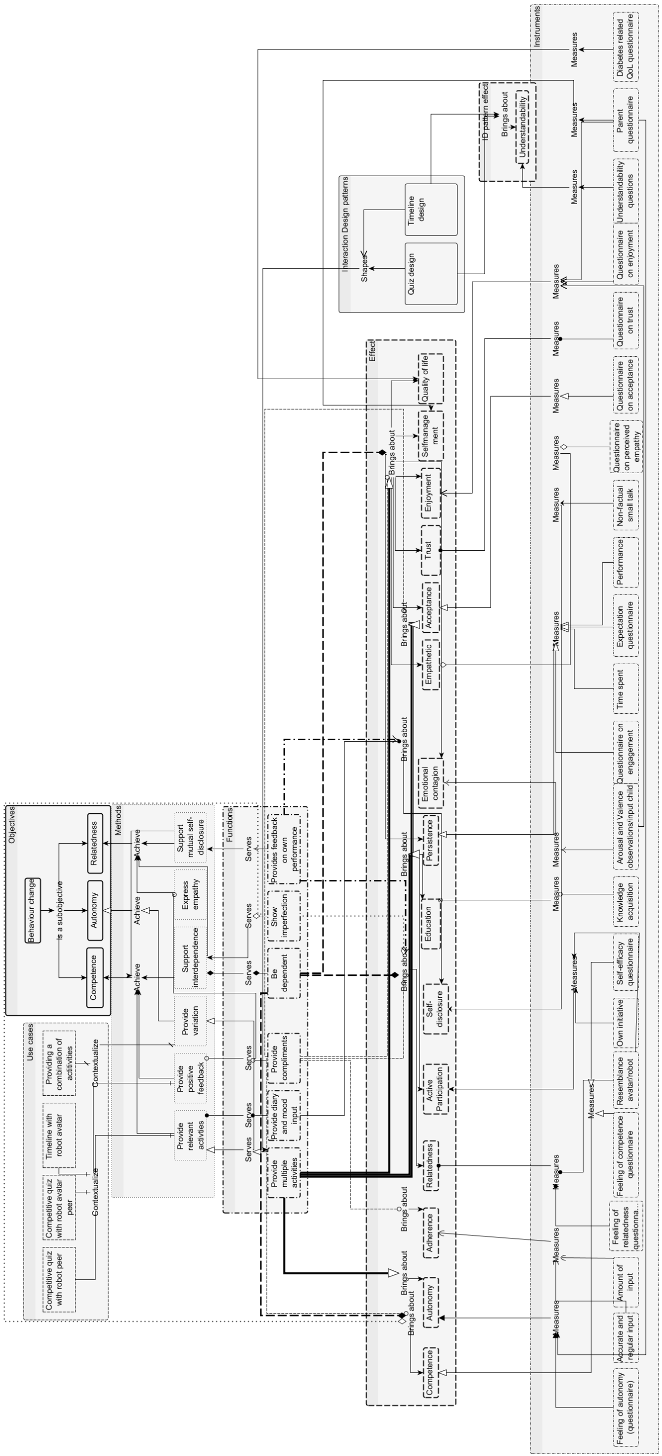


Figure 3: situated Design Rationale Cycle 1 - PAL system children

6 Pre-Cycle 1 evaluations

6.1 Pre-cycle 1 - Conversational fillers.

An extra evaluation focused on the effects of using conversational fillers (CFs), such as 'um', 'hmm', and 'ah', which may help to improve the human-robot interaction by smoothening the robot's responses. The controlled experiment with 26 (healthy) children showed that these CFs can improve the perceived speediness, aliveness, humanness, and likability of the robot, without decreasing perceptions of intelligence, trustworthiness, or autonomy. This knowledge is not implemented in the cycle 1 evaluation, but is expected to be used in the future cycles.

6.2 Pre-cycle 1 - Analysis Integrated Experiment Meander Medical Centre 2015.

Within the ALIZ-e project we performed many evaluations and this accumulated in an integrated experiment with end-users in the end of that project. In the beginning of the PAL project, the data were further analysed to derive relevant insigths from them. Children came to the hospital three times to interact with the robot for about 20 minutes. Each session was around an hour including questionnaires and interviews. They did three educational child-robot activities (quiz, sorting game and video watching), two intervening child-robot interactions (small talk and walking), and specific tests to assess the children and their experiences. Seventeen children (age 6-10) participated in the evaluation of this scenario, which provided new insights of the combined social robot support in the real environment. Overall, the children, but also their parents and formal caregivers, showed positive experiences. Children enjoyed the variety of activities, built a relationship with the robot and had a small knowledge gain. Parents and hospital staff pointed out that the robot had positive effects on child's mood and openness, which may be helpful for self-management. The experiences of this evaluation guided our first cycle protocol in having more knowledge on the differences between the children and the usability of different instruments. In Annex 11.2 the evaluation is described in more detail (note that robot behaviour was controlled by a "Wizard of OZ", for a main part).

6.3 Pre-cycle 1 - Diary support and self disclosure.

Based on a thorough analysis of operational demands, human factor knowledge and technological principles a selection of functions for an avatar that supports the self determination theory [7] was derived. By autonomously responding to the added content in a social fashion, e.g. matching the gestures and speech of the avatar appropriately to the mood of the child, children feel more supported in their competence and relatedness (see also Figure 4).

Functions	Purpose	Interaction design pattern	Human Factor Requirement
Display page, process button press, authenticate user	Basic application functionalities	-	-
Log user and avatar actions	Research	-	-
Create & read activity, measurement and picture	"Keeping the diary"	-	-
Update & delete activity, measurement and picture	Provide organizational ownership	-	AUT6
Ask to add activity together or alone	Provide procedural ownership	Question with answer options	AUT6
Add activity together	Provide structure with guidance	Ask activity item	AUT5
Add activity alone	Provide structure without guidance	-	AUT5
Create, read, update and delete activity type	Provide organizational ownership	-	AUT6
Respond to added activity	Relatedness support Competence support	- Mood matching - Question with answer options - Open question - Remark - Self-disclosure - Praise	REL2 REL1 REL1 REL1 REL3 COM1
Encourage adding a goal	Stimulate external motivation	Encourage	AUT3
Encourage completing active goal	Stimulate external motivation	Encourage	AUT3
Create goal	Autonomy support	-	AUT1, AUT2
Praise adding a goal	Competence support	Praise	COM1
Praise completing a goal	Competence support	Praise	COM2, AUT4
Read goal	Competence support	-	COM3
Delete goal	Provide organizational ownership	-	AUT6
Legend	Avatar	Web application	

Figure 4: Functional table of MyPAL that lists all the relevant functions and their relation to the interaction design patterns and human factor requirements. The blue rows are functions executed by the avatar and the white rows are executed by the web application.

Figure 5: Example of avatar self-disclosure. A picture that was accompanied by the text: "Do you do groceries sometimes? I'm always allowed to sit in the cart

A prototype was developed based on this, with for instance self-disclosure by the robot (see Figure 5). The evaluation centered on the avatar behaviors and its effects on the attitude of the children towards the robot, motivation support and performance. Performance is measured in terms of the amount and the consistency of the added content. A three-week user study with 13 children with diabetes was performed for this evaluation. Results show that almost all the avatar behaviors are picked up by the children and that those behaviors positively affect the motivation and performance of the children.

Finally, two design recommendations were derived that modulate the effectiveness of MyPAL. The first is *avatar quality over quantity.* The avatar behavior must be appreciated by the children in order to be effective. Simply showing more avatar behavior does not increase the appreciation. The behavior must match the childrens preferences. The second recommendation is *avatar sociability is key.* The more social the behavior of the avatar is, the more it is appreciated by the child and the more motivated the children are to add more content consistently.

The evaluation also provided another way for filling in questionnaires, we noticed in the Meander experiment that a more tangible way of filling in, instead of setting crosses, improved the answers and the thinking about answers. In this evaluation playmobil figures were used to put on the preferred answer. We are thinking about ways to combine this knowledge in a digital/tangible questionnaire.

In annex 11.3 the complete report can be found for more background.

7 Cycle 1 Experiment

7.1 Introduction and aims.

The cycle 1 experiments belong to the Initial Phase of the Project activities, named *First Development and Validation*. The aim of this phase is to obtain a first group of the results by the User Requirements Analysis implemented within the first prototype of the PAL technology. To achieve this objective an experimental campaign was set up during which the Consortium aimed to:

- Have a first explorative evaluation of *(i)the effects of the 1st PAL system prototype on the User Experience of the participants* (both children, their parents and their Health Care Professionals) and *(ii) the Determinants characterizing the end-users involved* (for more details please consult the protocol PAL-643783-2 (see `https://doi.org/10.5281/zenodo.291470`).
- Get a first qualitative indication of the possible effect on diabetes-, health- and lifestyle-related behaviours of the children involved (see Annex 11.4 and Annex **??** for more information).

7.2 Study's participants

Participants were enrolled among patients of the three different hospitals of the PAL Consortium (Meander Medical Centre Amersfoort and Ziekenhuis Gelderse Vallei Ede in the Netherlands, and Ospedale San Raffaele Milan in Italy). The inclusion criteria were set on an age from 8-10 and at least 6 months after diagnosis without other comorbidities. For the parents, at least one should be involved in the experiments. The Dutch health care professionals also recruited children that did not fall within this inclusion criteria, which explains the difference in number of children between Table 2 and Figures 6 and 7. For the comparison between questionnaires only the children and parents that met the inclusion criteria were evaluated, see Annex 11.4. In regard to the usage all children were taken into account, even when they did not meet the inclusion criteria, as they did provide relevant information for improving the system, see Annex 11.6.

Taking into account the inclusion criteria, there were 10 Italian and 11 Dutch children, 19 Italian and 11 Dutch parents (excluding the parents of older children). With the Health Care Professionals we had mainly informal interviews that were in Italy extended during the camp (see also paragraph 8.3.6) and in the Netherlands with a post interview with several nurses (see also Section 8.1).

	NL Sample (n=25)		IT Sample (n=10)		Total Sample (n=35)	
	M	SD	M	SD	M	SD
Age (years)	9.44	1.635	8.80	0.919	9.29	1.482
Experiment length (days)	22.44	4.119	33.90	3.780	25.71	6.587
	N	% of NL	N	% of IT	N	% Total
Nationality	25	100%	10	100%	35	100%
Gender (girl)	11	44.0%	4	40.0%	15	42.86%
Gender (boy)	14	56.0%	6	60.0%	20	57.14%
Parents together	21	84.0%	10	100.0%	31	88.57%
Uses Pump	17	68.0%	2	20.0%	19	54.29%
Uses Sensor	3	12.0%	2	20.0%	5	14.29%
Previous experience Charlie	13	52.0%	4	40.0%	17	48.57%

Table 2: Characteristics of the children

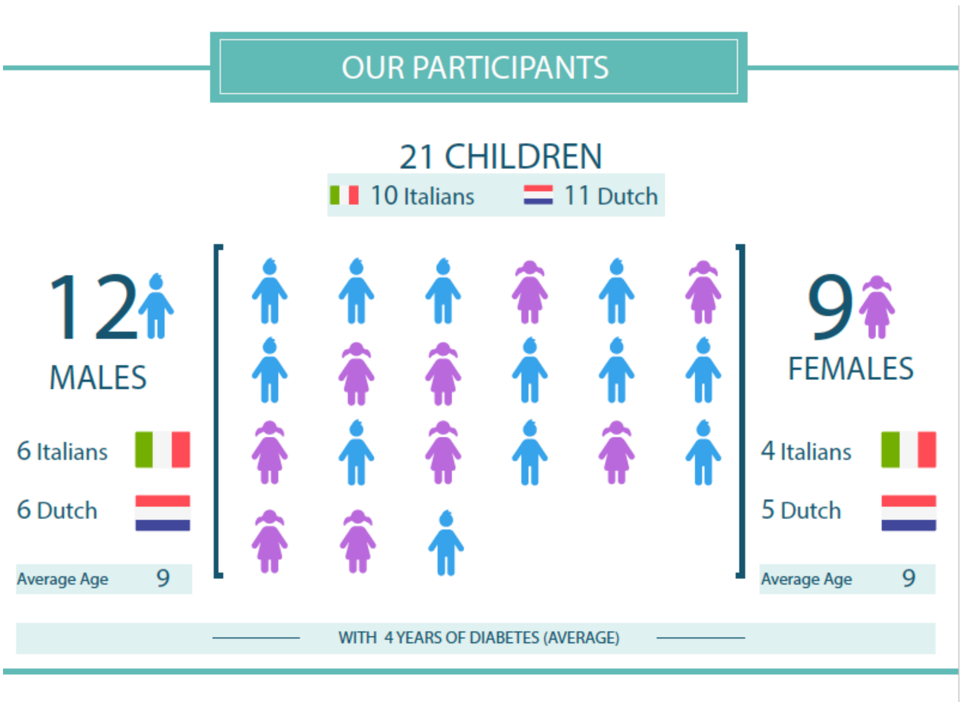


Figure 6: Characteristics of the children who participated to the May Experiments.

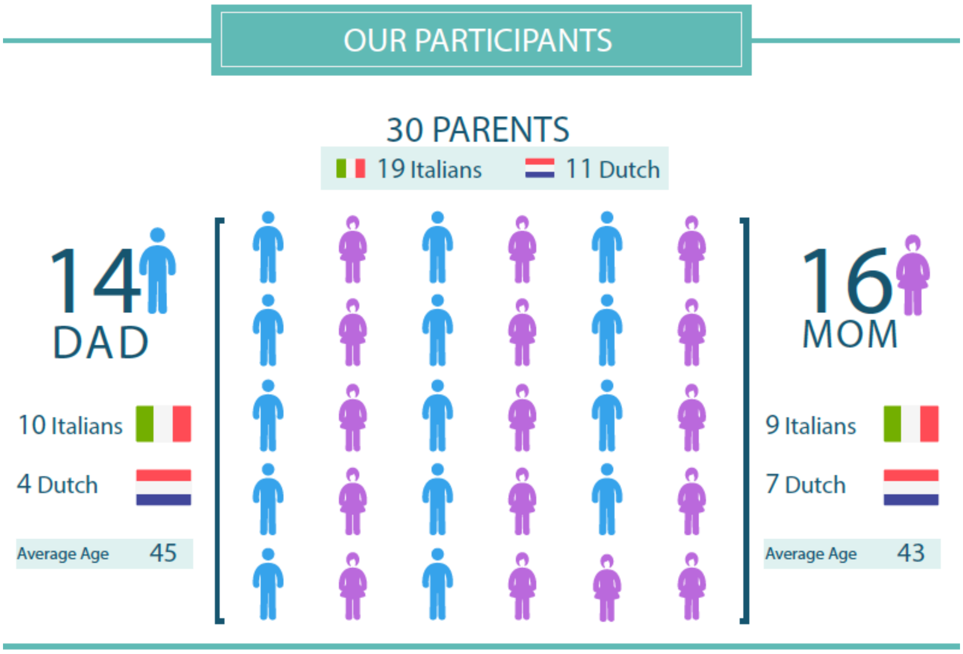


Figure 7: Characteristics of the parents from who the children who participated to the May Experiments.

7.3 Timeline and Settings of the experiments

The activities took place in the houses of the participants during the months of May and June 2016. Children and their families scheduled two visits to the hospitals (at the beginning and at the end of the study period), where each child had the possibility to interact with the NAO robot and the families met the research team and the Health Care Professionals in order to be instructed about the activities proposed. Moreover, in between, children and their families interacted with the 1st PAL System at home.

7.4 Technological components

The technological components of the 1st PAL System prototype consisted of:

- *The NAO robot* with which children played the Quiz game. Children and the NAO robot asked each other multi-choice questions (validated by Diabetological Pediatric Units cooperating to the project) from various domains, both general knowledge and diabetes-related.

- *Tablets* (one per child) with the MyPAL app which had the Quiz and the Timeline functionalities. The MyPAL-Timeline is intended to be a diary feature, in which children have the possibility to fill in a personalized report of their daily activities. Through the timeline children, according both to their specific diabetes management objectives and engagement in the System, could compile day by day: *(i) the therapy diary* - with details of glycemia checks and insulin doses; *(ii) a nutritional diary* - with the details of the meals; *(iii) an activity and emotion diary* - in which they could freely describe what they did during the day, also by uploading pictures, (e.g.: sports, party with friends, excursions, etc.) and which feelings they experienced in these occasions. The virtual avatar of the Nao interacted with the children during the Timeline use, for example with greetings and personalized motivational feedbacks tuned on the activities done (e.g.: Hi Sam, its nice to see you today, or Good work, youre accomplishing your objective very well today).

7.5 Instruments used to collect the data and information on their validation

The activities concerning the data collection phase were conducted through the use of several instruments, such as questionnaires and interviews (see `https://doi.org/10.5281/zenodo.291473` for an overview). Below are listed the kind of instruments used and their related purpose:

- *Semi-structured interview.* This type of interview is typically used in social sciences and in all the occasions in which the interviewer has a framework of themes to be explored. While a structured interview has a rigorous set of questions which does not allow the interviewer to divert, a semi-structured interview is open, allowing new ideas to be brought up during the conversation as a result of what the interviewer says [4]. In qualitative research and explorative studies, semi-structured interviews represent a flexible and effective method to derive the desired insights [8]. During these experiments, there were conducted three different interviews, on the base of the target. For example children participated to the *Initial Interview* which aim was to explore the attitude of the child towards robots before the experiments starts, their expectations and their current use of technology and the internet. They also participated to the *Mid-term Interview* which aim was intended to derive the first impressions on the usage in a real life setting of the PAL technology and its usefulness in achieving their own personalized diabetes-management goals. And finally, both the children and their parents participated in the *Final Interview* which aim was to derive the final impressions of the interviewed on the PAL System and discuss in detail their User Experience. The interviews have been structured differently for children and parents, mainly for the language implemented and detail of the questions submitted to their attention.

- *Questionnaires.* In order to carry out a specific qualitative analysis, many questionnaires were used during the experiments and, as for the Final Interview, they were structured differently both for the children and for the parents. For example, the *Family demographic questionnaire* was set up only for the parents and was used to derive a preliminary database of descriptors for the participating pool, to see if there could be found any differences among them (e.g.: age related, cultural related, etc). Another questionnaire used was the PedsQL^{TM}, which is a modular approach to measuring health-related quality of life (HRQOL) in healthy children and adolescents and those with acute and chronic health conditions (like T1DM in this case). It integrates seamlessly both generic core scales and disease-specific modules into one measurement system (generic [19] and diabetes specific [18] `http://www.pedsql.org/index.html`). The Consortium chose this questionnaire both for the children and for the parents, using two different versions bought from the editors. Other two questionnaires were submitted to the parents: the *Parents mid-term questionnaire* and the *Parents final questionnaire*. The first one was used to evaluate the first impression of the MyPAL app and the last one to collect their opinion on its usability and the experience at all.

- *The interactive Journey Map tool* provides a graphic and structured visualization about all the factors that influence the user experience, constructed from the user's perspective. In this occasion, the journey map is used with the purpose to extrapolate qualitative and quantitative data, involving directly the users (children with T1DM) by mapping their experience and telling their feedback and expectations about the PAL system in a playful and stimulating environment. Although the results were not analyzed in a structured way, the observations provided input for a subsequent system (see `https://doi.org/10.5281/zenodo.291471` for the guidelines on how to use the journey map tool).

7.6 Results, Usage by children

All the instruments used to collect data were very helpful to evaluate a lot of expectations, impressions and feedbacks of the children towards the Robot and the Avatar. All the questionnaires were submitted to the child before, and after the experiments started, in order to have a comparison of the data. However, in the middle of the experiment children gave their first impressions of the System through a phone call.

Below we summarize the results, but a more elaborated summary can be found in annex 11.5 and the complete thesis on this subject is annex 11.6.

The results showed that in general, there is an overall perception of a good quality of life related to diabetes and for children the mother is the major referee for therapy management within the family. In regard to expectations and motivation most children expected that the robot helped them in the management of the T1DM. They looked forward to play with both the Robot in the hospital and the Avatar on the tablet at home, and all of them were willing to participate, so neither the parents or the medical staff forced them. Results show also that there are no significant differences between the Robot in the hospital and the Avatar on the tablet, as calculated with a non-parametric paired Wilcoxon test.

Before the experiment, children mostly described the Robot as a technical toy and a helper and after the experiment most often as a helper and a friend or pal. This difference was not significant, but all cells in the cross table had a low expected count. Most children (around 70% or more) thought that a robot move, talk or see you. Only a quarter of the children were convinced that a Robot can show emotions. There are no significant differences in what children thought a robot can do before and after the experiment, as calculated with a non-parametric paired Wilcoxon test.

After the experiment ended, the children were asked how they felt towards the Robot and the Avatar. The children had a significant more positive attitude towards the Robot met in the hospital, compared to the Avatar on the tablet.

The children indicated that they trusted the robot a lot, learned a lot of new things by playing the Quiz game and think of the Robot as a friend. They do not need the help of their parents, but they shared a lot with them. For the nutritional diary instead, children sometimes needed help from their parents to fill it in, especially to calculate carbohydrates.

During the interviews with the children we learned that the voice of the Robot was sometimes unclear. The young children then had to read quick and indicated needing help with this and the older children said it went too slow. They like the appearance of the app, but they would like more colors and that the Robot makes more movements.

They suggested to improve the voice (in terms of speed which should be slower and more understandable) and the shape of the robot (e.g. more bigger), to add more games (e.g. memory games), to let the robot move more and respond to what they say or see. They also wished to interact with the Robot, sharing emotions, to share photos and feelings (songs, poems), music and have a nice screensaver on the tablet.

Three main trends were found in the system usage in which the majority of the users showed an overall low usage or quickly decreasing usage. A small number of users showed continuous and consistent usage throughout the entire experiment. As the personalization was only minimally implemented, the results are in line with common (digital) diabetes interventions. The results did not allow us to explore possible system usage and knowledge development predictors, but, however, they provided a solid baseline for further versions of the System in which the personalization will be further implemented. The main recommendations are to focus on the implementation of basic game design elements and personalized content to foster user engagement and continuous use. Maintaining the used measures (while adding some psychological predictors) and longitudinal experiment design will allow for comparative analysis in the further research cycles.

8 Post Cycle evaluations

8.1 Post Cycle - User-friendly Support of Health Care Professionals for Guiding Self-Management in Children with T1DM

During the first research cycle, to assist the health care professional, we aimed to study how PAL Control can aid them in guiding the children with T1DM self-management in a user-friendly manner. PAL Control is 1) a gatekeeper for information on the young patients to personalize their health care, 2) a tool to author and control the PAL robot and its avatar and aid the child in their self-management, and 3) a tool to provide explanations to the informal caregivers (e.g. parents) on the desired activities of the children. A first version of PAL Control (see Deliverable 2.1, PAL Control: a prototype

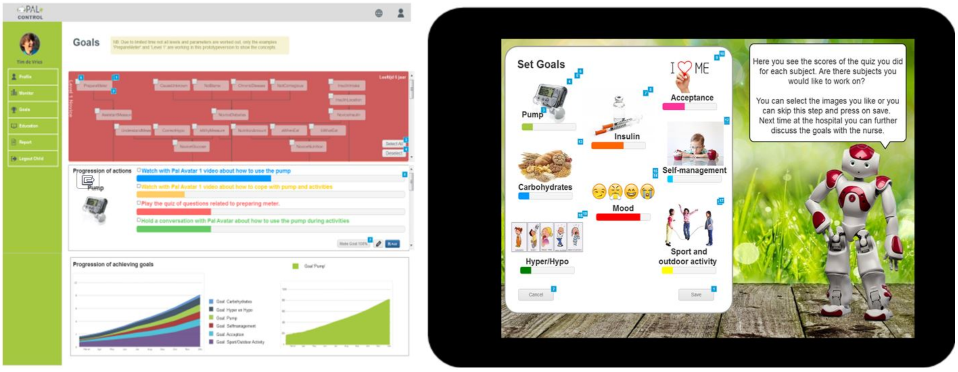


Figure 8: Example design progression Page for Health Care Professional in PAL Control and Assessment with PAL Robot or its avatar.

authoring tool interface for entering coaching protocols) enables health care professionals to set learning goals for children during consults, whereby the goals are visualized in a goal tree structure categorized by different levels. Furthermore, it enables the health care professional to enter data of the child including personal data and preferences such as sports and hobbies. Next, education materials can be added manually.

To fulfill this aim, a redesign (prototype) of PAL Control was developed and evaluated with end users (diabetes nurses) and an important stakeholder (diabetes doctor), following the situated Cognitive Engineering approach ([10]). This approach included 1) literature research and needs assessment to define the foundation; 2) specification of PAL Control redesign through scenarios, storyboards, use cases and requirements, and 3) evaluation of redesigned PAL Control prototype through think aloud, semi-structured interviews with health care professionals (N=5) and administering the System Usability Scale ([2]).

New functions of the the redesigned PAL Control were the following: an knowledge and skill assessment of the child with PAL Robot or its avatar in MyPAL, goals setting and actions selection, writing reports, an educational database for children and parents, monitoring solutions and a progression page (see Figure 8 for an example design).

Results showed that the Health Care Professionals (HCPs) evaluated PAL Control as unfriendly. Also, they felt that a combination of an assessment with a robot or its avatar, setting goals, selecting actions and the progression page, is a suitable and effective approach to HCPs in guiding children with diabetes self-management. It provided them support in making the consult with children and parents more meaningful, in comparison with the existing version of PAL Control, due to the fact that they can understand their needs better. In addition, valuable pointers were elicited for further development. The procedure and results are described more elaborately in 11.7.

8.2 Post Cycle - Self-Disclosure for Social Bonding.

An experiment that was performed in the follow up of the Cycle 1 evaluation looked at the development of a disclosure module. The idea is based on the Social Penetration Theory, which says a bond can be welded through the reciprocal disclosure of information about the self. The module was integrated into a mobile application with avatar presence for diabetic children and subsequently used by 11 children in an exploratory field study over the course of approximately two weeks at home. It was found that the relative amount of disclosures that children made to the avatar was an indicator for the relatedness children felt towards the agent at the end of the study. Girls were significantly more likely to disclose and children preferred to reciprocate avatar disclosures of lower intimacy. No relationship was found between the intimacy level of avatar disclosures and child disclosures. Particularly the last finding contradicts prior child-peer interaction research and should therefore be further examined in confirmatory research. More information on this study can be found in annex 11.8 and annex 11.9.

8.3 Post Cycle - Camp in Italy

8.3.1 Introduction and aims

After the Cycle 1 Evaluation the consortium participated also in the Summer Camps organized by the Diabetes Associations for children with T1DM. The experiments carried out in those occasions belong to the the Initial Phase of the Project activities, named *First Development and Validation*.

The research aim of the experiments carried out in Italy was to investigate the usability of the MyPAL app in terms of its effectiveness, understandability, learnability, operability and satisfaction from the perspective of the end user (e.g. children enrolled). In particular, its main objective was to evaluate, in the most objective manner, the usability of the app, in order to identify which were its still workable features and the functionalities that have to be implemented in the future to create a second release of MyPAL app.

8.3.2 Study's participants.

Participants were children with an age between 10 and 14 y.o. and none of them, as another inclusion criteria, used the MyPAL app in the past.

8.3.3 Timeline and Settings of the Experiments

The Summer Camp was organized by the association SOStegno70 and the experiments were carried out from the 25th of August 2016 to the 3rd of September 2016 in Misano Adriatico (Rimini) in a scholastic district.

In particular, two kind of activities were held:

- *MyPAL usability test*
- *Brainstorming with Healthcare Professionals*

8.3.4 Technological components

During that Summer Camp experiments were carried out using tablets with the MyPAL app, wich included Timeline and Quiz game functionalities, in order to conduct the *Usability Test.* During the camp, the NAO robot was not used.

8.3.5 MyPAL usability test

Usability testing is a technique used in user-centred interaction design to evaluate a product or service usability (how easy it is to use) by testing it with representative users. This technique can be seen as an irreplaceable usability practice, since it gives direct input on how real users use the system [16] The goal is to identify any usability problems in the use of the product, collect the related qualitative and quantitative data and determine the participant's satisfaction with the product or service. Basically, the usability testing focuses on measuring a human-made product's capacity to meet its intended purpose through the users interaction with it. The main aim would be to judge what changes to make to remove design problems that cause users frustration. Examples of products that commonly get benefits from usability testing are web sites, web application, computer and mobile interfaces and devices.

To carry on the Usability Test of the MyPAL app and provide answers to the research questions (for more details please consult the Protocol PAL643783-3), all the aspects belonging to three methods were considered *(*PACMAD, NIELSEN and ISO methods). In particular,the hypothetical positive aspects were identified that should be kept or emphasized, and the negative aspects that might provide troubles (issues, risks, and concerns) to the user. Then the following data were needed:

- *Verbal responses* provided by the spoken answers of user that demonstrate what, if anything, was most frustrating about MyPAL app.
- *Errors* provided by the user actions, that demonstrate if an icon, button, etc., is misleading and can be cause of misunderstanding.

- *Rating scale (Likert scale)* to collect the users satisfaction feedbacks

All these data were needed to quantify, through statistical analyses, the following usability *criteria*:

- *Navigability* which measures, in terms of usability, how fast and how many clicks does the user takes to complete the task
- *Readability* which measures, in terms of usability, how the contents are legible, understandable and enjoyable.
- *Task Success rate* which measures, in terms of usability, how easily a user completes the task (reach the goal) according to the number of errors detected during the task.

All the children enrolled participated to the MyPAL usability test once and for a duration of maximum 1 hour, in their spare time among all the other activities scheduled by the association SOStegno70 and the Healthcare Professionals involved.

The Usability Test required the presence of a two moderators: one *Facilitator* and one *Observer*. The *Facilitator* directly interacted and communicated with the participant and was responsible for all aspects of administration. The *Observer* does not interact directly with the participant, in order to avoid any changes of participants behaviour due to his or her presence. THe observer focuses more on the technical aspects of the product to be tested, ensuring that the product or service did not malfunction during the test and ensuring a double check of the work done by the *Facilitator*.

8.3.6 Brainstorming with Healthcare Professionals

Brainstorming is a technique used to foster creative thinking about a problem. The aim of brainstorming is to produce numerous new ideas, and to derive from them themes for further analysis [17]. This technique has been used to exploit the possibility to have several Health Care Professional, like for instance diabetologists and nurses at the same time during the Summer Camp. Thanks to the participants heterogeneity, the brainstorming has been used to focus all of them on a specific topic, in order to collect many possible solutions from several points of view.

To succeed the activity there were projected a set of slides to guide participants across the following topics:

- Feedback and suggestions about the MyPAL Control
- Health Care Professionals opinions about possible functionalities to be implemented in the system
- Example of MyPAL control User Interface (UI): Create a childs profile

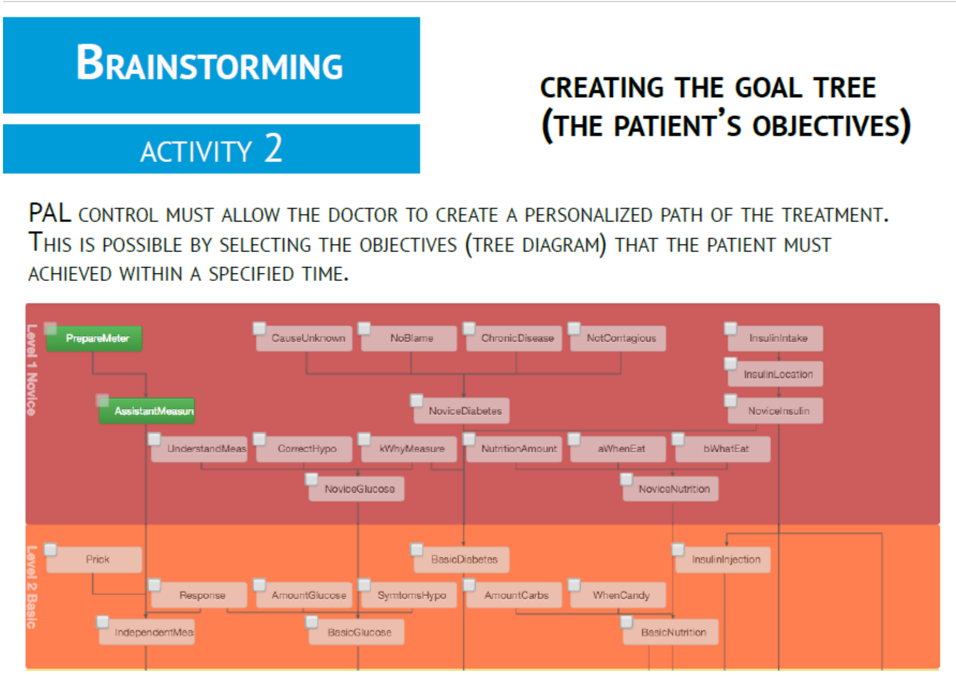


Figure 9: Goal Tree activity during the Brainstorming.

- Example of MyPAL control UI: re-design of the Goals tree (e.g. see Figure 9)
- Possibilities to improve the Quiz (considering both informative contents and engagement aspects in order to foster children to have a good management of the disease)

8.3.7 Instruments used to collect the data and information on their validation

The activities concerning the data collection during that experiment were conducted through the use of questionnaires. In details, the questionnaires used were:

- *The Technology Questionnaire Habits* which is questionnaire specifically designed for the projects purpose, and used also during the previous experimental campaigns and it consists of six questions in the form of multiple choices, to assess the level of confidence of the children with the technology.
- *System Usability Scale (SUS)* which is a simple, ten-item Likert scale (five points Likert scale from completely disagree to completely agree)

that can give a global view of subjective assessment of system usability [3].

For the brainstorming activity no questionnaire was used.

8.3.8 Results from the Italy Summer Camp

All the *criteria* quantified through the statistical analyses showed which were the most difficult functionalities for the children who participated to the Usability Test of the MyPAL app. We discuss the *Quiz* and the *Timeline*. For the *Quiz*, the main errors were:

- Time to read and answer the questions was not enough;
- Children didn't understand their role in turn with the Avatar;
- Children didn't understand which were the correct answer after have played a turn;

For what about the *Timeline*, the main errors were:

- Children didn't specify the day related to the activity they wanted to insert in the Timeline;
- They didn't recognize the *Activity* section of the Timeline;
- They didn't use properly the meal section (e.g. they didn't recognize the meanig of servings, or they filled all the foods eaten in the same string)
- They confused to define the time by AM/PM;
- They didn't recognize the link between the meals and the glycaemia check or the insulin injection;
- They didn't recognize some icons (e.g. the *sickbed* status in the emotional diary, or some icons related to the meals of the day);

Instead, the Brainstoming activity allowed researchers to collect many inputs about the PAL system.

According to the Healthcare professionals the *gamification* is very important for children since it gives the possibility to teach very important topics on T1DM through the entertainment.

The Goal Tree structure could be improved but it is hard to define a common structure for different countries, since the practices related to the management of the T1DM could be different.

The glycaemic diary could be useful only if there could be an integration with the glucometers, while the nutritional diary could be useful if it could be facilitate children and family in the carbohydrates counting.

8.4 Post Cycle - Camp in Netherlands

8.4.1 Introduction and aim

In the Autumn of 2016 an field test was conducted of PAL, during a four-day diabetes camp for children diagnosed with T1DM in the Netherlands (21 children (13 males and 8 females), aged 8 to 11 (M=9.24, SD=1.09)). The camp was hosted by the project partner Diabetes Association the Netherlands (DVN). The title of the camp was Robots & Heroes (http://sugarkidskampen.nl/robots-en-helden/). During the camp, multiple robot buddies were introduced, as part of the PAL system, to be childrens heroes and help them with their diabetes. A video was made that captures the atmosphere during the camp `https://youtu.be/7-Em2SHPbu8`. During the camp we had different research questions for which we got input during different activities. Each child got a personal tablet.

1. Interact with and evaluate the different activities (timeline, objectives, games)
2. Compare between robot and avatar (are they the same)
3. Perception of competence and warmth
4. Provide input on what the robot/avatar can do (drawing activity)
5. Provide input on what the robot/avatar should say in reaction to timeline input (theatre)
6. Evaluate what kind of explanation about robot/avatar is preferred (parents and children)
7. What do parents expect from PAL

Next to this, the NAO also provided fun activities, by telling a bed time story each evening and of course a performance at the last evening.

The results of these research activities are analyzed and written at this moment. Below we summarize in short the main findings.

8.4.2 Interact with and evaluate different activities

During the camp there were multiple interaction moments with the system. On the first evening the different activities of MyPAL were explained, the quiz, the timeline, the objectives and the break & sort. The objectives were personal and based on input from the parents before the camp.

During the four camp days, activities with both embodiment and virtual PAL Actor were organized by the research team. On day 1, the five robots held a plenary talk, introducing themselves and the system, followed by small talk conversations in groups of four or five children. On day 3, robot-rounds were organized with four different games (1: playing the quiz with the robot, 2: playing the quiz in the My PAL app, 3: playing the Break&Sort game with the robot, 4: playing the Break&Sort game in the My PAL app). All children participated in groups of two in all four games. By the end of day 3, there was a disco night with the robot giving a dance performance. In addition, every night two robots came to tell a bedtime story to the children with some funny anecdotes of their experiences that day. Finally, children could play with the My PAL app every evening of the camp after dinner. All children were given the same opportunities and playtime with robot and avatar.

The questionnaires are not analyzed yet, but the interactions were observed by the team and we noticed that some activities were quite hard to do (e.g. filling in the time line, understanding sort & break). Furthermore, children provided a lot of feedback on things they liked and disliked. This is really and advantage of being present during the camps as the children are really honest.

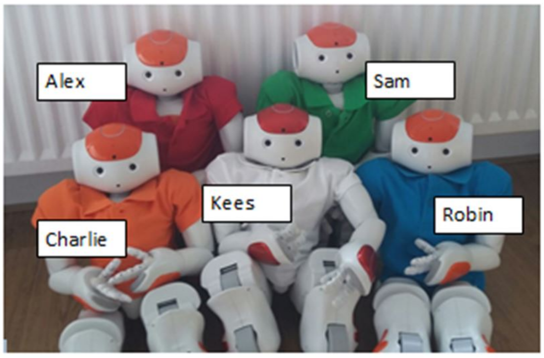
8.4.3 Compare between robot and avatar

A main idea from the PAL project is that the bond between robot and avatar is transferable. During the camp we tested three hypotheses related to this (also see Figure 12). First, children perceive the robot and avatar as the same entity (i.e., migration). Second, the more the children interact with the robot and avatar of a PAL Actor, the more they develop a bond with that PAL Actor. Third, a stronger bond with PAL Actor leads to a higher motivation to play on activities with that PAL Actor and higher usability of MyPAL.

To evaluate this the children were introduced to five PAL actors each with a different personality based on the BIG 5 [9]. Each robot represented an extreme end of one of the five personality. Each robot represented an extreme end of one of the five personality traits. In addition, we gave each robot a name and a different colour t-shirt. Robots that participated in the camp were: Charlie, the diligent/careful robot (conscientiousness), Alex, the happy robot (emotional stability), Sam, the sweet robot (agreeableness), Robin, the creative robot (openness to experience) and Kees, the shy robot (extraversion).

The children then had the interactions as described in the previous paragraph. To answer the hypotheses of this research we focused on four variables: similarity, friendship, usability and motivation. These were scored quantitatively on a Likert-scale with questionnaires. Children were asked to rate the PAL robot and avatar on the first and last day.

Qualitative aspects of our study consisted of childrens opinions about the PAL Robot and Avatar and were measured by open questions in the



Name	Robots personality	Shirt
Charlie	The diligent/careful robot. He is the leader, and likes everything to be organized.	orange
Robin	The creative robot. He is curious and likes to come up with new things to do.	blue
Alex	The happy robot. He doesn't worry about anything and he is always cheerful.	red
Sam	The sweet robot. He is caring and wants to make sure everybody is allright.	green
Kees	The shy robot. He is very quiet and likes to think a lot.	white

Figure 10: Five PAL Actors, their name, color and personalities

questionnaire as well as observations during the camp.

All children preferred the physical PAL Robot over the PAL avatar, when they were asked to choose. Children preferred the physical robot mostly because he was real, talked more often and showed more body movement and was able to dance. In regard to our hypothesis, results show the following (also see Figure 13). In general childrens scores on similarity between physical robot and avatar was not very high (M=1.76, SD=0.83). The open questions on this subject indicate that the main reasons for perceiving differences, were again the capabilities, talking and movements of the agents. Perceived similarity between PAL Robot and PAL Avatar positively affects the level of bonding between the children and PAL Avatar (R2=.79, p=.016). The level of bonding between the children and PAL Avatar affects children's motivation to play with MyPAL (R2=.39, p=.036) and children's rating of MyPAL's usability R2=51, p¡.001). Results also showed an interaction effect, as shown in Figure 12, of the age of children with perceived similarity on bond with the PAL Avatar. For older children, the perceived similarity had a additional effect on the bond with PAL Avatar (F(2,21)= 6.44, p=.02).

Figure 11: PAL Activities at camp

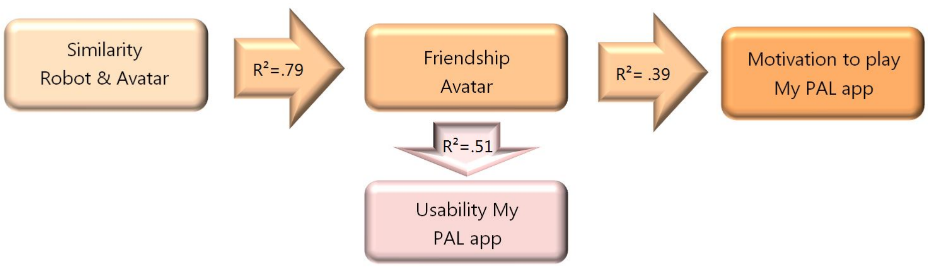


Figure 12: Relation between Perceived Similarity between PAL Robot and Avatar, Bonding between Child and PAL Avatar, Usability of MyPAL and Motivation to Play with MyPAL

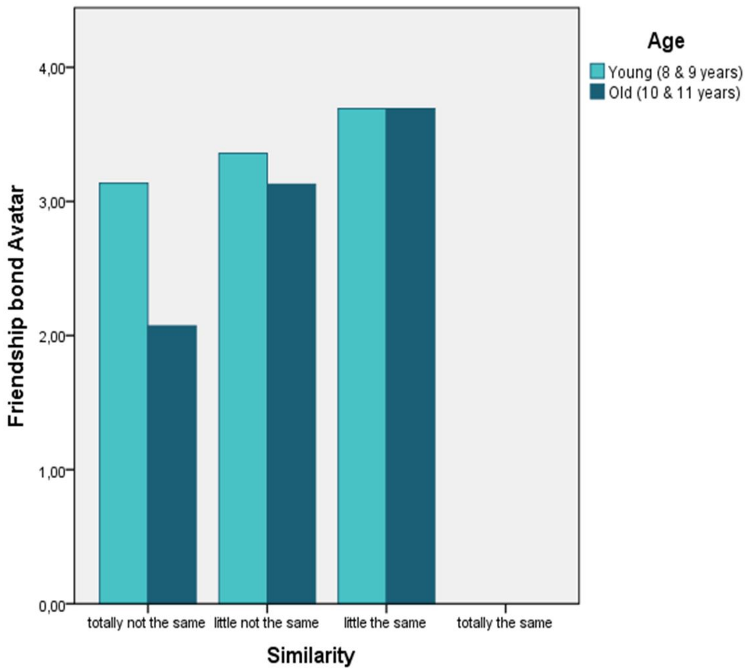


Figure 13: Interaction effect of the age of children with perceived similarity on bond with the PAL Avatar.

8.4.4 Perception of competence and warmth

In the PAL project we want to present the information in such a way that children pay most attention to it and learn from it. One way to do this in a personalized manner is to adapt the teaching style to the learners preference. As a first step towards this we evaluated if the teachers style was perceived as we intended it. For this we now looked at two dimensions, warmth and competence. All four possible combinations were evaluated in a class room, only the two extremes during camp. Low warmth, low competence had the following characteristics:

- Closed
- Gaze away
- Wobbling
- Fiddling

Where High warmth, high competence was open and up-front.

After the presentation the two different robots gave on the first day children were asked to rate the two robots using 20 words on a three point scale. The words were rated by experts on four dimensions (warmth, competence, dominance and affiliation).

The results show that the robot hat intended to express high competence was also perceived this way. The warmth was only perceived when there was also a high competence. We also saw that at camp the overall scores for perceived warmth and affiliation were lower. This is an interesting results, but we must take into account the small effect size.

8.4.5 Provide input on what the robot/avatar can do

To get input on how a robot can help and what it should do we did a creative co design session. Children were asked to finish a cartoon. The hypotheses was that the children would provide input on functions of the robot, the style it should interact and the non-verbal behavior. Unfortunately this seemed quite hard. Although we did get some good ideas from the discussion with the children during the drawing.

8.4.6 Provide input on what the robot/avatar should say in reaction to timeline input (theatre)

The question here was: What should the robot say and/or do in specific scenarios with input on the timeline?

To answer this question children were paired. One of the children got into a robot suit and played the NAO. The other was asked to reenact a specific scenario. The robot then had to react on the inputs from the child.

In the end 5 pairs performed this session and it resulted in an overview of what kind of feedback they preferred: immediate, context specific, motivational. This means that the feedback types that have been chose for the system are encouragement, mood matching, praise, and confirmatory response.

Furthermore, because the children spoke with each other it resulted in example sentences that can be used in the system. At the moment the final hand is laid on an overview of relevant sentences and triggers so this can be implemented in the PAL system and evaluated in the future.

8.4.7 Evaluate what kind of explanation about robot/avatar is preferred (parents and children)

The question here was: What explanation best helps child/parent to understand why the robot/avatar performed a certain action? To evaluate this both goal based and belief based feedbacks were proposed to the children during the 3rd day robot round and to the parents on the fourth day. Both groups preferred goal based feedback.

8.4.8 What do parents expect from PAL

The final day, parents of participating children were invited for group discussions on how PAL could support them in the care of their children with T1DM. Results of these group discussions provide input for PAL Inform, the support tool for informal care givers, such as parents. Results of the group discussion with parents (N=18) were the following.

Parents felt that PAL is for the children. Parents appreciate the possibility to monitor the development and health of the child, but do not want to be too strongly involved in the child-PAL interaction. PAL is fun for the children and this should not change. In addition, they experienced that in daily life some topics were difficult to discuss, such as challenges due to and feelings about T1DM. PAL Robot could help the child to voice them, as they are more likely to disclose this type of information to a 'robot buddy' than to their parents.

9 Refine specification

PAL follows a systematic design and evaluation process that aims at specific knowledge increments. The design specifications entail objectives, methods. use cases, requirements, claims, and interaction design patterns (with some additional illustrations like personas and storyboards; see Figure 2). These specifications are refined based on new insights that are derived from theory- and data-driven analyses of recent research. First, the analysis of specific outcomes of the cycle 1 evaluations lead directly to adjusted design specifications to improve, for example, the usability of MyPAL and the goal-structure in PAL control. Second, additional proposals for new functionalities (i.e., user requirements) will be worked out to improve PAL's support as described in research questions.

To guide the second process, the research questions are reified via so-called claims, which specify the specific functionality (requirements) and expected effects (measures) in the envisioned contexts-of-use (use cases). At the highest level, claims are structured as follows:

```
<preposition><use case><functionality><preposition><component><verb><noun><measure>
```

Currently, the following claims are being considered to be worked-out in a next design-test cycle:

1. *Immediate Feedback.* During all activities, immediate feedback by PAL-actor improves childs motivation, bonding and learning.

2. *Self-disclosure.* During small talk, self-disclosure by PAL-actor increases childs self-disclosures.

3. *Goal Structure.*

   - During all activities, the hierarchical goal structure representation improves childs understanding and motivation.
   - During goal selection, the hierarchical goal structure representation improves HCPs goal selection efficiency.

4. *Timeline.* During all timeline activities, the new mobile user interface improves usability.

5. *Guidance.* During all activities, the activity introduction and fading-out" guidance by PAL Actor, improves childs (a) understanding of the activity, (b) motivation to do the activity and (c) activity-related learning.

6. *Emotion annotation.* During emotion annotation, the affect button (a) is easy-to-use and (b) senses emotional state variety of the child.

10 Conclusions

The general iterative process from needs analyses, via design generation & specification, prototyping, to evaluation will be done in 3 cycles. This report reports the results of the first cycle and the setup towards the next cycles. In this report the results from MS9 (First evaluation completed) are reported. MS 10: design specifications, ontology and test protocol for cycle 2 is partially fulfilled. The design specifications and ontology are described, but the test protocol is expected to be finished beginning of M25, and the evaluation is expected to start in M27.

The different evaluations and large involvement of the different end user parties have provided us with lots of input. Conducting the experiments in the first year, as exploratory and formative evaluations, was a good decision: It provided us with feedbacks, but also discovered technical problems that we couldn't have uncovered without diverse and relatively long-duration usages. An example of this is the input from the HCPs on the control panel. First, they assessed a PAL prototype via expert reviews and walk-throughs, but only when they had to really use it they were able to provide sound and detailed requirement feedbacks. Participants knew they were working with a work in progress and were in generally quite forgiving and very enthusiastic. For a large part they are interested in participating again in next experiments and also the camp participants were really enthusiastic. In comparison with other camps in the Netherlands, the PAL camp has no trouble with finding volunteers and already has many applications (although it only is in October).

Overall, PAL functions were developed that proved to support the three high-level objectives that were derived from the Self-Determination Theory: autonomy, competence and relatedness. More specifically, the WP-tasks provided the following major outcomes:

1. The Meander hospital and cycle 1 experiments showed that playing PAL activities with the PAL actor in recurring sessions is associated with increased child's diabetes knowledge, awareness, attitude (i.e. relatedness), self-efficacy (i.e., autonomy and competence) and skills to use the acceptance and PAL system.

2. Diary support and self-disclosure support seem to provide further enhancements of the first observations mentioned above: Awareness (related to the timeline), attitude, self-efficacy and skills.

3. More specifically, conversational fillers proof to enhance child's attitude (i.e., relatedness) to the PAL robot.

4. HCP's are positive about the opportunities of PAL, but the system needs substantial improvement for acceptance and trust (i.e., improvement of the usability and goal structure).

5. Similar to the HCP's, parents are positive and prove to have a good attitude towards the PAL system and seem to experience already some knowledge gain.

6. There is a large diversity in child's usage of the PAL system; different classes of child's usage have been identified. Enhanced personalization is needed to establish adherence over the diverse user group.

7. A number of usability bottlenecks have been identified, which need to be solved to improve the PAL usage.

The different topics of this workpackage are on the cutting edge of current developments. This can be seen for instance by the different workshops we are organizing or invited to. Just a quick look at the workshops during HRI 2017 shows one about Child-robot interaction (co-organized with PAL partner), a workshop on benchmarking where we (in an earlier year) presented our methodology, robots for learning workshop that of course also discusses interesting aspects for the PAL project, and a workshop on socially assistive robots and data collection and sharing. Next to this our work on ontologies was presented during a workshop on Artificial Intelligence for Diabetes.

References

[1] Anika Boelhouwer. Analysis of system usage and knowledge development of the current PAL system for children with Type 1 Diabetes Mellitus. Master's thesis, University of Twente, September 2016.

[2] Simone Borsci, Stefano Federici, and Marco Lauriola. On the dimensionality of the system usability scale: a test of alternative measurement models. *Cognitive processing*, 10(3):193–197, 2009.

[3] John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[4] A Bryman. *Social research methods*, chapter 15.

[5] Franziska Burger. Agents Sharing Secrets Self-Disclosure in Long-Term Child-Avatar Interaction. Master's thesis, Radboud University Nijmegen, September 2016.

[6] Franziska Burger, Joost Broekens, and Mark A. Neerincx. A Disclosure Intimacy Rating Scale for Child-Agent Interaction. *International Conference on Intelligent Virtual Agents*, pages 392–396, October 2016.

[7] Edward L Deci, H Eghrari, Brian C Patrick, and D Leone. Facilitating internalization: The Self-determination theory perspective. *Journal of Personality*, 62(119-142), 1994.

[8] Rosalind Edwards and Janet Holland. *What is qualitative interviewing?* A&C Black, 2013.

[9] Aa Jolijn Hendriks, Willem KB Hofstee, and Boele De Raad. The five-factor personality inventory (ffpi). *Personality and individual differences*, 27(2):307–325, 1999.

[10] Jet Shin Hong. Healthcare professionals gain control of childrens diabetes self-management. Master's thesis, University of Amsterdam, 2016.

[11] Mike Ligthart. MyPAL: A Digital Diabetes Diary with a Responsive Social Avatar. Master's thesis, Radboud University Nijmegen, June 2016.

[12] Rosemarijn Looije, Mark A Neerincx, and Koen V Hindriks. Specifying and testing the design rationale of social robots for behavior change in children. *Cognitive Systems Research*, 2016.

[13] Rosemarijn Looije, Mark A. Neerincx, Johanna K. Peters, and Olivier A. Blanson-Henkemans. Integrating robot support functions into varied activities at returning hospital visits. *International Journal of Social Robotics*, 2016.

[14] Mark A. Neerincx. Situated cognitive engineering for crew support in space. *Personal and Ubiquitous Computing*, 15(5):445–456, 2011.

[15] Mark A. Neerincx and Jasper Lindenberg. Situated cognitive engineering for complex task environments. *Naturalistic decision making and macrocognition*, pages 373–390, 2008.

[16] Jakob Nielsen. *Usability engineering*. Elsevier, 1994.

[17] International Institute of Business Analysis. *A Guide to the Business Analysis Body of Knowledge (BABOK Guide), Version 2.0*. International Institute of Business Analysis, 2009.

[18] James W Varni, Bradley H Curtis, Linda N Abetz, Kathryn E Lasch, Elisabeth C Piault, and Andrea A Zeytoonjian. Content validity of the pedsql 3.2 diabetes module in newly diagnosed patients with type 1 diabetes mellitus ages 8–45. *Quality of Life Research*, 22(8):2169–2181, 2013.

[19] James W Varni, Michael Seid, and Paul S Kurtin. Pedsql 4.0: Reliability and validity of the pediatric quality of life inventory version 4.0 generic core scales in healthy and patient populations. *Medical care*, 39(8):800–812, 2001.

[20] N. Wigdor, J. de Greeff, R. Looije, and M. A. Neerincx. How to improve human-robot interaction with conversational fillers. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 219–224, Aug 2016.

11 Annexes

11.1 Specifying and testing the design rationale of social robots for behavior change in children

Bibliography Looije, R., Neerincx, M. A., & Hindriks, K. V. (2016). Specifying and testing the design rationale of social robots for behavior change in children. Cognitive Systems Research.

Abstract We are developing a social robot that helps children with diabetes Type 1 to acquire self-management skills and routines. There is a diversity of Behavior Change Techniques (BCTs) and guidelines that seem to be useful for the development of such support, but it is not yet clear how to work out the techniques into concrete robot support functions and behaviors. The situated Cognitive Engineering (sCE) methodology provides guidance for the design and evaluation of such functions and behaviors, but doesnt provide a univocal specification method of the theoretical and empirical justification. This paper presents an extension of sCE: a formal template that describes the relations between support objectives, behavior change theory, design specifications and evaluation outcomes, called situated Design Rationale (sDR) and the method to get this. As test case, the European ALIZ-e project is used to instantiate this design rationale and to evaluate the usage. This case study showed that sDR provides concrete guidance (1) to derive robot functions and behaviors from the theory and (2) to designate the corresponding effects with evaluation instruments. Furthermore, it helps to establish an effective, incremental and iterative, design and evaluation process, by relating positive and negative evaluation outcomes to robot behaviors at the task and communication level. The proposed solution for explicating the design rationale makes it possible for others to understand the decisions made and thereby supports replicating experiments or reusing parts of the design rationale.

Relation to WP This paper describes the design methodology of the PAL project.

Availablity Unrestricted. Available for download author copy (`https://drive.google.com/open?id=0BxB4LoFiKK5hX1VyNlB5bk5lOFU`)

11.2 Integrating Robot Support Functions into Varied Activities at Returning Hospital Visits

Bibliography Looije, R., Neerincx, M. A., Peters, J. K., & Henkemans, O. A. B. (2016). Integrating Robot Support Functions into Varied Activities at Returning Hospital Visits. International Journal of Social Robotics, 8(4), 483-497.

Abstract Persistent progress in the self-management of their disease is important and challenging for children with diabetes. The European ALIZ-e project developed and tested a set of core functions for a social robot that may help to establish such progress. These functions were studied in different set-ups and with different groups of children (e.g. classmates at a school, or participants of a diabetes camp). This paper takes the lessons learned from these studies to design a general scenario for educational and enjoying childrobot activities during returning hospital visits. The resulting scenario entailed three sessions, each lasting almost one hour, with three educational childrobot activities (quiz, sorting game and videowatching),two intervening childrobot interactions (small talk and walking), and specific tests to assess the children and their experiences. Seventeen children (age 610) participated in the evaluation of this scenario, which provided new insights of the combined social robot support in the real environment. Overall, the children, but also their parents and formal caregivers, showed positive experiences. Children enjoyed the variety of activities, built a relationship with the robot and had a small knowledge gain. Parents and hospital staff pointed out that the robot had positive effects on childs mood and openness, which may be helpful for self-management. Based on the evaluation results, we derived five user profiles for further personalization of the robot, and general requirements for mediating the support of parents and caregivers.

Relation to WP This evaluation provided knowledge on doing evaluations with (diabetic) children and provided input for improving and focusing the PAL system.

Availablity Unrestricted. Available for download (`http://dx.doi.org/10.1007/s12369-016-0365-8`)

11.3 MyPAL: A Digital Diabetes Diary with a Responsive Social Avatar

Bibliography Mike Ligthart. (2016). MyPAL: A Digital Diabetes Diary with a Responsive Social Avatar, Masterthesis Radboud University Nijmegen.

Abstract Diabetes Mellitus type I is an incurable disease that can be diagnosed at a young age. A structured lifestyle, where insulin use, carbohydrate intake and blood glucose are regularly monitored, is the only path to a relatively normal life. Children and their parents must remain vigilant.

This lifestyle is especially demanding for children and not every child is as good in self-management as they need to be. They can use some help with this. MyPAL is a digital diabetes diary that children can use to record their insulin use, carbohydrate intake and blood glucose values as well as write something about their day and how they feel. With that information the children can more easily link their diabetes values, what they eat and how they feel together. With this insight they can manage their insulin use and diet more efficiently. Besides children also medical professionals, parents and researchers benefit from this information. For example, a diabetes nurse can improve the treatment plan, parents can get a better idea how their child is doing and researchers can investigate the relationship between food, mood and blood glucose values more closely

Relation to WP This internship report and evaluation provides a design guideline and evaluation for the MyPAL diary/timeline.

Availablity Unrestricted. Available for download (`https://doi.org/10.5281/zenodo.268913`)

11.4 Report of the Cycle 1 May Experiments

Abstract An overview of the data from the year 1 experiment

Relation to WP Detailed description of all questionnaire data in Y1 cycle.

Availablity Restricted.

11.5 PAL Cycle 1 experiment child usage results

Abstract An overview of the data from the year 1 experiment

Relation to WP Detailed description of data children participating in Y1 cycle.

Availablity Restricted.

11.6 Analysis of system usage and knowledge development of the current PAL system for children with Type 1 Diabetes Mellitus

Bibliography Anika Boelhouwer. (2016). Analysis of system usage and knowledge development of the current PAL system for children with Type 1 Diabetes Mellitus, Masterthesis Radboud University Nijmegen.

Abstract Children who are diagnosed with type 1 diabetes mellitus need to learn a lot about diabetes and selfmanagement in a short period of time. A large problem in the support of this process is that health institutions cannot provide help at any given moment in the child life and are bounded by set face to face appointments. While digital interventions may address this issue by providing help and knowledge online which may be used at all times, this help and knowledge is general and not tailored to the individual. Also, actual usage of (digital) diabetes interventions has shown to be either extremely low or quickly decreasing. The Personal Assistant for a healthy Lifestyle project (PAL) strives to address these issues by providing a digital application with personalised communication and content. This study evaluated the current PAL application during a prolonged period of time with children diagnosed with type 1 diabetes mellitus between the ages of 6 and 12 years old. The main goals were to identify trends and possible predictors for both system usage and diabetes knowledge development. Three main trends were found in the system usage in which the majority of the users showed an overall low usage or quickly decreasing usage. A small number of users showed continuous and consistent usage throughout the entire experiment. As the personalisation was only minimally implemented the results are in line with common (digital) diabetes interventions. The results did not allow us to explore possible system usage and knowledge development predictors. They do however provide a solid baseline for further versions of the system in which the personalisation is further implemented. The main recommendations are to focus on the implementation of basic game design elements and personalised content to foster user engagement and continuous use. Maintaining the used measures (while adding some psychological predictors) and longitudinal experiment design will allow for comparative analysis in the further research cycles.

Relation to WP This internship report provides the usage results of the PAL Cycle 1 system

Availablity Unrestricted. Available for download (`https://doi.org/10.5281/zenodo.268909`)

11.7 Healthcare professionals gain control of childrens diabetes self-management

Bibliography Jet Shin Hong. (2016). Healthcare professionals gain control of childrens diabetes self-management, Masterthesis University of Amsterdam.

Abstract Children aged 8 to 12 with diabetes type I are motivated to get involved in their diabetes management to reduce the impact of their illness

on their short- and long-term health. Self-management of diabetes is an active and proactive process and it involves shifting and sharing responsibility for diabetes care tasks and decision-making in frequent collaboration with healthcare professionals. The research question this study sought to answer is: How can a healthcare management tool support healthcare professionals in guiding children with diabetes self-management involving a social actor (robot/avatar)?. To answer this question, a prototype of a healthcare management tool was developed and evaluated with end users (diabetes nurses) and an important stakeholder (diabetes doctor), following the situated Cognitive Engineering approach. Overall, this prototype of a redesigned PAL Control was perceived positively by the healthcare professionals and the findings suggested that a combination of an assessment with a robot or its avatar, setting goals, selecting actions and the progression page, is a suitable and effective approach to healthcare professionals in guiding children with diabetes selfmanagement. Healthcare professionals mentioned that this system has provided them support in making the consult with children and parents more meaningful due to the fact that they can understand their needs better on forehand. However, evaluations for a longer period of time is needed in order to validate if the needs are completely fulfilled. Nonetheless, useful suggestions were found during the evaluation of the prototype and provided important pointers for further development

Relation to WP This internship report provides the user requirements of the control panel for the health care professionals based on their initial use during cycle 1.

Availablity Unrestricted. Available for download (`https://doi.org/10.5281/zenodo.268911`)

11.8 Agents Sharing Secrets Self-Disclosure in Long- Term Child-Avatar Interaction

Bibliography Franziska Burger. (2016). Agents Sharing Secrets Self-Disclosure in Long- Term Child-Avatar Interaction, Masterthesis Radboud University Nijmegen.

Abstract A key challenge in developing companion agents for children is keeping them interested after novelty effects wear off. Self-Determination Theory posits that motivation is sustained if the human feels related to the agent. According to Social Penetration Theory, such a bond can be welded through the reciprocal disclosure of information about the self. As a result of these considerations, we developed a disclosure dialog module to study the self-disclosing behavior of children in response to that of a virtual agent. The module was integrated into a mobile application with avatar presence for

diabetic children and subsequently used by 11 children in an exploratory field study over the course of approximately two weeks at home. It was found that the relative amount of disclosures that children made to the avatar was an indicator for the relatedness children felt towards the agent at the end of the study. Girls were significantly more likely to disclose and children preferred to reciprocate avatar disclosures of lower intimacy. No relationship was found between the intimacy level of avatar disclosures and child disclosures. Particularly the last finding contradicts prior child-peer interaction research and should therefore be further examined in confirmatory research.

Relation to WP This internship report provides the model and evaluation of self-disclosure in interaction between avatar and child.

Availablity Unrestricted. Available for download (`https://doi.org/10.5281/zenodo.268912`)

11.9 A Disclosure Intimacy Rating Scale for Child-Agent Interaction

Bibliography Franziska Burger. (2016). A Disclosure Intimacy Rating Scale for Child-Agent Interaction. In: Traum D., Swartout W., Khooshabeh P., Kopp S., Scherer S., Leuski A. (eds) Intelligent Virtual Agents. IVA 2016. Lecture Notes in Computer Science, vol 10011. Springer, Cham

Abstract Reciprocal self-disclosure is an integral part of social bonding between humans that has received little attention in the field of human-agent interaction. To study how children react to self-disclosures of a virtual agent, we developed a disclosure intimacy rating scale that can be used to assess both the intimacy level of agent disclosures and that of child disclosures. To this end, 72 disclosures were derived from a biography created for the agent and rated by 10 university students for intimacy. A principal component analysis and subsequent k-means clustering of the rated statements resulted in four distinct levels of intimacy based on the risk of a negative appraisal and the impact of betrayal by the listener. This validated rating scale can be readily used with other agents or interfaces.

Relation to WP This paper provides the model and evaluation of self-disclosure in interaction between avatar and child.

Availablity Unrestricted. Available for download (`https://doi.org/10.1007/978-3-319-47665-0_40`)

Specifying and testing the design rationale of social robots for behavior change in children

Rosemarijn Looije[a,*], Mark A. Neerincx[a,b], Koen V. Hindriks[b]

[a] *TNO, Postbus 23, 3769ZG, Soesterberg, The Netherlands*
[b] *Interactive Intelligence Group, Delft University of Technology, Mekelweg 4, 2628 CD, Delft, The Netherlands*

Abstract

We are developing a social robot that helps children with diabetes Type 1 to acquire self-management skills and routines. There is a diversity of Behavior Change Techniques (BCTs) and guidelines that seem to be useful for the development of such support, but it is not yet clear how to work out the techniques into concrete robot support functions and behaviors. The situated Cognitive Engineering (sCE) methodology provides guidance for the design and evaluation of such functions and behaviors, but doesn't provide a univocal specification method of the theoretical and empirical justification. This paper presents an extension of sCE: a formal template that describes the relations between support objectives, behavior change theory, design specifications and evaluation outcomes, called situated Design Rationale (sDR) and the method to get this. As test case, the European ALIZ-e project is used to instantiate this design rationale and to evaluate the usage. This case study showed that sDR provides concrete guidance (1) to derive robot functions and behaviors from the theory and (2) to designate the corresponding effects with evaluation instruments. Furthermore, it helps to establish an effective, incremental and iterative, design and evaluation process, by relating positive and negative evaluation outcomes to robot behaviors at the task and communication level. The proposed solution for explicating the design rationale makes it possible for others to understand the decisions made and thereby

[*] Please address correspondence to Rosemarijn Looije

Email addresses: rosemarijn.looije@tno.nl (Rosemarijn Looije), mark.neerincx@tno.nl (Mark A. Neerincx), k.v.hindriks@tudelft.nl (Koen V. Hindriks)

Preprint submitted to Cognitive Systems Research *http://dx.doi.org/10.1016/j.cogsys.2016.07.002*

supports replicating experiments or reusing parts of the design rationale.

Keywords: Social robot, Cognitive engineering, Design rationale, Diabetes

---

1. Introduction

There is a need for social robot design methods, which provide theoretically and empirically founded implementations that can be systematically reused, compared and built upon progressively (cf., [1]). Current design methods do not (yet) meet these needs, holdings back the coming of age of the research field.

This paper focuses on the development of robots for behavior change. Although there is a substantial amount of research in social robots and behavior change techniques, it is hard to compare the results of studies due to a lack of agreement on (1) the (definitions of) relevant theoretical concepts, (2) the design specifications, (3) the methods for validation (or evaluation), and (4) the approach to relate these concepts, specifications and methods. Literature from the social robot domain on classification of robots (e.g. [2]) and evaluation (e.g. [3]) provides valuable information for design specifications and their evaluation. However, it is unclear how they relate and can be linked to behavior change theories. On the other hand, for behavior change techniques there is a taxonomy in development [4] which supports disambiguation of results, and therefore validation of effective techniques, but it does not relate these to design specifications (such as use contexts). Use contexts are taken into account in the research of Behavior Change Support Systems (BCSS), for instance in the persuasive systems design (PSD) model [5]. This model emphasizes the translation between method and design patterns for functionalities related to the context. Although method, requirement, design and implementation are related in PSD, it does not model the correlations and interrelations between different implementations.

An open question remains: "How can we conduct experiments in such a manner that it will be really possible to pinpoint a change to have been caused by a BCSS, or even more precisely, by a specific software feature in it?" [6] . Our social robot is in essence a BCSS and the question we want to answer is quite similar:

- How can we design and evaluate in such a manner that a) robot behaviors are derived from theory and b) evaluation effects can be designated to specific robot behaviors?

The situated Cognitive Engineering (sCE) methodology [7] can partially answer this question. sCE has been used in different domains, amongst which to systematically design and evaluate robot systems [8]. Although sCE supports iterative and incremental design and evaluation, it does not provide precise and concise translations and relations between the theory, functionalities of the system, hypotheses and instruments to evaluate (i.e. the concepts).

The situated Design Rationale (sDR) was developed as a refinement of the sCE methodology. This formal template supports the design of functionalities, the planning and performance of evaluations, and makes it possible to reason about the evaluation effects and decisions afterwards. To come to this formal template, we distinguish three sub-questions all in the context of the development of a social robot for supporting behavior change:

1. Which minimal set of concepts is needed to describe the what, when and why of design decisions?
2. How do these concepts relate to each other?
3. What is an adequate, concise and coherent, representation for describing the concepts and its relations for the design and evaluation process?

The research took place in the context of the development of a social robot that provides self-management support for children with diabetes (i.e., the European ALIZ-e project[1]). The structure of this paper is as follows: First in section 2, we provide background on diabetes, social robotics, behavior change and situated Cognitive Engineering. Second in section 3, we describe the sDR template, that describes the concepts and it relations, followed by the instantiation of sDR in section 4. In section 5 the use of the sDR is further exemplified with an experiment performed within the ALIZ-e project. And we finish with the conclusions and discussions on future work in section 6.

2. Background

Type 1 diabetes has an enormous impact on the daily life of children with this illness as we will discuss in section 2.1. There is a need for support of self-management and behavior change. A social robot might provide this support for this user group (age 7-12) (section 2.2). The behavior of the robot should

[1] `www.aliz-e.org`

be based on knowledge from behavior change theories and systems (section 2.3), and the design of the robot should be based on a state-of-the-art design methodology (section 2.4). Based on this background we can conclude what is lacking to come to a concise and precise situated Design Rationale.

2.1. Type 1 Diabetes Mellitus

To understand why we want to develop a social robot to support children with diabetes to increase their self-management it is necessary to understand what diabetes is and what this means for the life of the children, and their environment. There are two types of diabetes, Type 1 and Type 2 [9]. Type 1 typically presents itself at a young age, while Type 2 often occurs at a later age. Where Type 1 Diabetes Mellitus (T1DM) is a result from destruction of the insulin-producing cells in the pancreas by the autoimmune system, Type 2 is a metabolic disorder where the body does not make and absorb enough insulin. We will further focus on T1DM, because that is the type that is most prevalent in children and the incidence is rising [10]. For these children it is very important to keep their blood glucose levels as steady as possible. To reach this objective, children and their social environment (parents, teachers, siblings, friends etc.) need to have knowledge and skills to manage the disease. Examples of these are: Regularly measuring of blood glucose, counting of carbohydrates, calculating needed insulin and injecting (when pen is used) or bolusing (when pump is used) accordingly, and discounting the (interactive) effects of food intake, physical exercise, mental stress and hormones. Furthermore, a child and his or her environment need to be able to recognize symptoms of high and low blood glucose to act accordingly. Even when managed properly, a child will have periods of high imbalance due to for instance hormones or growth spurts. The effects of T1DM, even with our modern treatment, are quite severe. More than 50% of the children develop complications with regard to major organs like the heart and blood vessels 12 years after diagnosis [11]. The life expectancy of children diagnosed by age 10 is 19 years shorter than that of healthy children [12]. There are also effects on psychological well-being, feelings of embarrassment and on school performance [13]. The effects on psychological well-being are not limited to the children themselves, but also their parents are hugely influenced, because they understand the long-term effects better than a (young) child [14]. Other research suggests that high family stress negatively affects glycemic control [15]. To lower family stress it is important that children learn to manage their illness at a young age and that parents let them do this. A social robot can support in this, because it has a non-hierarchical relation with the child unlike a (in)formal caregiver. A social robot for changing the behavior of and/or educating children is not new as is shown by [16, 17] where they are applied for autistic children, [18, 19, 20] for education and [21] to acquire a healthy lifestyle. Aspects of behavior change and motivational theories can be, dependent on the features and form of the robot, implemented on the robot and applied to improve self-management.

2.2. Social robots

Below we provide a short overview of design and evaluation methods that are used in the field of personal social robotics on context, behaviors, appearances, and effects. We exclude work-oriented human-robot interaction (e.g., human-robot teamwork; [22, 23, 24]), because we focus on (non-work) social settings of the child. Robots can be classified according to their appearance (from mechanical to human-like for instance [25]) and their behavior. Bartneck et al. [26] for instance classify social robots on five factors: Form (abstract - anthropomorphic), modality (unimodal - multimodal), social norms (no knowledge on social norms - full knowledge on social norms), autonomy (no autonomy - full autonomy) and interactivity (no causal behavior - fully causal behavior). Fong et al. [1] provide a more elaborate classification specifically for socially interactive robots, robots for which social interaction plays a key role. First they identify two primary approaches to build socially interactive robots, biologically inspired or functionally designed. Decisions on the design and evaluation need to take the context into account. Fong et al. further identify other aspects that can be used to classify robots, e.g. embodiment, emotion, dialogue, personality, perception of humans, user modeling, socially situated learning and intentionality. It is meant as support for people designing socially interactive robots to make decisions on the form and behavior of the robot for the use in a specific context. This is further explained by providing different applications and examples of robots used in every application and a short indication of what aspects of the classification they adhere to. Dautenhahn [27] looks at the aspect of consistency of design and behavior. Examples are provided what happens when it is not consistent (e.g. very humanlike appearance of robots induces the uncanny valley effect, because it cannot perform as expected), but reaching consistency seems to be a matter of trial and error. With the design space provided it is possible however to place robots on the two dimensional axis of appearance (machine like vs. human like) and behavior (non-social and non-interactive vs. social and interactive). Spiekman et al. [28] uses the axis of machine to human like, next to an indication of toy like, body and facial realism to categorize and evaluate 3 robots (iCat, NAO and Nabaztag) and a human-like avatar. These different ways of classifying (social) robots shows that designers of robot systems make many choices, and these choices should be formalized to understand why these choices were made and also decide on the validity of the choices after evaluation.

It is important for comparability between different robot designs to measure the same type of effects and preferably also use the same measures. Weiss et al. [3] propose to use the following evaluation factors: Usability, social acceptance, user experience and societal impact. Which factor to use depends on the hypotheses. Furthermore, they propose, for the evaluation of hypotheses, to use a mix of interdisciplinary evaluation methods: Expert evaluation, user studies, (standardized) questionnaires (e.g. unified theory of acceptance and use of technology (UTAUT) questionnaire [29]), physiological measures, focus groups and interviews. Bartneck et al. [30] provides an instrument toolkit to measure how users perceive a robot on five factors relevant for HRI: Anthropomorphism, animacy, likeability, perceived intelligence and perceived safety. They developed five validated questionnaires for these five factors. These questionnaires are all relevant for evaluating the design of a social robot, but do not provide measures that are related to the objective of the robot use, e.g. education.

2.3. Behavior change

Behavior change is a large research field. We will focus on two topics: A taxonomy developed to describe behavior change methods and a model to design persuasive systems for behavior change. The taxonomy is interesting, because it is an effort to describe components of a behavior change method in a way to derive effectiveness in a similar way we want to describe the components of the robot. The persuasive systems model is of added value, because we also want to create a persuasive system, where we use the robot as ICT component.

Interventions to change behavior are complex and have many interacting components [31]. Therefore, the same problems occur as in social robot research: Research outcomes are hard to replicate, to implement in practical applications and to use for building theory [4]. We therefore need a better understanding of which components are effective within a behavior change intervention.

A first step is to get a common understanding of the components in an intervention. This helps in recognizing overlap between different interventions and identifying effective components. In Michie et al. [4] a hierarchically structured taxonomy of behavior change techniques (BCTs) is construed with the help of 55 experts in delivering and/or designing behavior change interventions from different countries. This resulted in 93 BCTs that were clustered in 16 groups of which 26 were used 5 or more times in different interventions. An example of a group is “Reward and Threat” covering seven BCTs (e.g., material reward, threat, incentive).

A selection of BCTs can be implemented in a social robot where the social robot is used instead of, or as a complement of, a human. The robot can be viewed as the IT artifact of a behavior change support system (BCSS). BCSS is defined by Oinas-Kukkonen [6] as a socio-technical information system with psychological and behavioral outcomes designed to form, alter or reinforce attitudes, behaviors or an act of complying without using coercion or deception. A BCSS is a complex system that is developed using theories of behavior change and persuasive technology by explicating functionalities of a system.

To support the design of a BCSS, Oinas-Kukkonen suggests the use of the Persuasive Systems Design (PSD) process. The design of a BCSS takes postulates from User Centered Design which are also used in persuasive design (e.g., ease of use), uses these in context (intent, event and strategy) and then a decision on the design of software features needs to be made. During the context step the intended outcome is decided on, using the outcome & change design matrix, which also influences the strategies. The combination of the PSD process and the outcome & change matrix provides a way of defining the system, context and intent clearly. This is necessary because these influence the outcomes, e.g., different IT systems will be able to implement persuasive strategies on different levels.

The behavior change literature provides objectives and methods that can be used to guide implementation of a social robot for behavior change. The PSD model guides the design of a BCSS by relating functions to behavior change techniques and always keeping the intended outcome in mind. The design thus takes as a starting point the intended outcome, but due to a lack of formalization between design decisions and evaluation measures the PSD model cannot pinpoint the effects to specific functions. This is also explicitly indicated by Oinas-Kukkonen who sees this as one of the open questions on the BCSS research agenda [6].

2.4. Situated cognitive engineering

The situated Cognitive Engineering (sCE) [32] methodology has its main strengths in the analysis of three system development components: the foundation, specification and evaluation. It has been applied, for example, in the domain of behavior change [33] and robots [8]. In sCE functions are incrementally developed. It can be viewed as a refinement of classical cognitive engineering methods [34, 35, 36], addressing the reciprocal adaptive behaviors of both human and machine (i.e., emergent human-machine cooperation patterns).

The classical methods are mostly focused on a thorough domain and task analysis, but the sCE method adds explicitly technology and human factor knowledge (methods, instruments) to establish a sound *foundation*. Technology is added for two reasons. First it provides focus in the process of specification and generation of ideas. Second, the effects of technology are made explicit and are integrated into the development and thereby evaluation process. The explicit use of human factors knowledge, e.g. knowledge on developmental age, behavior change, education and so forth, supports the development and the embedding of functions and experimental results in theories. Moreover, the sCE method is situated in a domain that is made explicit in use cases that contextualize the (robot) functions. The explication from foundation (e.g., tasks analysis, value sensitive design) to specification is guided by use cases.

The *specification* component encompasses, among other things, functions (requirements), interaction design patterns, use cases and expected effects (claims). Key (recurring) functions are shaped in interaction design patterns (i.e., the "look-hear-and-feel" of robot behaviours) and applied in specific use cases (i.e., contexts). The functions are justified by the expected effects.

In the *evaluation* components, experiments test the expected effects (claims) and provide guidelines about what to use and when to use it. As such, the results of the evaluation also provides input for theory development.

Our research aims the development of a social robot with the *objective* to enhance child's self-management by applying different behaviour change *methods* as the theoretical foundation, and to establish the empirical foundation via sound evaluation *instruments* that show how far this *objective* has been achieved. We have to explicitly relate the sCE concepts to these objectives, methods and instruments (see Figure 1) in order to reason about the design decisions made. The sCE method does insufficiently support this type of reasoning. There are for instance no explicit relations between a specific method and therefore objective and a function. Of course use cases take the objectives into account, but the relation is not made explicit. Furthermore, the expected effects are related explicitly to the functions and the instruments, but the interrelations between expected effects and functions are not made explicit. One function can have multiple effects, an effect can be related to different functions, multiple instruments can be used to measure the same effect, but it can also happen that one instrument measures multiple effects. These relations need to be explicated so that we can disambiguate the design and evaluation as much as possible by refining it, e.g. more specificity in instruments. Disambiguation will not always be possible, but explicating all relations makes it possible to see where there are still ambiguous relations. Knowing these ambiguities can guide further design and evaluation.

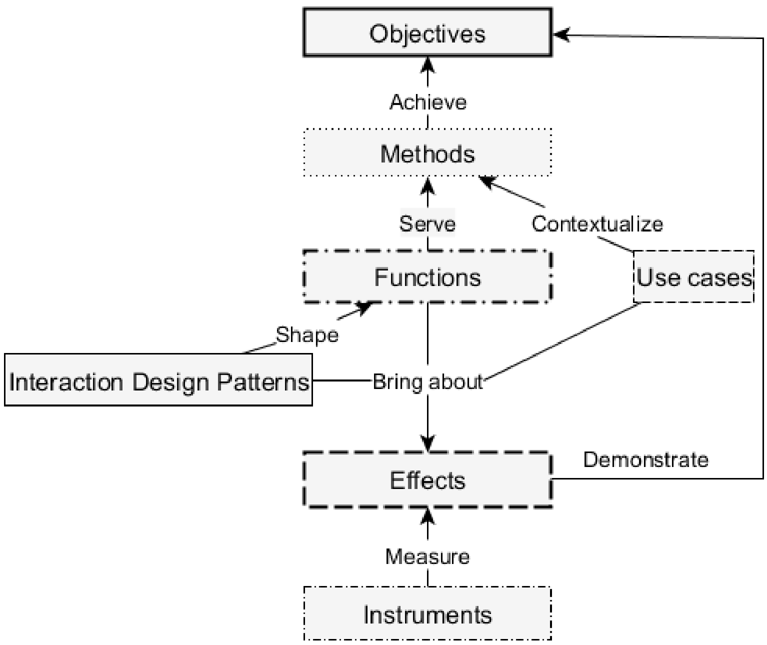


Figure 1: Generic concept map of the situated Design Rationale (sDR).

3. Situated design rationale

To create a situated Design Rationale (sDR) that specifies the relations between functional aspects and expected effects in a manner such that we can reason about the design decisions made and the interactions between effects and functions, we extend the sCE method. The concepts come from the sCE method and some of the relations also, but we add relations to make all relevant relations explicit in an sDR.

3.1. Concepts

In the previous section we distinguished the relevant concepts that have to be related to each other. The first relevant concept is *objectives* (e.g. support the forming of a relation between robot and user), second are *methods* (e.g. adapt the robot to the user's behavior) that are derived from literature or experiments to reach these objectives. The methods then have to be translated into, the third concept; *functions* (e.g. adapt the robot system to the state of the child) of the robot. The functions are shaped by, fourth concept, *interaction design patterns* (e.g. use of prosody to express emotions by the robot). Fifth, *use cases* (e.g. a quiz between the robot and child in which they act as peers) are used to contextualize the methods and show which *effects* (sixth concept) we expect towards the objectives (e.g. children relate more to the emotional expressive robot). But also the effects in relation to the implemented functions and design patterns are described (e.g. an expressive robot supports emotional contagion - i.e. the child is more expressive, emotions are recognized). Seventh and last *instruments* are then used to measure these effects (e.g. arousal and valence observations by the child). In Figure 1 the seven generic concepts and their relations are shown. In the following paragraph we explain how the generic sDR is developed.

3.2. Situated design rationale template

The situated Design Rationale is developed to support design of functionalities and evaluation before an experiment and reason about the effects and decisions afterwards. The explication from theory (objectives and methods) to functions and then to effects should thus be made clear. To make this possible we have to relate the concepts to each other, and as is said "a picture is worth a thousand words" we decided to use concept mapping [37] as a tool to describe the relations. In a concept map, relations between ideas, images, or words are linked with meaningful arrows. In our case the meaningful words are the concepts and the meaningful arrows the relations between the concepts[2].

The *objectives* come from the foundation of sCE, relevant theories (behavior change, education) are taken into account as well as knowledge on human factors (what are the capabilities of a child in the age group 7-10) and technology (what are the robot (in)capabilities) to come to a selection of *objectives*. Also based on the foundation of sCE *methods* are selected to **achieve** the *objectives* and which are supported by literature or derived from empirical experiments (e.g. provide variation, which supports competence and comes from educational theory). *Use cases* are then described to **contextualize** the *methods* and to show which *effects* are **brought about**. *Functions* are related to the *methods*. Only *functions* that **serve** a method are relevant here. In some cases, explicating the relation between *method* and *function* is quite straight forward. An example of this is a *method* that prescribes variation and a *function* "Provide multiple activities". *Functions* are **shaped** into *interaction design patterns*. An example of this is the *interaction design pattern* "Recognizable emotion expression" that supports the higher level *functions* "Exhibit social behavior" and "Adapt robot to child state within boundaries". The *interaction design pattern* **shapes** the function and defines what is needed to reach, in this case, "Recognizable emotion expression".

Then we specify the *effects* that the *interaction design patterns* and the *functions* **bring about**. This is a very important step. If a *function* cannot be related to an *effect* it should **bring about**, that *function* or *interaction design pattern* has to be reconsidered. The reason for this is that the relation between *functions, patterns* and *effects* is also the relation back towards the *objectives*. The *effects* **demonstrate** the result on *objectives*. An equally important relation is that from *effects* to *instruments* that **measure** the *effects*. When there is no *instrument* to **measure** an *effect*, the *effect* might be too specific or generic. The design is also guided by this step, because when there is one *instrument* that is used to **measure** many *effects* the results cannot be used to disambiguate between different *functions*. Therefore, either the *effects* have to be made more specific, or the *functions* need to be made more distinguishable from each other so that there is less ambiguity between the *effects*.

[2]Using yEd `https://www.yworks.com/en/products/yfiles/yed/` we created a general concept map of sDR 1 in which the concepts and their relations are visualized

When there is a first complete version of the sDR, it has to be checked and decided on what will be the focus of an experiment. The sDR can support deciding where experiments are needed to get more information, but also review the *instruments* to see if they are specific enough to derive conclusions from the results. The results can then be used to reason about the decisions made and refine and extend the sDR. Figure 1 provides a generic sDR, which we will instantiate using an experiment performed within the ALIZ-e project in the next section.

It's interesting to see the similarities between Worth Mapping [38] and sDR. Both take into account the values of the end users; in Worth Mapping these are the objectives of the design while in sDR these are part of the methods to reach the objectives and used to enrich the use cases. To satisfy the values both identify needed elements or methods and functions to reach a worthwhile outcome. This means that Worth Mapping guides the interaction design by making relations between values, elements and attributes clear, while sDR makes the transition to context and effects. sDR uses the values and attributes to describe the use cases and contextualize the methods which in its turn constrains the functions and interaction design patterns. The measured effects then demonstrate the progress towards the objectives, but also if user values are met.

4. Instantation of a sDR

We will now show how sDR can be used to describe the design and evaluation activities of the ALIZ-e project by instantiating the concepts with specific examples. We do this by going through the concepts, explaining decisions and showing parts of the sDR to exemplify the concepts. The complete sDR of the ALIZ-e project can be found here: `https://goo.gl/0HgUC8`.

In the complete sDR there are many intersecting lines, in a limited way this is also the case in the figures presented in this paper. As this problem can not be eliminated we used different arrows to make clear what the origin of lines are. In Figure 6 we added the outgoing arrow form to the text of the functions.

4.1. Objective

The overall objective of ALIZ-e is behavior change for self-management, with a focus on children with diabetes. The objective is thus behavior change and a decision needs to be made on which theory we will use to relate our progress to.

4.1.1. Choice for behavior change objective

Many theories for behavior exist, and the choice of one over the other guides the priority of objectives. We will briefly discuss Theory of Reasoned Action II [39], the Extended Parallel Process Model [40] and the Self-Determination theory [41].

In the Theory of Reasoned Action II (TRA II) behavior is determined by intention, which is determined by attitude, perceived norm and perceived behavioral control (similar to self-efficacy). Actual control is determined by environmental factors and skills to deal with these.

The Extended Parallel Process Model (EPPM) argues that changing behavior, attitude and intention results from an attempt to control threat, while not changing behavior comes from fear. According to EPPM people deal with threats and fear in three different ways. First, a threat can be seen as insignificant so there is no motivation to change. Second, a threat can be perceived as so serious they feel not able to deal with it, because they dont have enough perceived self-efficacy and response efficacy. The third option is that the threat is perceived as serious and they feel empowered to do something about it because of high self-efficacy and response efficacy.

The Self-Determination Theory is a motivational theory that supports a continuum of motivation, from external regulation (completely extrinsic) towards more and more internally motivated to end in intrinsic motivation [42]. The motivation can be influenced by supporting three basic psychological needs: autonomy, competence and relatedness. Autonomy is about the willingness to do a task, competence is the need for challenge and feeling of effectance, and relatedness refers to the connection with others [43]. Long-term interaction is seen as a prerequisite for behavior change in the long run and several behavior change methods endorse the reasoning that for long-term interaction there is a need for a bond with the interaction partner (e.g. Motivational Interviewing [44]).

All three example theories show differences, but also similarities (e.g. self-efficacy is important in all three). Because of these similarities and the complexity of these theories, there is an ongoing effort to analyze behavior theories until the level of behavior change techniques and then evaluate those on effect [4]. As a decision had to be made we chose Self-Determination Theory as our starting point (see objectives in Figure 2), because this theory

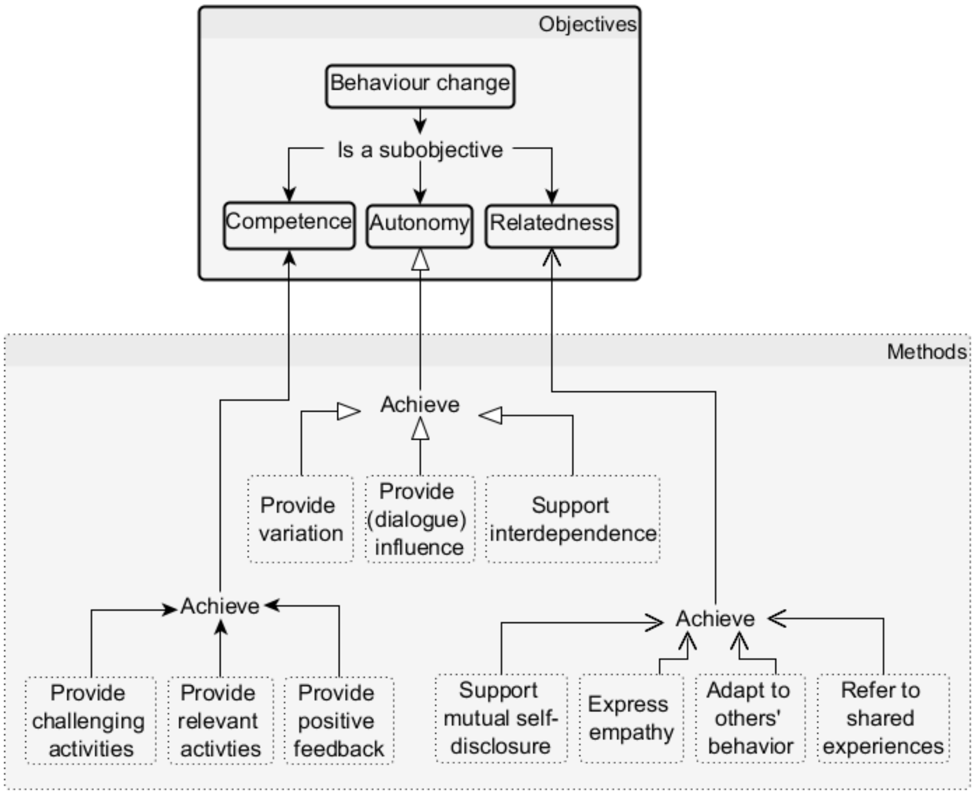


Figure 2: Objectives and methods that achieve them

is used not only in behavioral change but also in education [45], for children in the relevant age group (7-11) [46] and in games where it showed to be a predictor of enjoyment and future game play [47].

4.2. Methods

Another advantage of SDT is that there is an ongoing effort to connect the methodology of Motivational Interviewing (MI) to the theory of SDT [48]. Motivational Interviewing (MI) is a proven effective counseling style for promoting behavior change, but it is not grounded within a theoretical framework, SDT can provide this framework. MI techniques have also been used in persuasive technologies as the Health Buddy [49] and techniques from MI have been implemented in a social robot for adults with diabetes [50].

ID	Use case	Description
1	Competitive quiz with robot peer	The robot and child play a competitive Trivial Pursuit based quiz where they alternate in answering questions.
2	Collaborative sorting game with robot peer	The robot and child play a collaborative game on a large touch screen on which they have to swipe images, that are on the screen, to the correct categories (most of the time 2, that are on the left and right side of the screen).
3	Imitation memory game with robot peer	The robot makes a movement (e.g. arms up) and then the child imitates this and adds another movement, which the robot has to imitate. The string of movements gets longer and longer, so its both a movement and memory game. Variations are: that the robot can only add movements, some movements are prohibited, and there are different levels of sequences (more complex) and movements (more difficult).
4	Watching educational video with robot peer	Robot and child watch a video together.
5	Providing a combination of activities	Provide multiple activities as described in Use Case 1-4
6	Engaging in small talk with robot peer	Some interaction about hobbies, activities, friends, diabetes.
7	Support robot from one activity to another	The child has to help the robot from one activity to another, by walking with it (holding hands) or carry the robot.
8	Helping robot to stand up	When the robot falls over the child helps it in getting up.

Table 1: Overview of the ALIZ-e use cases.

To reach the objectives we can thus draw upon methods of MI, we further draw upon (amongst others) educational, gaming and persuasive methods and methods used for rapport building in human-human interaction. These

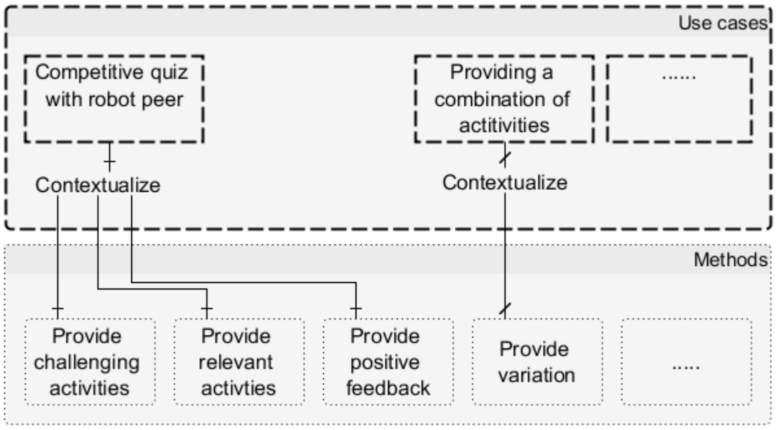


Figure 3: Use cases and how they contextualize methods.

methods are overlapping; for instance Vygotskys educational theory [51] and gaming theory [52] both endorse the importance of having challenging, relevant activities to support intrinsic (long-term) motivation. Literature supports the relation between challenging activities and self-determination theory [53]. Vygotsky and MI also state that the teacher/therapist should build up rapport with the student or client; in MI this is further elaborated in methods to build this up (e.g. express empathy). In Vygotsky the teacher can also be a peer in a collaborative learning sessions; the peers learn from each other and need each others help. In such a setting the rapport building will have another dimension than with a teacher/therapist, e.g. the shared experiences and matching the personal norm will be differently implemented. In Zhao et al. [54] an overview of methods to reach rapport is provided.

Figure 2 shows the methods used within the ALIZ-e project and their relations to the objectives. All methods come from literature; MI [48], educational [51, 55], gaming [52] and relation theories [54]. In some of this literature the methods are explicitly linked to SDT objectives (e.g. [48, 53]), other relations need to be derived.

As can be seen, there are three different objectives. These objectives are not completely unrelated, but all have their main focus which is depicted in the figure. The functions will connect the different methods to each other.

4.3. Use cases

The objectives, methods and (later on) related effects and measures wont change a lot during the course of a project. A method can be added, but as the objectives are the starting point these will be relatively stable. The choice for a method also guides the expected effects and with these the measures. This is different for the other concepts we discuss, the use cases, functions and interaction design patterns. The instantiations of these concepts will be refined and added on during the whole project. Within the ALIZ-e project we focused on developing a social robot for long-term interaction with children and as the domain we chose behavior change for improving self-management of children with diabetes. To further specify this setting, taking into account the knowledge on the domain and users, we created eight use cases over the course of the project (see Table 1) describing the interaction in more detail. For more information on how these use cases were incorporated in experiments we refer to [56], in which an experiment is described containing most of the use cases.

Each use case contextualizes the methods and provides situational context of the effects that are measured. The competitive quiz for instance contextualizes methods which focus on competence, while providing a combination of activities is related to provide variation (see Figure 3).

4.4. Functions

Based on the methods and use cases a selection of functions was implemented during the project. In Table 2 the functions used in the ALIZ-e experiments are named with a short exemplification next to it. We evaluated (parts of) these functions. Some of the more complex social behaviors like maintain social relationships are encompassed in for instance the function "personalize activities". Choosing the right level of function description is a bit of trial and error. We don't want the functions on implementation level, because this would complicate the picture sDR too much. The functions should be with enough detail to be able to relate them to specific methods and specific effects. You dont want the functions to encompass too little or too much, because the sum of the parts can be different than the sum of the whole. Some functions contribute to one method, others contribute to multiple methods. In Figure 4 this is shown, the functions related to the methods of Figure 3 are shown, but it is also shown that most functions are related to multiple methods and that these methods can be related to different objectives (see Figure 2).

Function	Exemplification
Personalize activities (based on personal info, performance, history etc.)	A game should be challenging and relevant, and small talk should be relevant
Provide multiple activities	The child should be able to switch between activities and the same objectives should be presented in different ways
Provide open questions	The child should have the opportunity to express him/herself
Disclose robot information	The robot should disclose personal information about itself, a background story
Adapt robot to child and activity within boundaries	The robot should adapt its emotions to child and activity state. Be happy together with child, but also a bit sadder when losing. Recognizable emotion expression is necessary for this.
Provide acknowledgements	The robot should acknowledge what the user is doing
Provide compliments	The robot should provide compliments to the user on its actions
Exhibit social behavior	The robot should behave according to standard social norms; look behavior, turn taking, use of natural (non-verbal) cues (e.g. thinking behavior- uhmmm and gestures)
Show imperfection	The robot should not be all knowing and also need the help of the child sometimes

Table 2: The different functions and an exemplification

4.5. Interaction design patterns

There are many interaction design patterns possible for the use cases we looked at in ALIZ-e, but as social behavior and the emotions that come with this are very important. We looked at this in more depth. We looked for instance at the recognition of robot emotion expression for different robots (iCat and NAO) [57] and at the effect of embodiment (virtual or physical) on the effectiveness of social behavior [58]. Figure 5 shows how the different aspects of the voice and body influence the emotion expression and thereby the social behavior.

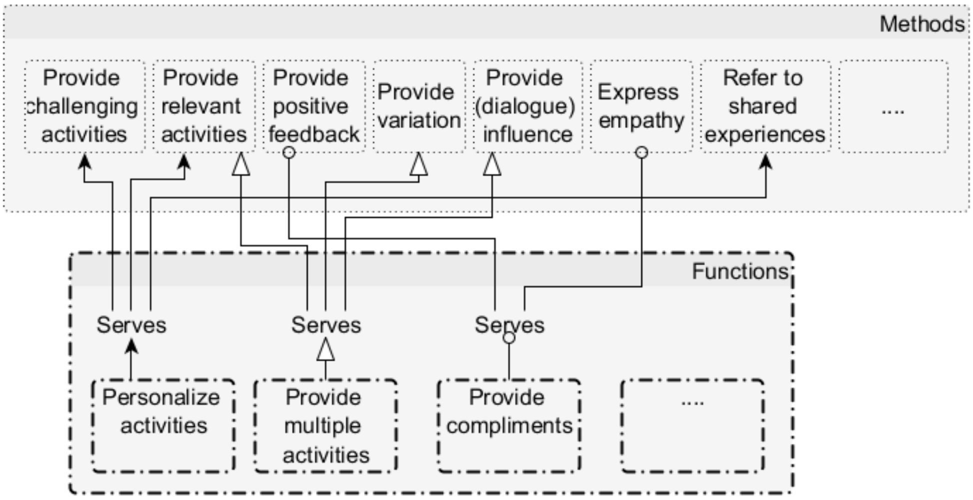


Figure 4: Methods and functions that serve them

4.6. Effects

The expected effects are derived from literature about the objective and used techniques and from the functional design. Both the up- and downsides of an implementation should be defined so that in an experiment it can be validated if the upsides outweigh the downsides. We identified three levels on which these up- and downsides can be reported within human-robot interaction (leaving out pure technical evaluation):

1. The child perceives and comprehends the intentions of the robot
2. The robot perceives and comprehends the intentions of the child
3. The situated Human-Robot interaction

Within the ALIZ-e project we looked at "perceive and comprehend 'intentions' of robot" (1) and "situated human-robot interaction" (3) in the experiments. The experiments on recognizable emotion expression were on level 1, while the situations where there was interaction with the robot during an activity (quiz, sorting game, small talk etc.) were on level 3. On level 1 the interaction design patterns are evaluated and on level 3 the functions. The effects of the interaction design patterns are related to the functions they shape and of course the interaction design patterns. The effects of the functions are not only related to the functions, but also to the methods and objectives where the expected effects are derived from.

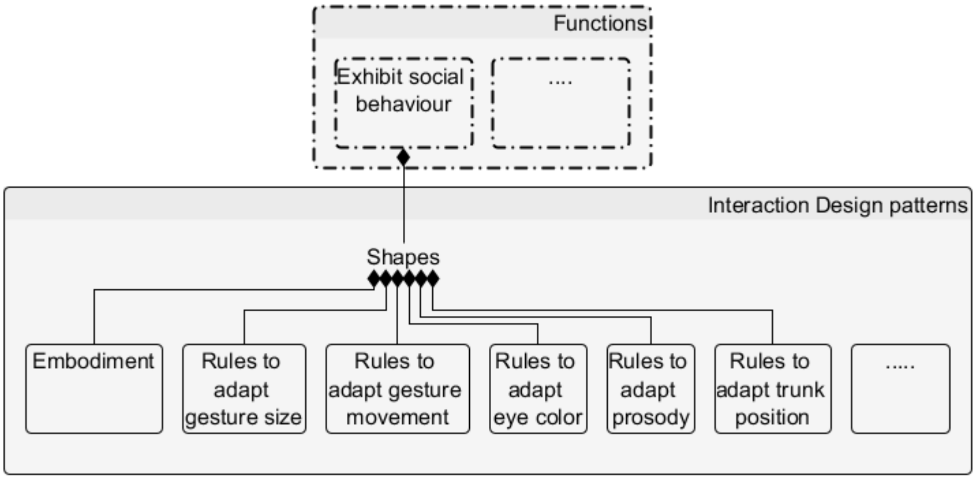


Figure 5: Interaction design patterns shape functions

In Figure 6 a selection of the effects, and their related functions and instruments are shown. The effects show a direct relation with the objectives as effects on competence, autonomy and relatedness are expected. Next to this it can be seen that it is expected that most of the implemented functions, even all for this specific set of functions, contribute to the acceptance, trustworthiness, enjoyment and the robot being seen as empathetic. This set of expected effects is derived from the objective relatedness, from which this set is derived as being important. The relation back to the objectives is not drawn to make the sDR not unnecessarily complex, as these relations can also be found going back in the sDR. The interaction design patterns relate to their specific effects directly and indirectly via the function it shapes. The rules to adapt prosody for instance has a direct effect on understandability and an indirect, together with other patterns that shape the social behavior, on for instance trust.

4.7. Instruments

After the expected effects are described there is a need to measure these. We prefer using objective instruments in combination with subjective instruments. Especially because it is known that children have the tendency to score extreme on questionnaires and there is thus a high chance on a ceiling effect. In Figure 7 it can be seen that although we would like to have ob-

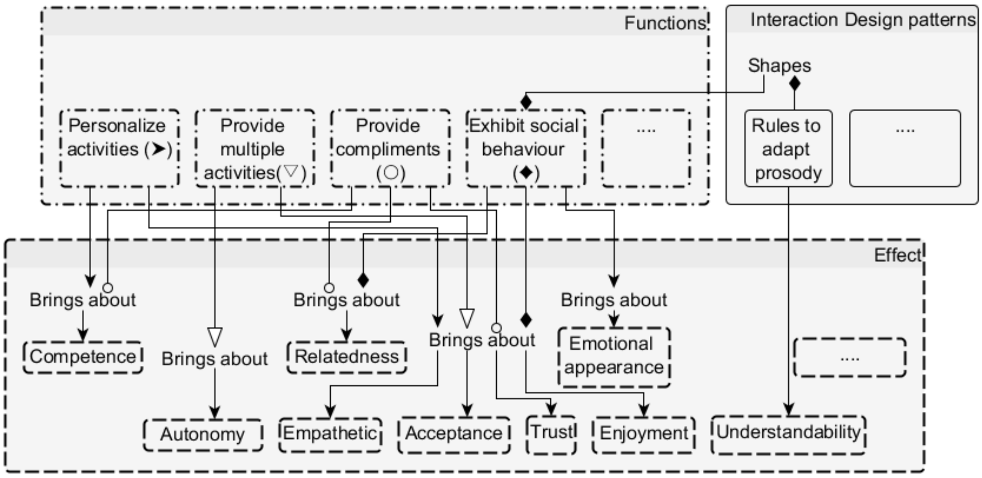


Figure 6: Functions and interaction design patterns bring about effects

jective measures, many are still subjective. Enjoyment is measured with a questionnaire and observations and emotional appearance and understandability both have questions for the child to check recognition of either emotion or spoken text of the robot. Having a forced choice question does eliminate some of the problems of a questionnaire, but it also means there is a need for a within subject design and this is not always feasible with specific user groups.

5. Evaluation of the sDR

The previous section described the sDR using the ALIZ-e project as an example. This section will show how a specific design and evaluation cycle can be supported by the creation of an sDR. In this cycle, a model for adaptive emotion expression for a NAO robot was developed. The robots internal valence and arousal values were influenced by emotional state of the child and emotional occurrences in the activity (e.g. winning the game). This adaptation of internal values led to a change in voice, posture, whole body pose, eye color and gestures to express its emotional state. In an experiment 18 children (mean age 9) played a quiz with two NAOs consecutively (within subject design). One of the NAOs adapted its emotions according to the model and the other did not. A more detailed description of the method is provided in [59]. The objective this experiment focused on *relatedness* and

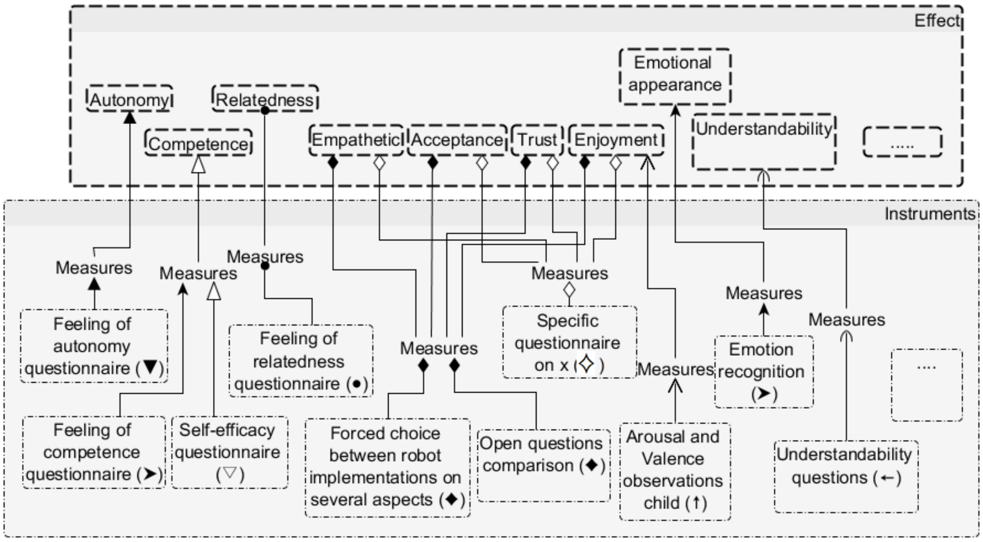


Figure 7: Effects are measured with specific instruments

the method *adapt to others' behavior*. The function to serve this method was *adapt robot to child and activity within boundaries* in the use-case *quiz*. Effects were expected on *emotional contagion, preference, relatedness, empathy, acceptance, trust, fun and motivation*. Relatedness as effect is directly related to the objective of relatedness, the other expected effects are derived from literature on relatedness as being contributing factors to relatedness. The instruments were *arousal and valence observations, forced choice preference, specific questionnaires for relatedness, empathy, acceptance, trust, fun and motivation and open questions related to these aspects*. Figure 8 shows the sDR of this specific experiment, we limited the number of relations in comparison to Figure 1 by excluding the relation between effects and objectives and use cases and functions, both can be derived by following the other relations.

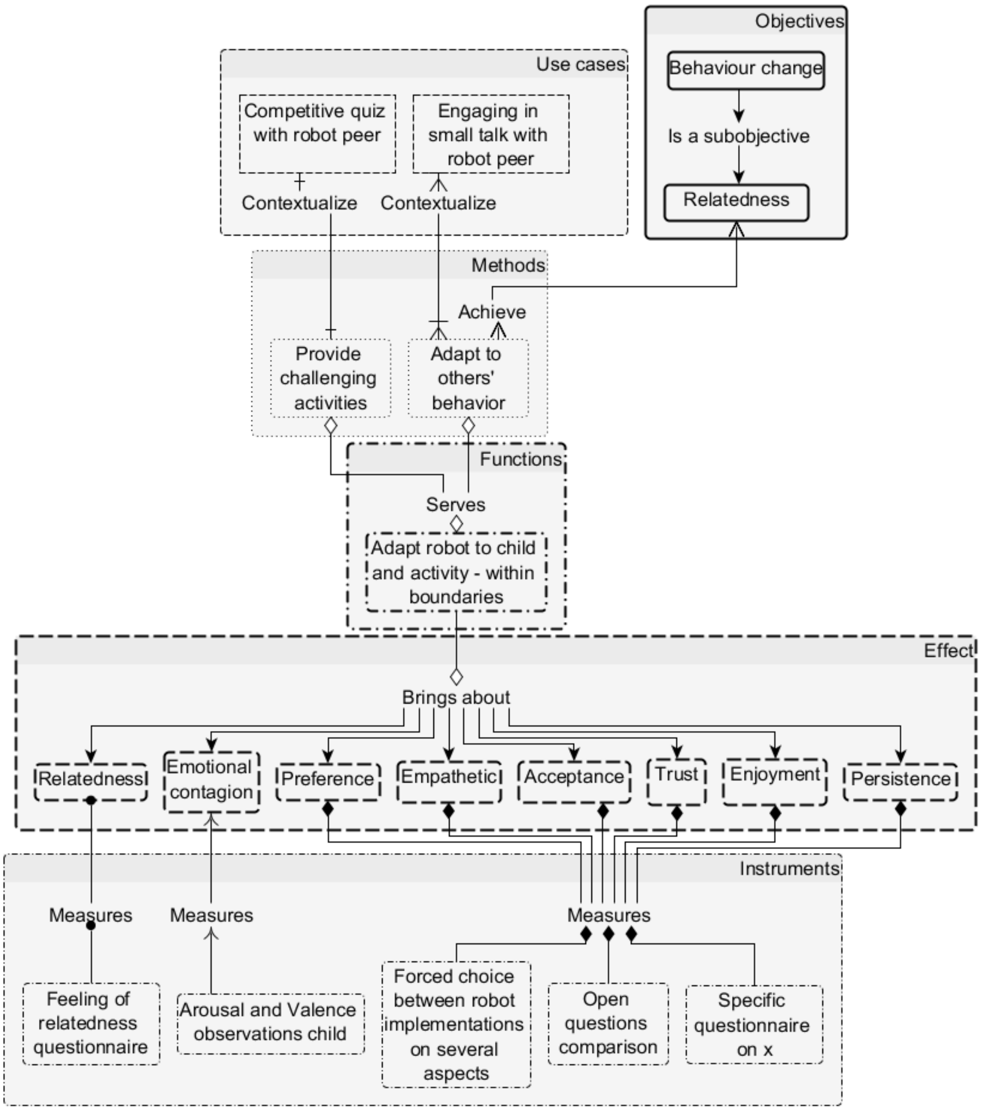


Figure 8: situated Design Rationale of emotional contagion experiment. The use cases "bring about" the effects, but for readability reasons we excluded this line from the overview as we did with the "demonstrate" lines from effects to objectives. [59]

5.1. Results

The objective results on arousal and valence observations showed that the children were significantly more expressive (smiling more) when interacting with the affective robot in comparison with the non-affective robot (M=33.59, SE=17.34) than for the non- affective robot (M=29.06, SE=13.53), (t)(16)=2.156, (p) <0.05, (r)= 0.47 (one-tailed). The answers on the questionnaire on robot-child interaction showed a ceiling effect. Both robots scored very high and the difference was not significant for any of the question topics. In the second questionnaire the children had to choose between one of the two robots on different aspects (e.g. fun, trust) and in the end prefer one of them. There were differences, but non were significantly different, although on trust there seemed to be a trend in favor of the non-affective robot. Finally they were also asked about their motivations to choose one or the other. The most noticeable motivations were clearly that the non-affective robot was more understandable, while the affective robot was preferred most often because it showed emotions.

5.2. Experiment and sDR conclusions

The expression results are quite clear and show a significant effect for the emotional contagion, but this positive effect is not supported by the questionnaires. These suffer from the ceiling effect; only with forced choice some differences can be seen, but still not large. Notwithstanding these ceiling effects we can conclude from the observations that adaptive emotional expressivity influences children to engage in more positive expressivity.
Another interesting result is "trust" where we see that the non-affective robot scores (non-significantly) higher than the affective one. Looking back at the sDR this means that a robot that adapts its state to the child is less trustworthy and might involve lower relatedness. Based on the results we are not ready to conclude this, because it could also be that the sDR is not complete. Reinvestigating literature we see that emotional voices can suffer from understandability issues [57]. This is also supported by the responses the children provide, where they indicate the non-affective robot is more understandable. Understandability is a known factor for trust in automation [60]; in addition, literature on trust of children in caretakers with an unfamiliar accent [61] indicates that understandability influences trust. We have to add understandability thus as a possible downside for prosody which can be measured asking directly about understandability and in concurrence look at effects on trust and acceptance. Figure 9 shows the changed portion of the sDR.
The sDR shows the decisions made for the design of the experiment, this makes it possible to relate the negative result on trust back to the function that was implemented. It shows the sDR is not discriminatory enough on

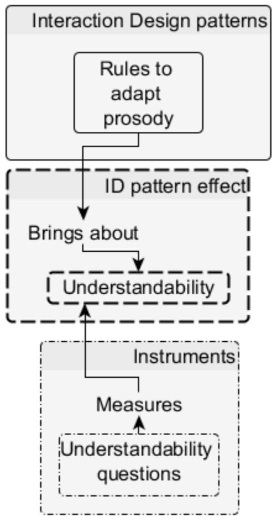


Figure 9: Refinement of sDR, based on emotional contagion experiment.

the effects and that this can be improved by adding a branch to indicate that an interaction design pattern could have influenced the trust. Finally, the experiment provides confirmation that adapting the emotion of the robot to the emotion of the child and activity has a (mainly) positive influence, which can be used for theory building on emotional adaptivity.

6. Conclusion and discussion

The objective of this paper was to provide a formal template that supports the systematic design and evaluation of an experiment and reason about the effects and decisions afterwards. We reached this objective by formalizing the relations between theory, design specifications and evaluations and guidelines for creating it. The developed sDR supports the systematic, iterative and incremental design and evaluation of social robots for behavior change.

To come to this sDR we had to answer three questions. First, we had to specify the relevant concepts. We used the concepts as defined in the sCE method. Second, the relations had to be identified. For this identification we used knowledge on behavior change, social robotics and design specifications.

To make the decisions visible and to support reasoning about the effects and reusability we had a third question on representation of the concepts and their relations. We decided on using concept mapping to visually relate the concepts.

After answering these three questions the sDR method was explained by instantiating the generic sDR template with the European ALIZ-e project. We walked through every concept and its relations to other concepts and also showed how the knowledge from theories and empirical evaluations are taken into account in this process. The complete sDR of the ALIZ-e project can be found here (`https://goo.gl/0HgUC8`). It is interesting to see that, when multiple experiments are concatenated in one overall project sDR, the objectives, methods and their related effects and instruments stay stable over the course of the project. Use cases, functions and interaction patterns on the other hand are added, removed and refined according to the projects progress. This relatively stability of the sDR supports adapting and extending.

At this moment it is not hard to create an sDR for one experiment, as the decisions that are described in the sDR are decisions you take anyway. Which objective do you have with the project, what methods can be used, what functionalities do you want to address in this specific experiment and what effects do you expect and how do you measure these effects? By creating the sDR before performing the experiment shortcomings in the experimental setup can be found.

After the experiment is finished and you would like to do another experiment with the same objectives but other functionalities the sDR can be extended, the easy thing is that the sDR already shows decisions you don't have to think about anymore, the hard thing is to incorporate the new experiment in the old sDR. Sometimes this is easy, e.g. when the functions and expected effects are really different. Other times this is harder, when new interrelations between for instance functions and effects appear. When this happens it means you have to rethink the definitions and try to concatenate or split functions to make the relations less complex or ambiguous. This stipulates the importance of having an ontology in which the concepts are defined, so others also know what is meant by it and can reuse it.

The use of sDR was further exemplified with a specific experiment. In this experiment we could see how sDR supports design and evaluation, the sDR can be adapted after interpretation of the results of the evaluation. With sDR we can reason on why a certain effect occurred (e.g. why did the effect on trust differ from the other effects?). As can be seen Figure 6 there is quite some overlap in effects for different functions in the current sDR of the ALIZ-e project, showing the interactions but also resulting in ambiguous results. This could be improved by identifying claims that are specific for a function or by changing the level of function description, but it will never be perfect needs continuous improvement. By making this possible it also creates the opportunity to identify elements that need to be added to aid the design and evaluation (e.g. experimental support on the design pattern of prosody).

Finally, sDR supports iterative and incremental theory building by showing which elements are validated, which are invalidated and which need more research and/or validation, all within a specific context. Theory building is possible, because the reasoning of the whole chain, from theory to instrument is clarified in the sDR, making it also possible to transfer the ideas to other domains and evaluate it there for more generalizable theories.

Although this is all desirable, it asks for well thought over decisions of the chosen effects and instruments. A further complication that we will not solve is that there can be relations we did not foresee resulting in unexpected effects or incorrect attribution of effects to certain functions.

Notwithstanding these complications sDR provides a method to evaluate a complex system, such as a social robot for behavior change, meanwhile getting an idea of the interaction between functionalities. These interactions are important, because a complex system is never just a combination of its parts. The awareness of interrelations makes it possible to create theories on a level that is fitting to what is "really" known. Furthermore, we will be able to distinguish between groups of outcomes and combine this with user characteristics to develop user profiles which can be used for fast adaptation of the interaction. This will be further explored in the PAL project, a H2020 project on behavior change for self-management of children with diabetes. We foresee reuse of the objectives, most of the methods, effects and instruments with refinement and extension of functionalities more focused on behavior change from the ALIZ-e sDR.
Next to this, by putting relations and concepts in an ontology we further formalize the sDR and make it in this standard format available for people outside the projects. This way, the research community can make use of the knowledge progress on social robots and avatars for children. The complete overview and the experiment specific sDR provide an elaborate guidance in understanding the decisions and the possibility to replicate it. We believe this will open the way to generalizing the results and applying it in other domains.

6.1. Future work

This paper focused on formalizing, reporting and sharing of the design rationale. It's essential to share this rationale with the research and design community and for this we will need an easy to use, preferably interactive, tool. This tool should support the creation of sDRs so they are easier to create, extend and understand. The sDR is now lacking a tool for visualization, the structuring of lines is currently a (mostly) manual and labour intensive job. This is a drawback for creating, adapting and extending an sDR. We would therefore like to develop a tool like sCE has for relating use-cases, expected effects and functions to each other `www.scetool.nl`. This should be extended with a good visualization tool, like they exist for network analyses (e.g. cytoscape.js - `js.cytoscape.org`). With the addition of the related ontology, code and information on the experiment it should then be possible to reproduce the experiment. At the moment the experimental code for the PAL project is stored at a GitHub repository with version numbers for each experiment, and we have the relevant sDR. Sharing this to the research community in a more structured manner should be possible in the future.

Another addition could be to visualize the expected positive and negative effects, this would be similar to sCE where positive and negative claims are made explicit. This will make the sDR both more informative and more complex, so we should think about how to visualize this.

Acknowledgement

This work is (partially) funded by the FP7 ALIZ-e project (grant number 248116), `www.aliz-e.org` and the H2020 PAL project (grant number 643783), `www.pal4u.eu`.

References

[1] T. Fong, I. Nourbakhsh, K. Dautenhahn, A survey of socially interactive robots, Robotics and autonomous systems 42 (2003) 143–166.

[2] K. Dautenhahn, B. Ogden, T. Quick, From embodied to socially embedded agents - implications for interaction-aware robots, Cognitive Systems Research 3 (2002) 397–428.

[3] A. Weiss, R. Bernhaupt, M. Lankes, M. Tscheligi, The USUS evaluation framework for human-robot interaction, in: AISB2009: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction.

[4] C. Susan Michie DPhil, C. Marie Johnston PhD, C. Charles Abraham DPhil, C. Jill Francis PhD, M. P. Eccles, The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions, Annals of Behavioral Medicine (2013) 1–15.

[5] H. Oinas-Kukkonen, M. Harjumaa, Persuasive systems design: Key issues, process model, and system features, Communications of the Association for Information Systems 24 (2009) 28.

[6] H. Oinas-Kukkonen, Behavior change support systems: A research model and agenda, in: Persuasive Technology, Springer, 2010, pp. 4–14.

[7] M. A. Neerincx, Situated cognitive engineering for crew support in space, Personal and Ubiquitous Computing 15 (2011) 445–456.

[8] G.-J. M. Kruijff, M. Janíček, S. Keshavdas, B. Larochelle, H. Zender, N. J. Smets, T. Mioch, M. A. Neerincx, J. Diggelen, F. Colas, et al., Experience in system design for human-robot teaming in urban search and rescue, in: Field and Service Robotics, Springer, pp. 111–125.

[9] IDF, 5th IDF diabetes atlas, Technical Report, International Diabetes Federation (IDF), 2012.

[10] C. C. Patterson, G. G. Dahlquist, E. Gyürüs, A. Green, G. Soltész, E. S. Group, et al., Incidence trends for childhood type 1 diabetes in europe during 1989–2003 and predicted new cases 2005–20: a multicentre prospective registration study, The Lancet 373 (2009) 2027–2033.

[11] T. Danne, O. Kordonouri, What is so different about diabetes in children, Diabetes Voice 52 (2007) 16–19.

[12] K. V. Narayan, J. P. Boyle, T. J. Thompson, S. W. Sorensen, D. F. Williamson, Lifetime risk for diabetes mellitus in the united states, Jama 290 (2003) 1884–1890.

[13] M. Peyrot, How is diabetes perceived? the results of the dawn youth survey, Diabetes voice 53 (2008) 9–13.

[14] K. K. Boman, J. Viksten, P. Kogner, U. Samuelsson, Serious illness in childhood: the different threats of cancer and diabetes from a parent perspective, The Journal of pediatrics 145 (2004) 373–379.

[15] E. Tsiouli, E. C. Alexopoulos, C. Stefanaki, C. Darviri, G. P. Chrousos, Effects of diabetes-related family stress on glycemic control in young patients with type 1 diabetes systematic review, Canadian Family Physician 59 (2013) 143–149.

[16] B. Robins, K. Dautenhahn, P. Dickerson, From isolation to communication: a case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot, in: Advances in Computer-Human Interactions, 2009. ACHI'09. Second International Conferences on, IEEE, pp. 205–211.

[17] E. I. Barakova, T. Lourens, Expressing and interpreting emotional movements in social games with robots, Personal and ubiquitous computing 14 (2010) 457–467.

[18] F. Tanaka, T. Kimura, Care-receiving robot as a tool of teachers in child education, Interaction Studies 11 (2010) 263–268.

[19] S. Tejada, N. Traft, M. Hutson, H. Bufford, M. Dooner, J. Hanson, A. Radler, G. Mauer, Educational robots: Three models for the research of learning theories and Human-Robot interaction, pp. 70–76.

[20] J. Kennedy, P. Baxter, T. Belpaeme, The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning, in: Proc. HRI, volume 15.

[21] E. Short, K. Swift-Spong, J. Greczek, A. Ramachandran, A. Litoiu, E. C. Grigore, D. Feil-Seifer, S. Shuster, J. J. Lee, S. Huang, et al., How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play, in: Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on, IEEE, pp. 924–929.

[22] H. Yanco, J. Drury, Classifying human-robot interaction: an updated taxonomy, International Conference on Systems, Man and Cybernetics 3 (2005) 2841–2846.

[23] D. R. Olsen, M. A. Goodrich, Metrics for evaluating human-robot interactions, in: Proceedings of PERMIS, volume 2003, p. 4.

[24] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, M. Goodrich, Common metrics for human-robot interaction, in: Proceedings of the HRI 2006, ACM, pp. 33–40.

[25] S. Woods, K. Dautenhahn, J. Schulz, The design space of robots: Investigating children's views, in: ROMAN 2004, IEEE, pp. 47–52.

[26] C. Bartneck, J. Forlizzi, A design-centred framework for social human-robot interaction, in: ROMAN 2004, IEEE, pp. 591–594.

[27] K. Dautenhahn, Design spaces and niche spaces of believable social robots, in: Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on, IEEE, pp. 192–197.

[28] M. E. Spiekman, P. Haazebroek, M. A. Neerincx, Requirements and platforms for social agents that alarm and support elderly living alone, in: Social Robotics, Springer, 2011, pp. 226–235.

[29] V. Venkatesh, M. G. Morris, G. B. Davis, F. D. Davis, User acceptance of information technology: Toward a unified view, MIS quarterly (2003) 425–478.

[30] C. Bartneck, D. Kulić, E. Croft, S. Zoghbi, Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots, International Journal of Social Robotics 1 (2009) 71–81.

[31] P. Craig, P. Dieppe, S. Macintyre, S. Michie, I. Nazareth, M. Petticrew, Developing and evaluating complex interventions: the new medical research council guidance, Bmj 337 (2008) a1655.

[32] M. Neerincx, J. Lindenberg, Situated cognitive engineering for complex task environments, Aldershot, UK: Ashgate Publishing Limited, p. coming soon. Editor: Schraagen, J.M., Militello, L., Ormerod, T., & Lipshitz, R.

[33] O. A. B. Henkemans, P. van Empelen, G. L. Paradies, R. Looije, M. A. Neerincx, Lost in persuasion a multidisciplinary approach for developing usable, effective, and reproducible persuasive technology for health promotion, in: Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2015 9th International Conference on, IEEE, pp. 49–56.

[34] E. Hollnagel, D. Woods, Cognitive systems engineering: New wine in new bottles, International Journal of Man-Machine Studies 18 (1983) 583–600.

[35] D. Norman, User-Centred System Design: New perspectives on human-computer interaction, Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 31–61.

[36] J. Rasmussen, Information processing and human-machine interaction: An approach to cognitive engineering, Elsevier Science Inc. New York, NY, USA, 1986.

[37] J. D. Novak, A. J. Caas, The theory underlying concept maps and how to construct and use them, Florida Institute for Human and Machine Cognition Pensacola Fl, www. ihmc. us.[http://cmap. ihmc. us/Publications/ResearchPapers/T heoryCmaps/TheoryUnderlyingConceptMaps. htm] 284 (2008).

[38] G. Cockton, Diffusion of worth mapping: The worth of resource functions (2013).

[39] M. Fishbein, I. Ajzen, Prediction and change of behavior: The reasoned action approach, 2010.

[40] K. Witte, Fear control and danger control: A test of the extended parallel process model (eppm), Communications Monographs 61 (1994) 113–134.

[41] E. L. Deci, R. M. Ryan, Handbook of self-determination research, University Rochester Press, 2002.

[42] R. M. Ryan, E. L. Deci, Intrinsic and extrinsic motivations: Classic definitions and new directions, Contemporary educational psychology 25 (2000) 54–67.

[43] E. L. Deci, R. M. Ryan, The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior, Psychological inquiry 11 (2000) 227–268.

[44] W. Miller, S. Rollnick, Motivational interviewing: Preparing people for change, The Guilford Press, 2002.

[45] C. P. Niemiec, R. M. Ryan, Autonomy, competence, and relatedness in the classroom applying self-determination theory to educational practice, Theory and Research in Education 7 (2009) 133–144.

[46] S. J. Sebire, R. Jago, K. R. Fox, M. J. Edwards, J. L. Thompson, Testing a self-determination theory model of children's physical activity motivation: a cross-sectional study, Int J Behav Nutr Phys Act 10 (2013) 111.

[47] R. M. Ryan, C. S. Rigby, A. Przybylski, The motivational pull of video games: A self-determination theory approach, Motivation and emotion 30 (2006) 344–360.

[48] D. Markland, R. M. Ryan, V. J. Tobin, S. Rollnick, Motivational interviewing and self–determination theory, Journal of social and clinical psychology 24 (2005) 811–831.

[49] J. H. Bigelow, S. Cretin, M. Solomon, S.-Y. Wu, M. O'Connell, Patient compliance with and attitudes towards Health Buddy, Technical Report, DTIC Document, 2000.

[50] R. Looije, M. A. Neerincx, F. Cnossen, Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors, International Journal of Human-Computer Studies 68 (2010) 386 – 397.

[51] L. S. Vygotsky, Mind in society: The development of higher psychological processes, Harvard university press, 1980.

[52] M. Csikszentmihalyi, S. Abuhamdeh, J. Nakamura, Handbook of Competence and Motivation, New York: The Guilford Press, pp. 598–698.

[53] W. S. Grolnick, S. T. Gurland, K. F. Jacob, W. Decourcey, The development of self-determination in middle childhood and adolescence, Development of achievement motivation (2002) 147–171.

[54] R. Zhao, A. Papangelis, J. Cassell, Towards a dyadic computational model of rapport management for human-virtual agent interaction, in: Intelligent Virtual Agents, Springer, pp. 514–527.

[55] A. Assor, H. Kaplan, G. Roth, Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork, British Journal of Educational Psychology 72 (2002) 261–278.

[56] R. Looije, M. A. Neerincx, J. Peters, Integrating robot support functions into varied activities at returning hospital visits, International Journal of Social Robotics (in press).

[57] J. M. Kessens, M. A. Neerincx, R. Looije, M. Kroes, G. Bloothooft, Facial and vocal emotion expression of a personal computer assistant to engage, educate and motivate children, in: 3rd IEEE Int. Conference on Affective Computing and. Intelligent Interaction (ACII), Amsterdam, the Netherlands.

[58] R. Looije, A. Van der Zalm, M. Neerincx, R.-J. Beun, et al., Help, i need some body the effect of embodiment on playful learning, in: RO-MAN, 2012 IEEE, IEEE, pp. 718–724.

[59] M. Tielman, M. Neerincx, J.-J. Meyer, R. Looije, Adaptive emotional expression in robot-child interaction, in: Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction, ACM, pp. 407–414.

[60] R. R. Hoffman, M. Johnson, J. M. Bradshaw, A. Underbrink, Trust in automation, Intelligent Systems, IEEE 28 (2013) 84–88.

[61] K. D. Kinzler, K. H. Corriveau, P. L. Harris, Childrens selective trust in native-accented speakers, Developmental science 14 (2011) 106–111.

Int J of Soc Robotics (2016) 8:483–497
DOI 10.1007/s12369-016-0365-8

Integrating Robot Support Functions into Varied Activities at Returning Hospital Visits

Supporting Child's Self-Management of Diabetes

Rosemarijn Looije[1] · Mark A. Neerincx[2] · Johanna K. Peters[3] · Olivier A. Blanson Henkemans[1]

Accepted: 13 June 2016 / Published online: 23 June 2016

Abstract Persistent progress in the self-management of their disease is important and challenging for children with diabetes. The European ALIZ-e project developed and tested a set of core functions for a social robot that may help to establish such progress. These functions were studied in different set-ups and with different groups of children (e.g. classmates at a school, or participants of a diabetes camp). This paper takes the lessons learned from these studies to design a general scenario for educational and enjoying child–robot activities during returning hospital visits. The resulting scenario entailed three sessions, each lasting almost one hour, with three educational child–robot activities (quiz, sorting game and video watching), two intervening child–robot interactions (small talk and walking), and specific tests to assess the children and their experiences. Seventeen children (age 6–10) participated in the evaluation of this scenario, which provided new insights of the combined social robot support in the real environment. Overall, the children, but also their parents and formal caregivers, showed positive experiences. Children enjoyed the variety of activities, built a relationship with the robot and had a small knowledge gain. Parents and hospital staff pointed out that the robot had positive effects on child's mood and openness, which may be helpful for self-management. Based on the evaluation results, we derived five user profiles for further personalization of the robot, and general requirements for mediating the support of parents and caregivers.

Keywords Diabetes · Children · Social robots

✉ Rosemarijn Looije
rosemarijn.looije@tno.nl

Mark A. Neerincx
mark.neerincx@tno.nl

Johanna K. Peters
johannapeters.jkp@gmail.com

Olivier A. Blanson Henkemans
olivier.blansonhenkemans@tno.nl

1 TNO, Postbus 23, 3769 ZG Soesterberg, The Netherlands

2 Interactive Intelligence Group, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

3 Department of Artificial Intelligence, University of Groningen, Postbus 407, 9700 AK Groningen, The Netherlands

1 Introduction

1.1 Diabetes Type 1

The growing burden of chronic illness on health and health care has globally led to health policy responses increasingly referring to self-management. This applies to the increasing number of children and adolescents in Europe with a chronic illness. For example, the incidence of childhood type 1 diabetes mellitus (T1DM) in Europe, now ranging from 3.9/100,000 cases per year in Macedonia to 57.4/100,000 in Finland [26], is rising rapidly. In the below 5-year-old age group, there is a doubling time of less than 20 years [13]. T1DM is associated with serious physical and psychological complications [8,27], which may appear sooner or later, cause high morbidity and mortality, affect the quality of life, and increase health-care costs [14]. Complications can be prevented by performing self-management (e.g., monitoring blood glucose, recognizing symptoms and injecting insulin). However, self-management is not an easy goal to attain for young patients. First, it requires motivation and long-term perseverance, in order to become a way of life. However, children's illness regularly causes feelings of embarrassment (approximately 25 % of the youth involved in a study of Peyrot [27]), and negative effects on school performance and psychological well-being. Improving the way they feel about diabetes, might be a first step in improving the self-management. Second, the children need not only to learn to self-manage their lifestyle-related diseases to improve their situated health-related habits, but also to be prepared for the physical and social changes at adolescence. Third, the specific self-management goals of children and adolescents are strongly affected by a diversity of personal and environmental factors, such as the childs developmental stage, parents support and health care providers. So, children and their social environment have to find a personalized strategy to establish pervasive self-management.

1.2 Improving Self-Management

There is a broad source of literature on theories that are relevant for self-management support: Changing behavior [9,20,30], persuasive design [11], gaming theory [7], education [40] and behavior change support systems [24]. These theories have some common principles. According to the first principle, intrinsic motivation is key and requires that someone feels in control of the situation (experience autonomy). This can be reached for instance by providing variation and influence of dialog. The second principle emphasizes the feeling of competence: The user should feel capable of reaching an objective. This principle originates from educational and gaming theory [7,40], and from behavior change literature [9,20,30], stating that relevant activities and objectives should be provided, which are challenging and achievable, and for which positive feedback should be provided. The third and final principle concerns relatedness: Education and self-management are improved when there is a relation between tutor and trainee. The tutor can be a peer or teacher with whom a form of relatedness (or rapport) is build up [20,25,40]. The three factors: autonomy, competence and relatedness are the building blocks of the self-determination theory (SDT) [9].

1.3 Social Robots

Social robots show human-like (social) characteristics, e.g. they express emotions and use natural cues as gaze to share point of focus [12]. For prolonged self-management support, rapport should be build up between child and robot resulting in a positive effect on relatedness [4]. In Zhao et al. [25], several behaviors are identified to create rapport between an agent and a person. Examples are the initiation of mutual self-disclosure, praise and acknowledgement, and referring to shared experiences. It is interesting to note that these behaviors are also prescribed in behavior change methods, e.g., express empathy in Motivational Interviewing [21]. So, the social robot can be viewed as an embodiment of a behavior change support system [24]. Such robots are being used for behavior change support, for instance, to support persons with autism [1,29], to acquire a healthy lifestyle [33], and to educate persons (e.g. [18,35,36]). A robot has a rich set of possibilities to incorporate behavior-change methods from social sciences, but the specific translation from these methods to a coherent and concise set of robot functions is complex and difficult to evaluate.

1.4 Situated Cognitive Engineering

The European ALIZ-e project aimed at a social robot that 7 to 11 year old children could use recurrently and possibly help these children to progress on self-management (i.e., autonomy, competence and relatedness, [2,3]; see Sect. 1.2). An iterative situated Cognitive Engineering method was applied [22], to (i) derive use cases, requirements and claims for the self-management support (i.e. the design rationale), and (ii) build prototypes to test and refine the design rationale. The tests were conducted at schools and hospitals, focusing on specific parts of the design rationale, i.e. one or more “core functions” of the social robot that were hypothesized to have effect on relevant SDT-factors. For example, the idea that relatedness is stimulated by having a background story for the robot [39]. These functions were studied in different setups and with different groups of children (e.g. classmates at a school, or patients in a hospital). Often it was not (yet) required (for a first test and refinement cycle) to involve the target group, children with diabetes. This paper takes the lessons learned from these tests to design a general scenario, incorporating a variety of use cases. This way, an integrated set of core functions was prototyped and tested with children with diabetes in a hospital (i.e. the real target environment).

The next sections provide an overview of the earlier experiments conducted and their results. The current study incorporates the “proven” functions and makes use of the insights on the experimental setup that we built up in these experiments. The resulting social robot and scenario are evaluated with diabetic children in a hospital setting, studying the influence on autonomy, competence and relatedness. Furthermore, the perceptions and opinions of the children, their parents and their medical caregivers on the short and long-term are investigated. Conclusive evidence on the effects of the specific metrics could not be found, but the interactions with the children, parents and caregivers during the evaluation and afterwards gave valuable insights. Parents and caregivers became more enthusiastic over time and reported results in increased self-management and lower thresholds in hospital visits.

2 Lessons Learned from Previous Experiments

Over four years, several tests were conducted, in which children interacted with a social robot (Philips iCat or Aldebaran

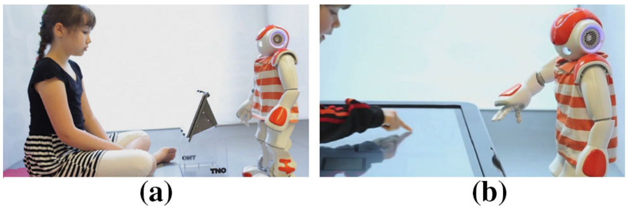


Fig. 1 The games: quiz and sorting game (**a**) The quiz (**b**) The sorting game

NAO) and performed one or several activities with the robot. These activities were designed to examine the effects of specific support functions, e.g. on specific learning objectives. Four educational activities were developed. The first was a Trivial Pursuit®based quiz in which robot and child played against each other. This educational quiz had a textual and competitive nature Fig. 1a. The second activity was an educational sorting game Fig. 1b, in which the child and robot classified objects in categories and could cooperate to reach the highest classification score. Due to its collaborative nature and visual orientation, the sorting game involved another learning style than the competitive and textual quiz (whereas, they could support the same learning objective). The third educational activity entailed different versions of movement games [31], which could address the same learning objective, but in a kinesthetic learning style. The fourth activity used educational videos that are both visual and aural. With this variety of activities, the social robot could support a variety of learning styles [10]. Next to these educational activities, there were "intervening" activities, such as small talk, to establish continuous child–robot interactions. All robot support functions were designed to address the objectives of SDT: autonomy, competence and relatedness. Table 1 provides an overview of the relevant experiments, their relations to the objectives of SDT, the context (setting and users) in which the experiment was conducted, the results and the transfer of these results into the integrated social robot (that will be tested subsequently).

According to SDT, a feeling of autonomy can be enhanced by providing choices. To stimulate this, the ALIZ-e project aimed at providing numerous activities that robot and child could do together. The quiz and sorting game were developed to support this. They both focus on education, but where the robot and child are playing against each other in the Trivial Pursuit®based quiz, in the sorting game they have to cooperate to get the highest score. In [15] it was shown that the possibility to switch between activities is beneficial for the motivation (see experiment 1 in Table 1).

The second factor of SDT, competence, can be supported by adapting the difficulty of the exercises to the child [16]. This adaptation proved to be beneficial for the motivation of the children (see experiment 2). It should be noted that the robot was not an expert in this interaction, i.e., the robot made the same amount of errors as the child [32]. Showing that the robot was not an expert was emphasized by making the robot exhibit thinking behavior [42]. Overall, this resulted in a positive experience of the robot (see experiment 3 and 4 in Table 1). In addition to competence, experiment 3 and 4 also addressed relatedness by encouraging self-confidence.

The third pillar of SDT is relatedness, meaning that the robot is experienced as a "pal". Firstly we made sure that the robot can exhibit recognizable emotions [6,19] (see experiments 5 and 6). We also looked at adapting the robot to the personality of the child [38], but we found that personality is probably not a good aspect to adapt to (experiment 7). We still expect that adapting to energy level, and perhaps modulating the energy level of the child will support the relatedness, but this was not evaluated. We did evaluate the adaption of robot's emotional state to the state of the user and state of the situation (within boundaries) [37]. The results from this experiment showed that children who interacted with the robot that adapted its emotional state to the child and situation, showed more, and more positive, emotional expressions than children who interacted with a robot that did not adapt its emotions to the child and situation (experiment 8). However, recognizing child's emotions in an interactive situation is still very hard. Therefore, we studied the effects of remembering small facts about their life (e.g. name, hobbies, information provided in a previous session) [5]. This is rather easy to implement and proved to have a very positive effect on the children (see experiment 9). Another easy to implement functionality is that the robot tells something about itself (e.g. age, hobbies), which proved to increase the willingness of the children to disclose information about themselves [39] (experiment 10). Finally, we looked at the willingness of children to touch the robot [34]; experiment 11 showed that they are quite willing.

In addition to the conclusive results, interesting observations were acquired during the experiments that are relevant for the further development of the robot. For example, changing activities by the robot and the child themselves proved to be stimulating (e.g., to transfer from quiz to sorting game without the help of the experiment leader [15]; see results experiment 1 in Table 1). Another observation was that providing a confined, shared environment for the robot and child proved to reduce child's feeling of being observed and part of an experiment [34] (experiment 11).

3 Constructing an Integrated Set of Child–Robot Activities for Hospital Visits

Table 1 provides an overview of the 11 experiments that examined the specific robot support functions for child's self-management with their relations to the self-determination theory (SDT), the location of the experiment, the participants

Table 1 The 11 experiments that examined specific robot support functions for child's self-management with their relations to the self-determination theory (SDT), the location of the experiment, the participants and the results

Nr.	Experiment	SDT objectives focus	Location	Users (nr, age)	Results (implemented in current evaluation y/n)
1	Multiple activities (quiz, sorting game) in hospital setting [15]	Autonomy	Hospital	Non-diabetic children (13, age 7–11)—hospitalized	Multiple activities are beneficial (y) child and robot should go from one activity to another (y)
2	Difficulty of math assignments adapted to performance child [16]	Competence	School	Children (20, age 9–10)	Adapting the difficulty of exercises to the child increases motivation (n)
3	Make the robot fallible in his answers [32]	Competence	School	Children (22, age 10–12)	Robot is able to adapt its performance to performance child (n) fallibility is not proven effective, but theory is convincing (y)
4	Make the robot think about its answers. The robot takes some time and expresses thinking behavior [42]	Competence	School	Children (26, age 9–11)	The thinking robot is experienced as faster, more humanlike and more likeable, without decreasing perception of intelligence, trustworthiness and autonomy (y)
5	Make robot express recognizable emotions [19]	Relatedness	Research institute	Children (18, age 8–9)	Emotions are recognizable and add to likeability (y)
6	Compare capability to express emotions between robots [6]	Relatedness	School	Children (14, age 8–9)	Emotions of NAO are equally well recognized as thos of iCat (y)
7	Adapt robot personality behavior to personality child [38]	Relatedness	School	Children (16, age 7–9)	Personality was very hard to determine (no correlation between what child, parent and teacher indicated) so not a good factor to adapt to. (n)
8	Adapt robot emotions to child emotions and state of activity (within boundaries) [37]	Relatedness	School	Children (18, age 8–10)	Children that interacted with the robot that adapted its emotional state to the child and situation showed more, and more positive, emotional expressions than children that interacted with a robot that did not adapt its emotions to the child and situation (y)
9	Make robot remember small facts about their life (e.g. name, hobbies, information provided in a previous session) [5]	Relatedness	Hospital and At home	Diabetic children (30, age 6–12) - not hospitalized	Very positive effect on relation children towards robot (y)
10	Make robot disclose information about itself (e.g. age, hobbies) to stimulate disclosure of child [39]	Relatedness	At home	Diabetic children (6, age 9–12)—not hospitalized	Increased willingness of children to disclose information about themselves (y)
11	Ask the child to touch the robot in an interactive move session [34]	Relatedness	School	Children (22, age 9–11)	Most children like to touch the robot (y) The children like an enclosed environment where they are more alone? with the robot (y)

and the results. The following subsections will elaborate on these results and describe how they will feed into the next version of the robot and the set of child–robot activities for returning hospital visits.

3.1 Child–Robot Interaction Environment

Based on the knowledge gathered in experiment 11 [34] we developed a physical setup for this evaluation. Firstly, we used the robot playground as used in [34] again. The playground (see Fig. 3) consists of three walls of 150cm high on which a robot landscape is depicted in soft grays. The floor consists of grey playtiles and one red and one blue depicting the positions the child should sit for the different games. All cables are hidden under the floor and behind the walls and two cameras are unobtrusively placed behind the walls so they just peek over it. The playground provides a shared environment for robot and child and since we did this experiment inside the hospital it also makes the surroundings more friendly. Furthermore, because children sit on the ground with the robot they are naturally on the same level as the robot, which is different when the robot stands on a table and the child sits at the table, which is also more static. Finally, the shared environment closes of the rest of the environment more, so the experimenter, who is in the same room, is easier to forget about.

3.2 Child–Robot Activities

Next to the environment we made sure the interactions were in concordance with what we learned from previous experiences. The evaluation was a wizard-of-oz evaluation, which meant that the experimenter/wizard did the speech and state recognition of the child and there was a protocol that was followed that described the possible dialog and behavior actions. The wizard had some freedom to put in new text for the robot to say. The wizard had camera images from the playground, could switch from camera dependent on the activity, and had an elaborate wizard interface to direct the interaction. Overall, the activities consisted of three educational child–robot activities (quiz, sorting game and video watching), two intervening child–robot interactions (small talk and walking), and specific tests to assess the children and their experiences.

3.2.1 Educational Activities

The child and robot could do three activities together, following experiment 1 that concluded that multiple activities are beneficial [15]. The two games as developed within the ALIZ-e project and an educational video. The quiz was based on Trivial Pursuit®. Child and robot each stand on opposite sides of a tablet in a kind of see-saw construction (see Fig. 1a). The tablet is turned towards the robot and it can then ask the first multiple choice (A–D) question. After posing the question the robot turns the tablet towards the child and the child can answer, by saying the answer out loud (no touch). The robot reacts on the answer and congratulates when it is correct and provides the argumentation when it is incorrect. There is no judgment when the answer is incorrect. Then the next question appears on the tablet and the child can pose it to the robot. The robot thinks about the answers it provides (experiment 4) [42] and makes errors (experiment 3) [32]. The game can be set up competitive, but we did not incorporate a scoring mechanism.

The sorting game shows pictures on a large touch screen (see Fig. 1b), the pictures need to be swiped into one of two categories that are named/depicted on the sides of the screen. The categories are for instance "high in" and "low in" carbohydrates and pictures shown on the screen are "a salad", "chips", "bread", "sweets", "milk" etc.. Child and robot stand on opposite sides of the table and they can both, one at the time, swipe a picture in the correct category. The aim here is to get a high score together, so it is a collaborative game setup. During the game the robot acknowledges the actions of the child with exclamations as "too bad", "you did great".

The difficulty of both the quiz and the sorting game was not adapted to the users' performance although it was found to be effective (experiment 2) [16]. We did not do this because of a limited number of questions/assignments per session and a high variability between children. The questions/assignments were related to diabetes and thus relevant for the children.

The final activity is not a game, but an educational video the robot and child can watch together. The video is for instance about the symptoms of high blood glucose levels (a "hyper").

After a certain number of questions of the quiz (8), or a certain amount of time with the sorting game (5 min) the robot initiated a change activity dialog. The child could then choose to proceed or change activity, although in the first and second session they had to do all activities so there was a time limit on how long they could do each game (10 min max). The child could also initiate the dialog to change the activity. When this was really soon after starting the activity, the robot tried to convince the child to do it a little longer ("just a few more questions"), otherwise it would agree on changing.

3.2.2 Small Talk

Based on experiment 9 and 10 [5,39] we incorporated small talk in the evaluation. At the start of the evaluation the robot asked the child some personal information: Name, age, hobby. The robot did also ask if the child had questions for the robot, so it could also answer questions about its age

Fig. 2 Walking with the robot

and hobbies. Furthermore, the robot asked at the end of the first and second session if the child had plans for the coming weeks (until the next session) and referred back to these in the next session. Finally, during the activities the robot asked questions about diabetes. The robot for instance said "The holiday period seems to be really hard to me, with all the candy and strange food, how do you deal with that?". During the small talk and the activities the robot displayed emotions that correlated with the situation (experiment 5, 6 and 8 [6,19,37]).

3.2.3 Walking

Because we did not want a detrimental effect on the interaction when switching activities, because of interference of the experimenter (experiment 1 [15]), we decided the child was responsible for getting the robot from one activity to another (see Fig. 2). We thought this would work because experiment 11 [34] showed no hesitation of most children to touch the robot. We explained how to walk with the robot, but when some children started to lift the robot we also accepted this. Something else that came up after a few of the first sessions were finished, was that the robot fell over sometimes and most children felt the need to help it up. Therefore we added a function that made sure the robot would not hurt the child, shutting down the automatic stand up function and removing motor stiffness, so that the child could support the robot standing up. We also explained to the children how they could help the robot in standing up by putting it in sitting position.

4 Evaluation

In order to get a feeling of how diabetic children interacted with the NAO when different activities are offered and physical interaction is possible we carried out an experiment.

Fig. 3 The robot playground

4.1 Evaluation Method

4.1.1 Participants

17 diabetic children in the age of 6–10 (M = 8.24, SD = 1.25) participated in the experiment. They were selected by their diabetic nurses of the Meander Medical Centre (Amersfoort, the Netherlands) and on basis of the parents willing to come three times extra to the hospital. All children got the diabetes diagnosis more than a year and a half ago, the range was 23–108 months (M = 51, SD = 29,64). Most children used a pump to regulate their insulin intake (11), the others used insulin injections (6).

4.1.2 Materials

To execute the experiment in an adequate way the following materials are needed for the experimental setting: The child with the robot on the robot playground and the execution of the experiment including measurement material.

- Robot playground: playtile floor of $2 \times 3\,\text{m}^2$ with walls (Fig. 3)
- See-saw tablet holder, a device enabling turntaking by flipping the tablet
- Samsung Tablet
- 15″ screen to watch little movies about diabetes
- 27″ television touch screen with table legs, to play the sorting game
- Questionnaires
- Wizard Laptop
- Movie Laptop
- An extra screen to watch interaction
- Cameras to record interaction
- 3 NAO robots (2 minimum needed for third session and backup when technical failures occur)

4.1.3 Measures

We used observations, tests and questionnaires to quantify and qualify the interaction with the robot.**Tests**

Knowledge test This questionnaire is used to assess whether there is knowledge improvement. This test is filled out before the first and after the last interaction and consists of 32 knowledge questions (e.g. What is important for you to know about your physical education class? (a) If you're going to do something fun, (b) If it is active or calm what you're going to do, (c) If you are going to play football, (d) If you're clothes look good: b is correct). The questions one until eight occur in the first session of interaction on the tablet, questions nine until 16 in the second session and questions 17 until 24 in the third session (for the children who chose the quiz). When questions or answers were not understood or the children were not able to read they received help.
Self-efficacy test The SE card-sorting questionnaire is used to assess the current autonomy of the child. To measure SE, a card sorting questionnaire based on Karoly and Bay [17] is used together with diabetes-care activities proposed by the diabetes specialists of the Meander Medical Center.
Memory test With the aid of a memorizing task we examine whether children memorize more information given by a familiar robot, as is expected when intrinsic motivation is higher due to a peer teacher that applies SDT strategies [23]. In the third session every child listens to two robot stories. One story is based on the English Wechsler Intelligence Scale for adults (Williams [41]) and the other one thought up using the same build up. One story is given from the familiar robot (called Charlie) and the other story is provided by another NAO robot (called Robin), who is introduced as a friend of Charlie. This robot is exactly the same as Charlie, but has a different voice and wears a grey striped shirt. The order of the stories and the robots is counterbalanced. After each story there is a short recall memory test. First the children are asked to reiterate the story as best as they can (immediate free recall). After this they are asked nine questions about the story (Immediate cued recall). An example of such a question is: "what was the name of the lady in the story?"

Questionnaires

Fun Questionnaire To measure the pleasure and fun the children experienced the children filled in a Likert scale questionnaire about the robot and the activities. First there were three 7-point questions on fun with the robot, quiz and sorting game, after which four 4-point questions were asked related to different aspects of the robot. The questionnaires used were based on the Smileyometer from the Fun Toolkit of Read and MacFarlane [28].
SDT Questionnaire To measure the feeling of self-determination we asked the children 10 questions on a 7-point Likert scale. Question 2,3,8 and 9 were regarding feeling of competence, question 4,6,7,10 were about feeling of relatedness and question 1 and 5 were related to feeling of autonomy.

Observations

Game preference In the second session the children could say which game they preferred and were asked if they wanted to start with this game and in the third session they could only choose one game.
Online analysis and offline video and logging analysis For the analysis of the whole interaction in each session we used notes that were taken during the interaction, video analysis and analysis of the logs. We looked at walking, time with activities, game order, attention of child, interaction with robot (talking general, talking diabetic related, touching), reaction on technical failures, empathy with robot, and how much the experiment leader is involved.

4.1.4 Procedure

Every child had three sessions of about an hour in the hospital. These appointments were at least 14 days apart (see Fig. 4).

In the first session the NAO robot, called Charlie, is introduced as a robot that helps children to manage their diabetes but still has to learn many things about diabetes himself. The experiment leader explains the activities in short and shows how the children can walk with Charlie. The interaction with Charlie starts with small talk and walking followed by one of the games. With the quiz Charlie has to be put exactly in front of the bars on the ground to be able to turn the tablet. In each session at least eight questions are played so that after three

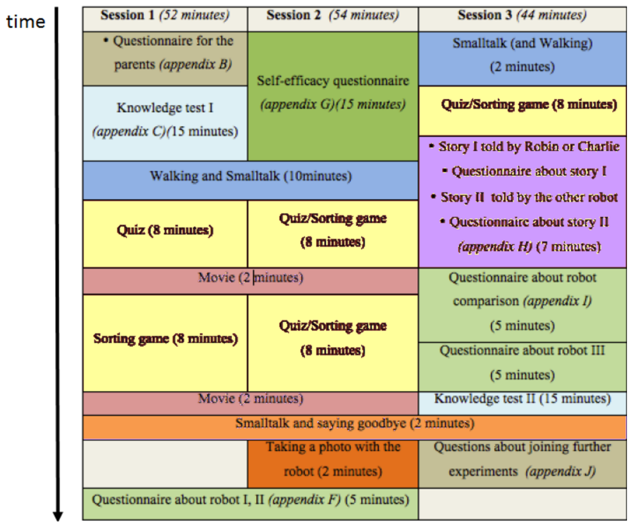


Fig. 4 Planning for the three sessions

sessions they practiced 24 of the 32 knowledge test questions (if they chose to do the quiz in the last session). The sorting game is on the other side of the playground on a large touchscreen. Several pictures are shown and the child and NAO have to put them in the correct category (on one of the sides of the display). Examples of categories are: hyper/hypo, low/high carbohydrates. During each game open questions related to diabetes are asked to support self-disclosure (e.g. "Did it ever happen to you that you had a hypo or hyper and did not notice? How come did this happen?"). In between the games in the first and second session the children can watch an 1-min movie about dealing with typical diabetes situations, which is presented on a 15" screen. Dependent on the time left another short movie can be presented to the child after the games. After the interaction with the robot the children always fill in a questionnaire concerning judgment of the robot and the games they have played. The first session starts with the quiz. In the second session the children are allowed to choose with which game they want to begin with but they have to play them both. In the third session only one game is played, chosen by the child, because of a new scenario where the children meet Charlie's friend Robin. Both Robin and Charlie tell a short story after which the children have to do a test with free and cued recall about the story.

5 Results

Below we will describe the results from the evaluation. These results are divided in results that can be directly derived from the instruments and observations used in the evaluation and in feedback we got afterwards.

The tests were analyzed using t-tests and the questionnaires using the non-parametric Wilcoxon and Friedman tests. Game preference was counted and compared between the second and third session. The video and logging analysis was performed using Grounded Theory as starting point. This was because the 17 children differed in age, phase in their illness and interaction with the robot so much that we couldn't compare between them. What we could do was analyzing the data looking for similarities and differences, to create preliminary user profiles, on which the robot could adapt its interaction in the future. All videos and logging files were watched and we looked at similar behavior between the participants on aspects as speech and touch interaction (time spent, manner of interaction, extravert behaviour etc.).

5.1 Tests

The self-efficacy test is excluded because most children had some difficulty filling it in. Furthermore, the test took too long to do a pre- and post test.

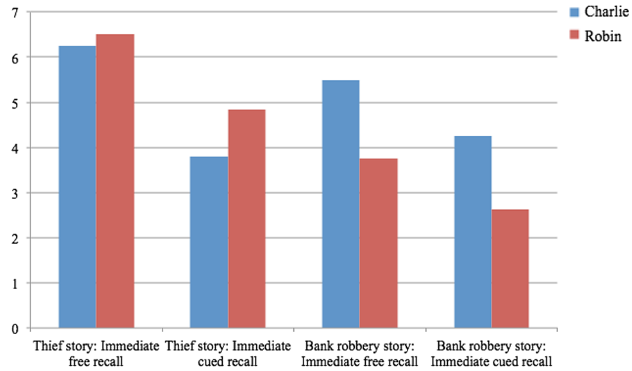


Fig. 5 Story recall comparison between Charlie and Robin

Knowledge test Questions 7, 8 en 18 are excluded because we noticed that multiple answers were correct. A paired sample t-test shows that there was a significant difference in knowledge acquisition between the pre- and post test for the questions that were presented during the experiment (1-24). First session M=11.35, SE=.77, second session M=13.7, SE=.66 and a paired t test t(16)=5.6, p<0.001 (2-tailed). The final eight questions (25-32) did not show significant improvement t(16) =1.19, p=.25 with M=5.94 and SE=.34 for the first session and M=6.29 and SE=.44 for the second session.

Memory test We did an independent samples t-test to test whether there is a significant effect of the robots in the immediate free recall and in the immediate cued recall (see Fig. 5). There are no significant differences assessed between the scores reached after the stories told by Charlie and the scores reached after the story told by Robin in the immediate free recall ($p = .114$, $p = .521$) and in the immediate cued recall ($p = .869$, $p = .306$).

5.2 Questionnaires

Fun We had separate questions on fun with robot, quiz and sorting game. Over the sessions these did not change significantly. The same was true for the separate questions on interaction with the robot (see Table 2).

Self-determination For the self-determination questionnaire we aggregated the questions related to competence, to relatedness and to autonomy per session.

Competence Overall, 49 % of the children rated their feeling of competence a 7 (highest) and only 4 % rated their competence under 4. In session 2 this was 56 and 7 % and in the third session 50 and 4 %. This means that no improvement was possible for almost half of the children and only very little for the children who scored initially under 7.

Relatedness We performed the same procedure as for competence and counted the number of times a 7 (highest) was chosen. 69 % of the time children felt very related to the robot

Table 2 Fun questionnaire means and (SD) [* 1 NA due to technical failure sorting game, ** for quiz: 2 NA and 6 children who filled in something while they didn't play the quiz, for sorting game: 3 NA and 6 children who filled something in while they didn't play the sorting game, *** 1 NA (missed question)]

Question (scale)	Session 1	Session 2	Session 3
How much fun did you find Charlie the robot? (1–7)	6.5 (0.87)	6.8 (0.43)	6.8 (0.44)
How much fun did you find the quiz? (1–7)	5.8 (0.88)	6.2 (0.66)	5.7 (1.16)**
How much fun did you find the sorting game (1–7)	5.9 (1.14)	6.1 (1.12)*	5.7 (0.73)**
How friendly did you find Charlie the robot (1–4)	3.9 (0.33)	3.8 (0.39)	6.1 (0.40)***
How well could you play together with Charlie the robot? (1–4)	3.6 (0.51)	3.8 (0.39)	3.8 (0.47)
How "cosy" is Charlie? (1–4)	3.7 (0.44)	3.9 (0.33)	3.8 (0.62)
How warm (hospitable) is Charlie? (1–4)	3.5 (0.62)	3.4 (0.61)	3.8 (1.35)

and only 6 % chose a rating under 4. In the second session this was 76 and 1 % and the third session 74 and 3 %. So as with competence the ratings were already so high in the first session there was little room for improvement, 54 % of the questions were rated a 7 on all three sessions.

Autonomy The autonomy was rated a 7 (highest) for 38 % of the time in the first session (15 % under 4), 44 % in the second session (6 % under 4) and 53 % in the third session (6 % under 4). Because of this increase we performed a Friedman test, but this was not significant $p=0.29$ ($df=2$, $chi^2=2.45$).

5.3 Observations

Game Preference In the second session 9 of the 17 children chose the sorting game as their favorite and 8 chose quiz and they also agreed starting with this game. In the third session 8 children chose to play the sorting game and 9 the quiz. 16 of the 17 children chose the game they preferred in the second session to play in the third. Only one child switched from sorting game to quiz.

Video and logging analysis From the video and logging analysis we extracted five user profiles as shown in the following and we did some additional observations.

User profiles The profiles were based on observations made during the experiment itself and observations from the videos afterwards. During the experiment the wizard, who was the same in almost all sessions, made notes about the behavior of the child in the experiment. Afterwards the same person identified some aspects, based on the notes and rewatching a few videos on which the children could be categorized in profiles. The scoring aspects were discussed with colleagues. Then taking these aspects all sessions were watched and scored. Aspects we looked at were related to dialog and actions of the robot, e.g. naming the child, falling. But also to dialog and actions of the child, e.g. reaction on falling of the robot, attention towards robot, time spent in activities, talking with robot (only telling or also listening), walking with robot, reaction on diabetes related questions. Finally we did also some general observations about the child, e.g. happy, open, shy, technology minded.

1. *Children who "just deal with it"(pp3, pp10, pp12, pp13, pp16)* In this group there are children who know very much about diabetes and how to deal with it. They can tell about it in an open manner, even about the difficult parts (see Table 3). They seem to feel good and do many things on their own. In the group of children who "just deal with it" there are also children whose parents have diabetes. The children who indicated that their parents have diabetes seem to be much more relaxed and open for diabetes related questions and providing information to Charlie in a positive way. Diabetes for these children seems to be a shared (and not problematic) lifestyle together with a parent.
2. *Children who feel to fall outside the group(pp2, pp9, pp11)* Children who seem to feel not that comfortable yet with having diabetes and the integration of it in their life belong to this group. Different reasons can be listed for this feeling. For example when children do not know enough about their diabetes, cannot connect the consequences of the diabetes to their feelings and are therefore more dependent on their parents. In the interaction this becomes clear by difficulty answering the open questions related to diabetes. They also see Charlie immediately as a friend, this is shown by having a picture of Charlie above the bed at home (pp2), having lots of empathy for Charlie when it falls (pp9) and more then passing interest in how many friends Charlie has (pp11).
3. *Children who are afraid to make errors(pp4, pp5, pp14, pp17)* When children look away very often during interaction, give answers which are not consistent with their behavior or are ashamed to say anything, it seems that children react only like that because of someone listening or watching (for example Table 4). These children seem not that sure in what they know about diabetes and do not dare to say something, because it could be wrong.

Table 3 Example of 'children who "just deal with it"'

Kind (pp12)	Ik ga heel goed opletten, wat ik eet en ik kijk goed op de verpakkingen. En dan onthou ik dan. Als ik bijvoorbeeld bij Sinterklaas pepernoten wil eten, weet ik hoeveel in 50 g zit en dan hou ik dat in mijn hoofd als ik dan de volgende keer 100 g wil eten weet ik dat dubbele moet doen
Charlie	Oh, wat goed zeg! Nu ik dat weet, kan ik het ook aan andere kinderen leren
Kind (glimlacht)	Oh dat is fijn!
Translated	
Child	"I'm very careful with what I eat, look on the packaging and remember that. So If I want to eat ginger nuts at Sinterklaas for example I know how much sugar there is in 50 g. I keep that in mind and when I want to eat 100 g I know that I have to do twice as much insulin"
Charlie	"Oh, great! Since I know that now, I can tell it to other kids"
Child (smiling)	"Oh, that is good"

Table 4 Example of 'children who are afraid to make errors'

Kind (lijkt arrogant/onzeker en kijkt vaak weg) (pp4)	Nou, als ik iets wil eten, dan spuit ik gewoon
Translated	
Child (seems arrogant/unsure, often looking away)	So, when I want to eat something, I just inject insulin

4. *Children who are shy (pp7, pp8, pp16)* These children take a longer time to tell something or do something with the robot. Often they whisper their answers, or just laugh a bit uncomfortable.
5. *Children who have difficulty with multitasking (pp1, pp5, pp6, pp17)* Some children in this experiment were still very young and had difficulties with talking with Charlie and playing the games at the same time. Sometimes these children could not read the quiz questions themselves. The experiment leader plays a big role in these interactions. Social desirable behavior is almost unpreventable in those situations. In the most cases they also know less about diabetes than the other children and do less diabetes related actions on their own.

Other observations In general some children touch the robot from the first meeting on, curious about how it feels. Especially in the last session Charlie gets many questions of how it works. All children are interested in unpredictable facts about Charlie as for example the name and colors of Charlie's soccer club and the outcome of the last game. Furthermore, compliments seem to support all children: They

Fig. 6 Drawing and paper craft gifts

react positively on them, some react more reserved whereas others give the robot compliments in return immediately.

Walking the robot is not very easy, at least not to bump into anything, but it is appreciated by most and when it goes not fast enough they just carry it to the intended spot. Also the falling seemed to support most children in feeling useful, but not all children liked to help the robot after it fell. All children had to help the robot to the other activities and all children experienced at least one fall during their three sessions. For some children this occurred more often than for others. Our feeling was that although helping to stand up was beneficial the falling had a negative influence when it occurred often.

In the dialogs we saw some progression in what was disclosed towards the robot, they really wanted to tell the robot about their experiences in between the visits. Very noteworthy is that 4 children gave a present to the robot (drawing, paper craft, loom bracelet and World Football Cup goodie) (see for example Fig. 6).

5.4 Feedback After Evaluation

At the moment of completing this paper the experiment has finished a year and a half ago; since that time we received great feedback from parents and medical staff. Parents have told us of more independence since the three 20-min interactions session. Medical staff tells us that children still ask when the robot returns and that they notice children are more at ease at the hospital since the experiment. In follow up contacts we noticed that parents, children and medical staff are more willing to participate in a follow up study than they were to participate in this study. This is also apparent in the fact that the Meander Medical Centre is now part of the H2020 project PAL that also looks at the use of the robot, in physical and virtual form, for children with diabetes.

6 Conclusions and Discussion of the Evaluation Results

6.1 Tests

After negative experiences with other questionnaires, we decided to use this self-efficacy questionnaire with the sorting cards. This method seems to work well: It encourages the children to think about their answers and vary them. But the questionnaire was not enough adapted to the target age and took too long to fill in. So although it did not have the desired result now, we would like to refine it and use it as pre- and posttest for self-efficacy in the future. In the Netherlands there is a list of "Know and Do" objectives for different age groups (6/7, 8/9, etc.); we are looking in to using this to measure the level of self-management. Of course we will also look for alternatives to measure variation in self-efficacy related to diabetes over time. It should be noted that parents and medical personnel indicated (after the experiment) that self-efficacy was improved. One of the parents for instance told that their daughter made more decisions on her own, like adapting the insulin before a meal because she wouldn't eat a lot of it. The parent said that the fact the robot made errors did have a positive effect. Furthermore, although not significant, there was an increase in autonomy according to the questionnaire.

The knowledge test had good results, but improvements are possible. Some (more interesting) questions had multiple possible answers, because in many situations there are multiple solutions for the problem at hand for diabetics (this is just one of the things we want to learn the children). Also the reaction to high or low bloodsugar is dependent on the situation: Illness, stress, physical activity and food influence the bloodsugar and to keep the variation at a minimum it is necessary to know why the body reacted in this way to come to the best reaction. Furthermore, we noticed that children answered lifestyle questions truthfully. So when asked how they handled a situation like telling a parent of a friend they had diabetes, they did not provide the "correct" answer, which was very obvious ("I do x because then I show I'm the boss of diabetes"), but said they rather not tell because it would make them different. We were very suprised, but also happy with this. We rather have the answer about how they handle such a situation so that we can make them understand why they should change behaviour than that they provide the "correct" answer.

The memory test did not result in a significant difference between the familiar (Charlie) and unfamilair robot (Robin), but we did see some opportunities to improve the test. First we need to make absolutely sure that both robots are equally understandable, while speaking with different voices and we should use a validated, for the specific age group, verbal memory test.

6.2 Questionnaires

All questionnaires suffered from the same problem, a ceiling effect. A score below the 6 was low which makes it impossible to have an increase over time. Next to this we saw that the sorting of cards in the self-efficacy test had a positive effect on thinking about a question, whereas some items of the questionnaires stimulated putting crosses automatically. This could be seen for instance in the questions of session 3 where many of the children (12 out of 17) answered questions about the activity they did not perform. It keeps being a challenge to have questionnaires that are informative, but they are still an important measurement method, so we will keep adapting them and hope to create an informative questionnaire. Furthermore, we will look further into ways to decreasing the effect, like make the answering more tangible (e.g. no cross but moving something to the answer), more forced choice, implicit association tests, providing parents with questionnaires for some effects, longer evaluation periods, and more.

6.3 Observations

6.3.1 Game Preference

It was nice to see that some children preferred the quiz while others preferred the sorting game. This encourages us to proceed with having different activities that are performed with the same robot to reach the same objective and that which activity is performed depends on the child's preference, state and current objective.

6.3.2 Video and Logging Analysis

User profiles The user profiles indicated in this experiment are a starting point for us to focus on some parts of the interaction and see if we can recognize these same profiles in another experiment or that they need to be adapted. The profiles as they are now, are solely based on the interpretations of one coder and thus need to be verified. After a set of stable user profiles is identified we want to use these profiles in the future to make a fast adaptation to the user possible. Below we provide per user profile a first idea on how the user profile influences the adaptation.

1. *Children who "just deal with it" (pp3, pp10, pp12, pp13, pp16)* The robot can tell the children who are more uncomfortable with their diabetes how these children could deal with it. The children mention that the robot needs to know more and get a teacher role which can give them more self-confidence. This group is challenging for the interaction because in particular the children who are easily comfortable in the interaction with the robot are also the first who get bored by the robot and its games. Fortunately, this group seems to be interested in a robot and how it works. In the interaction with this group this could be taken advantage of. Although the children in this group are already quite confident with their diabetes they might benefit from short interactions to provide

them with a bit more confidence to take the next step in self-management. This idea is fed by the feedback we got from some parents with children in this group.

2. *Children who feel to fall outside the group (pp2, pp9, pp11)*: For these children the robot has to be a real friend. Remembering what the children said adds great value. It seems to be nice especially for these children to share interests with the robot, for example playing cards (pp2) or wearing bracelets (pp11). The robot should combine friendship and dealing with diabetes. To not break the bonding with the child, the robot has to be careful with its questions and for example not ask a question like "What do you do with Santa-Claus, so many weird food, how do you deal with it?" in the beginning of the interaction to not bring the child in an unpleasant situation.

3. *Children who are afraid to make errors (pp4, pp5, pp14, pp17)* The robot can show the children that it doesn't matter to make errors by making errors itself. It can give the children self-confidence through playing the games and praise when the children did something good. The bonding can grow and the child can grow too.

4. *Children who are shy (pp7, pp8, pp16)* When children are very shy, the robot should be patient, and should play and walk with the children instead of talking too much. Some children need more time to talk about difficult issues. The robot has to try to estimate such children's state and help them managing their diabetes without being too pushy.

5. *Children who have difficulty with multitasking (pp1, pp5, pp6, pp17)* To improve self-efficacy and knowledge with children who have difficulty with multitasking, the robot should catch the attention and hold the attention of the children. That is very challenging especially because children are very good in ignoring other things when they are engrossed in something else. The bonding with the robot could grow in first instance via playing and later via dialogue.

6.4 Feedback After Evaluation

The feedback after evaluation provided us with lots of information, but in a semi-structured manner. Our experiences during this experiment with small talk with parents and health care professionals when they were watching the sessions and afterwards has shown us the importance of involving them in a more structured manner. In the future we will do this by involving them more in the design and evaluation via focus groups, structured interviews, participation in the experiment and questionnaires.

7 General Conclusion and Discussion

7.1 Main Outcomes

Overall, the general scenario for educational and enjoying child–robot activities during returning hospital visits, proved to capture the lessons learned well. The children had very positive experiences in the three sessions of almost one hour (i.e., quiz, sorting game and video watching, and small talk and walking). The children, but also their parents and formal caregivers, showed positive experiences. Children enjoyed the variety of activities, built a relationship with the robot and had a small knowledge gain. Parents and hospital staff pointed out that the robot had positive effects on child's mood and openness, which may be helpful for self-management. Based on the evaluation results, we derived five user profiles for further personalization of the robot, and general requirements for mediating the support of parents and caregivers.

More specifically, personalization to developmental age, interests and objectives of a specific child, proves to be important for both the interaction as the questions asked. Furthermore, we should not only focus on improving self-efficacy of the child, but also on improving confidence of the parents in their child. Many of the parents were over-protective. Involvement of children, parents and medical staff is thus essential. Fortunately we have seen that formal and informal caregivers changed from skeptic to enthusiastic, based on the reactions of the children who showed increased self-management and more positive hospital experience. The robot showed to have a new role for self-management that is different from that of the caregiver and peer. If the long-term effects follow the same line is to be seen, the positive attention the children received now in relation to their illness can already explain many of the beneficial effects of the robot intervention. On the other hand, if we can have such an effect with three 20-min sessions with a robot it is worth the effort.

7.2 Importance of Evaluation "in the Wild"

Performing an evaluation with children with diabetes in a care environment provided us with knowledge and experiences we could not have acquired doing evaluations at schools. We noticed that diabetic children's experiences with the robot differed from "healthy" children. They seemed to be more open for social interaction with the robot and also the fact that the robot was not all-knowing and dependent on the child seemed to influence these children more than healthy children. This was the first evaluation the robot received gifts from children, which shows that there is some kind of bond/relationship forming. The shared space of child and robot added to this experience as did the dependence of the robot on the child when it fell or had to go to another activity.

Because the children were brought to the experiment by their parents who often waited in the same room as the experiment leader (outside the experiment room) it was the first time we could interact with parents for a longer period. We of course knew that parents of children in this age group are of a huge influence on the child, and that this might be even more so with chronically ill children, seeing it first hand does change how you look at this influence. There were parents who already said at the beginning that they did not know if their child could perform well in the evaluation and we saw this back in the shyness of the mentioned child that changed a lot during the three sessions. Furthermore, having a child with diabetes has tremendous influence on family life. So caretakers and social environment influence the child, but the child also influences his or her environment. In future research we will take the influence and experiences of family and social environment into account.

The evaluation took place in the room next to the coffee corner of the hospital staff involved in the care of the diabetic children. This was great because they could look through a window and see what was happening, but also talk to parents and experiment leaders while getting coffee and thereby getting a better feel of the aim of the robot. They could see the enjoyment of the children, and also see and hear that the robot will not substitute them.

One of the main challenges we found is that because of the bond the children seemed to form and the things they discussed with the robot it did not feel ethically right to strictly follow protocol. For example when a child discussed his or her problems with diabetes because of a birthday party the robot did not react with “I don’t understand”, but the wizard typed in a relevant comment for the robot to say. Due to this and technical problems, no session was the same and the applicability of inferential statistics was limited.

7.3 Future Work

This evaluation showed that parents, medical staff and children enjoyed working with the robot and saw advantages of the use. The next step is now to develop a prototype that can stand alone, might also be used at home (in virtual form) because there are only a few hospital visits, and that involves all stakeholders. This means we need at least a solution to deal with speech recognition and dialog management, personalization on at least child interests, developmental age and objectives towards self-management, and evaluating effectiveness so that care institutions can argue for the costs of using the robot. Currently, these aspects are being addressed in the European H2020 project PAL (www.pal4u.eu).

Acknowledgments First we would like to thank the children and their parents, who had to come three times outside of regular appointments. Without the Meander Medical Centre and the support of the diabetic team there, especially Roos Nuboer and Mirjam Schouten who made the whole team enthousiastic and made sure that our schedule was full and tight, this experiment would not have been possible. Finally, not only the authors of this paper contributed to this experiment, but also Joris Janssen was indispensable as experiment leader and Bert Bierman for his never ending and never failing technical support. This work is (partially) funded by the EU FP7 ALIZ-E project (Grant Number 248116) and the H2020 PAL project (Grant Number 643783).

References

1. Barakova EI, Lourens T (2010) Expressing and interpreting emotional movements in social games with robots. Pers Ubiquitous Comput 14(5):457–467
2. Belpaeme T, Baxter PE, Read R, Wood R, Cuayáhuitl H, Kiefer B, Racioppa S, Kruijff-Korbayová I, Athanasopoulos G, Enescu V et al (2012) Multimodal child–robot interaction: building social bonds. J Hum Robot Interact 1(2):33–53
3. Belpaeme T, Baxter P, De Greeff J, Kennedy J, Read R, Looije R, Neerincx M, Baroni I, Zelati MC (2013) Child–robot interaction: perspectives and challenges. In: International conference on social robotics. Springer, pp 452–459
4. Bickmore T, Cassell J (1999) Small talk and conversational storytelling in embodied conversational interface agents. In: AAAI fall symposium on narrative intelligence, pp 87–92
5. Blanson Henkemans OA, Van Dooren MM, Bierman BP, Janssen J, Looije R, De Vries JL, Neerincx M (2016) Design and evaluation of a personal robot playing a self-management education game with children with diabetes type 1. International Journal of Human–Computer Studies. (Submitted)
6. Cohen I, Looije R, Neerincx M (2014) Child’s perception of robot’s emotions: effects of platform, context and experience. Int J Soc Robot 6(4):507–518
7. Csikszentmihalyi M, Abuhamdeh S, Nakamura J (2005) Handbook of competence and motivation, chap 32:Flow. The Guilford Press, New York, pp 598–698
8. Danne T, Kordonouri O (2007) What is so different about diabetes in children. Diabetes Voice 52:16–19
9. Deci EL, Ryan RM (2002) Handbook of self-determination research. University Rochester Press, Rochester
10. Fleming ND, Mills C (1992) Not another inventory, rather a catalyst for reflection. Improve Academy 11:137
11. Fogg B (2002) Persuasive technology: using computers to change what we think and do. Ubiquity 2002(December):5
12. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. Robot Auton Syst 42(3):143–166
13. Gale EAM (2012) Epidemiology of type 1 diabetes. Diapedia 13:21042
14. Garattini L, Chiaffarino F, Cornago D, Coscelli C, Parazzini F, del COsti e Risorse del DiabeteSGRRE. et al (2004) Direct medical costs unequivocally related to diabetes in italian specialized centers. Eur J Health Econ Form HEPAC 5(1):15–21
15. Greeff JD, Janssen J, Looije R, Mioch T, Alpay L, Neerincx M, Baxter P, Belpaeme T (2013) Activity switching in child–robot

interaction: a hospital case study. In: 5th international conference on social robotics, ICSR 2013, 27 October 2013 through 29 October 2013, Bristol, 8239 LNAI, pp 585–586
16. Janssen JB, van der Wal CC, Neerincx MA, Looije R (2011) Motivating children to learn arithmetic with an adaptive robot game. In: Proceedings of the Third international conference on social robotics, Springer, pp 153–162
17. Karoly P, Bay RC (1990) Diabetes self-care goals and their relation to children's metabolic control. J Pediatr Psychol 15(1):83–95
18. Kennedy J, Baxter P, Belpaeme T (2015) The robot who tried too hard: social behaviour of a robot tutor can negatively affect child learning. In: Proceedings of the HRI, vol 15
19. Kessens JM, Neerincx MA, Looije R, Kroes M, Bloothooft G (2009) Facial and vocal emotion expression of a personal computer assistant to engage, educate and motivate children. In: 3rd IEEE international conference on affective computing and intelligent interaction (ACII), Amsterdam
20. Markland D, Ryan RM, Tobin VJ, Rollnick S (2005) Motivational interviewing and self-determination theory. J Soc Clin Psychol 24(6):811–831
21. Miller W, Rollnick S (2002) Motivational interviewing: preparing people for change. The Guilford Press, New York
22. Neerincx M, Lindenberg J (2008) Situated cognitive engineering for complex task environments. In: Schraagen JM, Militello L, Ormerod T, Lipshitz R (eds) Situated cognitive engineering for complex task environments. Aldershot, Ashgate Publishing Limited, p coming soon
23. Niemiec CP, Ryan RM (2009) Autonomy, competence, and relatedness in the classroom applying self-determination theory to educational practice. Theory Res Educ 7(2):133–144
24. Oinas-Kukkonen H (2010) Behavior change support systems: a research model and agenda. In: Ploug T, Hasle P, Oinas-Kukkonen H (eds) Persuasive technology. Springer, Berlin, pp 4–14
25. Papangelis A, Zhao R, Cassell J (2014) Towards a computational architecture of dyadic rapport management for virtual agents. In: Brinkman W-P, Broekens J, Heylen D (eds) Intelligent virtual agents. Springer, Berlin, pp 320–324
26. Patterson CC, Dahlquist GG, Gyürüs E, Green A, Soltész G, Group ES et al (2009) Incidence trends for childhood type 1 diabetes in europe during 1989–2003 and predicted new cases 2005–2020: a multicentre prospective registration study. The Lancet 373(9680):2027–2033
27. Peyrot M (2008) How is diabetes perceived? The results of the dawn youth survey. Diabetes Voice 53:9–13
28. Read JC, MacFarlane S (2006) Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In: Proceedings of the 2006 conference on Interaction design and children, ACM, pp 81–88
29. Robins B, Dautenhahn K, Dickerson P (2009) From isolation to communication: a case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot. In: Advances in computer–human interactions, 2009. ACHI'09. Second international conferences on, IEEE, pp 205–211
30. Rollnick S, Mason P, Butler C (1999) Health behavior change: a guide for practitioners. Elsevier Health Sciences, London
31. Ros R, Demiris Y (2013) Creative dance: an approach for social interaction between robots and children. In: Salah AA, Lepri B (eds) Human behavior understanding. Springer, Berlin
32. Schadenberg B (2012) Modelling the userGs skill and performance with the use of a bayesian rating system. Master's thesis, Rijksuniversiteit Groningen
33. Short E, Swift-Spong K, Greczek J, Ramachandran A, Litoiu A, Grigore EC, Feil-Seifer D, Shuster S, Lee JJ, Huang S, et al (2014) How to train your dragonbot: socially assistive robots for teaching children about nutrition through play. In: The 23rd IEEE international symposium on robot and human interactive communication, 2014 RO-MAN, IEEE, pp 924–929
34. Solms L (2014) An exploration of the effects of touch on social bonding between robot and child (unpublished)
35. Tanaka F, Kimura T (2010) Care-receiving robot as a tool of teachers in child education. Interact Stud 11(2):263–268
36. Tejada S, Traft N, Hutson M, Bufford H, Dooner M, Hanson J, Radler A, Mauer G (2006) Educational robots: three models for the research of learning theories and human–robot interaction, pp 70–76
37. Tielman M, Neerincx M, Meyer JJ, Looije R (2014) Adaptive emotional expression in robot–child interaction. In: Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction, ACM, pp 407–414
38. Van Dam I (2013) Meet my new robot best friend: an exploration of the effects of personality traits in a robot on enhancing friendship. Master's thesis, Universiteit Utrecht
39. Van Der Drift EJ, Beun RJ, Looije R, Blanson Henkemans OA, Neerincx MA (2014) A remote social robot to motivate and support diabetic children in keeping a diary. In: Proceedings of the 2014 ACM/IEEE international conference on human–robot interaction, ACM, pp 463–470
40. Vygotsky LS (1980) Mind in society: the development of higher psychological processes. Harvard university press, Cambridge
41. Wechsler D, Laicardi C, Orsini A (1997) WAIS-R: Wechsler adult intelligence scale revised: manuale. OS
42. Wigdor N (2014) Conversational fillers for response delay amelioration in child–robot interaction. Master's thesis, University of Utrecht

Rosemarijn Looije is a researcher at TNO, department of Perceptual & Cognitive Systems. She has a BSc and MSc Artificial Intelligence—Human–Machine Interaction from the University of Groningen. She is project coordinator of the PAL project and pursuing her PhD on designing and evaluating a social robot for children with diabetes. Her interests lie in understanding why certain functionality works or doesn't work and how to pinpoint this by making decisions explicit. Furthermore she is interested in how to combine data sources to support personalization and present this to the end user.

Mark A. Neerincx is full professor Human-Centered Computing at the Delft University of Technology and principal scientist Perceptual & Cognitive Systems at TNO (Netherlands organisation of applied research). Recent projects focus on the situated cognitive engineering of electronic partners (ePartners) that support the social, cognitive and affective processes in human-automation collaboration to enhance performance, resilience, health and/or wellbeing. Examples are robotic and virtual assistants that help patients to cope with their chronic disease (e.g., diabetics) in different self-managements activities, and assistants that help older adults with dementia and care givers to establish positive experiences of life in the care centers. Important results are: (1) cognition and emotion models for performance and health support, (2) models of human-machine partnership for attuning assistance to the individual user and momentary usage context , (3) prototypes of electronic partners and social robots for self-efficacy and behavioral change, (4) cognitive engineering methods and tools , and (5) a diverse set of usability best practices.

Johanna K. Peters is an MSc student at the University of Groningen in Human–Machine Communication with a specialisation in Cognitive Engineering. She is currently doing her graduation project at Philips

Drachten. At the University of Twente she finished her BSc in Psychology with a specialisation in Human Factors and Engineering Psychology on a project at Volkswagen AG. She is particularly interested in being part of interdisciplinary teams to get the best results in innovations that have a focus on the interaction with the end user.

Olivier A. Blanson Henkemans is an applied researcher in e/mHealth for health promotion at TNO, department Child Health with a PhD in Human–Machine interaction from the Delft University of Technology. He works on usability engineering of highly interactive and intelligent systems in the health care domain. This concerns both hard- and software. He has a special interest in strong involvement of the end-users during the whole process from design to evaluation. He ensures this with the use of various user-centered design technniques such as task analysis, personas, scenarios, concepts and story boards, and requirement analysis. To support end user involvement with (special) user groups he masters a broad range of innovative methods, including but not limited to, creative sessions (e.g., image theater, draw and tell), stakeholder meetings and focus groups, literature reviews, large scale surveys and randomized controlled trials.

Master Thesis Artificial Intelligence

MyPAL: A Digital Diabetes Diary with a Responsive Social Avatar

Mike Ligthart, BSc

Radboud University Nijmegen

June 2016

Internal Supervisor:	dr. Pim Haselager *Radboud University Nijmegen*
External Supervisor:	prof. dr. Mark Neerincx *TNO and Delft University of Technology*
Reviewer:	dr. Khiet Truong *University of Twente*

Radboud University

TNO innovation for life

Abstract

Diabetes Mellitus type I is an incurable disease that can be diagnosed at a young age. A structured lifestyle, where insulin use, carbohydrate intake and blood glucose are regularly monitored, is the only path to a relatively normal life. Children and their parents must remain vigilant. This lifestyle is especially demanding for children and not every child is as good in self-management as they need to be. They can use some help with this.

MyPAL is a digital diabetes diary that children can use to record their insulin use, carbohydrate intake and blood glucose values as well as write something about their day and how they feel. With that information the children can more easily link their diabetes values, what they eat and how they feel together. With this insight they can manage their insulin use and diet more efficiently. Besides children also medical professionals, parents and researchers benefit from this information. For example, a diabetes nurse can improve the treatment plan, parents can get a better idea how their child is doing and researchers can investigate the relationship between food, mood and blood glucose values more closely.

The only constraint is that the required information is added regularly. And if children find something difficult it is consistently keeping a diary. One of the most mentioned causes of why children have trouble to consistency use a diabetes diary is the lack of motivation. MyPAL provides several solutions for this problem. Following the situated Cognitive Engineering (sCE) design method, myPAL is specifically designed to support the development of motivation.

The first deliverable included in this thesis are the system specifications of myPAL. These specifications are based on operational demands, human-factor knowledge, and technological principles. The self-determination theory, that identified the feeling of autonomy, competence and relatedness as antecedents of motivation, is the largest human-factor contribution. An avatar, a technological principle, is used to support those antecedents of motivation. By autonomously responding to the added content in a social fashion, e.g. matching the gestures and speech of the avatar appropriately to the mood of the child, children feel more supported in their competence and relatedness.

The second deliverable is a thorough evaluation of the avatar behaviors and its effects on the attitude of the children towards the robot, motivation support and performance. Performance is measured in terms of the amount and the consistency of the added content. A three-week user study with 13 children with diabetes was performed for this evaluation. Results show that almost all the avatar behaviors are picked up by the children and that those behaviors positively affect the motivation and performance of the children.

The final deliverables are two design recommendations that have been found to modulate the effectiveness of myPAL. The first is avatar *quality over quantity*. The avatar behavior must be appreciated by the children in order to be effective. Simply showing to avatar more does not increase the appreciation. The behavior must match the children's preferences. The second recommendation is *avatar sociability is key*. The more social the behavior of the avatar is, the more it is appreciated by the child and the more motivated the children are to add more content consistently.

The system specifications, evaluation and design recommendations bring the sustainable application of autonomous avatars in diabetes care a step closer to being realized.

Abstract (Dutch)

Diabetes Mellitus type I is een ongeneselijke ziekte dat zich al op jonge leeftijd kan openbaren. Een gestructureerde dagindeling, waarbij insulinegebruik, inname van koolhydraten en de bloedglucosewaarden regelmatig worden bijgehouden, is de enige weg richting een relatief normaal leven. Kinderen en hun ouders dienen voortdurend waakzaam te zijn. Deze manier van leven is voor kinderen in het bijzonder erg veeleisend. Niet elk kind is zelfredzaam genoeg. Ze kunnen hulp gebruiken om dat te worden.

MyPAL is een digitaal diabetesdagboek dat kinderen kunnen gebruiken om hun insulinegebruik, koolhydrateninname en bloedglucosewaarden bij te houden. Verder kunnen ze hun dagelijkse bezigheden en gevoel hierbij noteren. Met die informatie kunnen kinderen het verband leggen tussen hun diabeteswaarden, wat ze eten en hoe ze zich voelen. Met dat inzicht kunnen ze hun insuline inname en dieet beter reguleren. Naast kinderen kunnen ook medische professionals, ouders en onderzoekers profiteren van de data. Een diabetesverpleegkundige kan bijvoorbeeld het behandelplan bijstellen, ouders krijgen meer inzicht in hoe het gaat met hun kind en onderzoekers kunnen de relatie tussen voedsel, gevoel en diabeteswaarden nog gedetailleerder onderzoeken.

Een voorwaarde is echter wel dat de vereiste informatie regelmatig wordt toegevoegd aan het dagboek. Als kinderen ergens moeite mee hebben dan is het een dagboek consistent bijhouden. Eén van de meest genoemde redenen hiervoor is dat kinderen nauwelijks gemotiveerd zijn. MyPAL biedt meerdere oplossingen voor dit probleem. Met behulp van de 'situated Cognitive Engineering' ontwerpmethode is myPAL specifiek ontworpen om de motivatieontwikkeling bij kinderen te ondersteunen.

Het eerste wat deze scriptie oplevert zijn de systeemspecificaties voor myPAL. De specificaties zijn gefundeerd op de functionele eisen, human-factors kennis en technologische principes. De zelfbeschikkingstheorie, wat het gevoel van autonomie, competentie en verbondenheid heeft geïdentificeerd als voorlopers van motivatie, wordt ingebracht vanuit de human-factors kennis. Een avatar, als technologisch principe, kan vervolgens ingezet worden om die voorlopers van motivatie te ondersteunen. Door de avatar autonoom en op sociale wijze te laten reageren op de toegevoegde dagboekinhoud, bijvoorbeeld door de bewegingen en de spraak van de avatar af te stemmen op de gemoedsruststand van het kind, voelen kinderen zich competenter en meer verbonden.

Het tweede wat deze scriptie oplevert is een grondige evaluatie van het gedrag van de avatar en zijn effecten op de attitude van de kinderen ten opzichte van de avatar, ontwikkeling van motivatie en de prestatie van de kinderen. Een experiment van drie weken met dertien kinderen met diabetes is uitgevoerd ten behoeve van de evaluatie. De resultaten tonen aan dat het meeste van de avatargedragingen worden opgepikt door de kinderen. Belangrijker nog, de opgepikte gedragingen hebben een positief effect op de motivatie en prestatie van de kinderen.

Het laatste wat deze scriptie oplevert zijn twee aanbevelingen die van invloed zijn op de effectiviteit van myPAL. De eerste is *kwaliteit boven kwantiteit*. Het avatargedrag moet gewaardeerd worden door de kinderen om effectief te kunnen zijn. Simpelweg het laten zien van de avatar maakt het niet meer gewaardeerd. Het gedrag moet passen bij de voorkeuren van de kinderen. De tweede

aanbeveling is: *het sociale vermogen van de avatar is de sleutel tot succes.* Des te sociale het gedrag van de avatar, des te meer het wordt gewaardeerd door de kinderen en des te meer de kinderen gemotiveerd zijn om het dagboek consistent van inhoud te voorzien.

De systeemspecificaties, de grondige evaluatie en de ontwerpsuggesties brengen de duurzame inzet van autonome avatars in de diabeteszorg een stap dichter bij de werkelijkheid.

Table of contents

1. Introduction

More than half a million children suffer from Diabetes Mellitus type I worldwide. This number increases every year. A quarter of the children with Diabetes live in Europe. Children benefit from a structured lifestyle where insulin use, carbohydrate intake, and blood glucose are regularly monitored. Starting at a young age with a well-balanced self-management plan including self-reports, physical activity, and a healthy diet increases the changes for a relatively normal life [1].

Helping children with diabetes to increase their self-management skills is the path to a better quality of life. This is the main force that drives the Horizon2020 project: Personal Assistant for a healthy Lifestyle (PAL). The `personal assistant' comes in different forms depending on the user, location and context. For example, for children visiting the hospital the personal assistant is a robot while at home it is a virtual version of that robot. At the hospital the children learn more about diabetes together with the robot in a playful fashion. At home children can use various mobile health applications. In those apps the robot is virtually present. Besides children also their parents and medical professionals can use the PAL-system. The project is carried out by a European consortium with partners from The Netherlands, Italy, Germany and the United Kingdom. Research institute TNO has the role of project coordinator [2].

My master graduation project is a part of the PAL project. During a 10-month internship at TNO Soesterberg I developed one of the mobile health apps that will be part of PAL, namely a digital diabetes diary called myPAL. Research shows that keeping a digital diabetes diary, including for example insulin use and carbohydrate intake, result in more insight and empowerment for the patient [3], [4]. Furthermore, it enables caregivers to make more informed adaptations to the treatment plan [5]. For adults dozens of mobile apps are available to keep a diabetes diary [6] but for children different design choices are necessary. For example, motivating children to keep a diary on a regular basis is challenging. Coercing a child often is counterproductive [7].

In the first half of this thesis I describe the groundwork that is required to design, build and evaluate myPAL: a digital diabetes diary with a responsive avatar. The goal of myPAL is to support the motivation of children to add content to the diary as consistently as possible. Human-factor knowledge and technological principles come together to achieve this goal. The two main building blocks are the self-determination theory, which is a framework for supporting intrinsic motivation, and a social responsive avatar. Using rapid-prototyping techniques and a larger pilot study a suitable configuration of all the building blocks is created in the form of system design specifications and a prototype. The design process mainly revolved around finding suitable ways of supporting motivation with myPAL.

If previous research shows us anything it is that children respond differently to systems like myPAL [5]. Responsive social systems vary in effectiveness and are not appreciated by everybody equally. The keyword for tackling these problems is personalization. If responsive systems were able to adapt their behavior to a specific user a higher effectiveness and appreciation is to be expected. The thing is, we are not there yet. More insight is needed into how children respond to a responsive social system such as myPAL.

In the second half of this thesis I describe a three-week user study of children using myPAL. I monitored the development of motivation and the performance over time. Performance is defined by the amount and the consistency of the content added to the diary. The first aim for this user study is to evaluate the design. Does the implemented avatar behavior get picked up by the children? Does the avatar behavior influence motivation and performance? The second aim of the user study is to achieve more insight in how children respond to myPAL in terms of performance and motivational development. To do this I will categorize the children based on their motivation and performance. Can we subsequently identify predictors that are able to indicate whether children are currently on a path for a high performance or not?

This thesis is structured as follows. In the next chapter fundamental concepts relevant for this thesis are introduced in more detail. First, the impact of diabetes on children is introduced and I motivate why this research can have an impact on their life's. Secondly, the position of myPAL in the PAL project is discussed. In the third section a definition is given of social robots in the health care domain. Because we are doing research with children it is important to design the prototype and the user study accordingly. In the fourth background section I reflect on several ethical implications of doing research with children. In the final background section, I introduce, situated Cognitive Engineering (sCE), the user-centered design method I used to develop myPAL.

A technical solution is never stands alone; it is always embedded in a certain context. The context provides operational demands for the solution. An important principle of user-centered design is to really take the perspective of the user. Therefore, besides technological principles also human factor knowledge is included in the foundation of the solution. The foundation, including design and research questions, is further discussed in chapter 3.

The design cycle of sCE continues with extracting specifications for the system from the foundation. These specifications, discussed in chapter 4, lead to a prototype which is described in chapter 5. The pilot study and the three-week user study are discussed in chapter 6. The formulated design and research questions are also evaluated in chapter 6. Finally, the whole process as well as the overall results are evaluated in chapter 7.

UNIVERSITY OF TWENTE

ANALYSIS OF SYSTEM USAGE AND KNOWLEDGE DEVELOPMENT OF THE CURRENT PAL SYSTEM FOR CHILDREN WITH TYPE 1 DIABETES MELLITUS.

Master Thesis

By

Anika Boelhouwer

Department of Cognitive Psychology and Ergonomics

For the degree of:

Master of Science, Psychology

Supervisors:

Martin Schmettow

Jan-Maarten Schraagen

Mark Neerincx

30-09-2016

Abstract

Children who are diagnosed with type 1 diabetes mellitus need to learn a lot about diabetes and self-management in a short period of time. A large problem in the support of this process is that health institutions cannot provide help at any given moment in the child life and are bounded by set face to face appointments. While digital interventions may address this issue by providing help and knowledge online which may be used at all times, this help and knowledge is general and not tailored to the individual. Also, actual usage of (digital) diabetes interventions has shown to be either extremely low or quickly decreasing. The Personal Assistant for a healthy Lifestyle project (PAL) strives to address these issues by providing a digital application with personalised communication and content. This study evaluated the current PAL application during a prolonged period of time with children diagnosed with type 1 diabetes mellitus between the ages of 6 and 12 years old. The main goals were to identify trends and possible predictors for both system usage and diabetes knowledge development. Three main trends were found in the system usage in which the majority of the users showed an overall low usage or quickly decreasing usage. A small number of users showed continuous and consistent usage throughout the entire experiment. As the personalisation was only minimally implemented the results are in line with common (digital) diabetes interventions. The results did not allow us to explore possible system usage and knowledge development predictors. They do however provide a solid baseline for further versions of the system in which the personalisation is further implemented. The main recommendations are to focus on the implementation of basic game design elements and personalised content to foster user engagement and continuous use. Maintaining the used measures (while adding some psychological predictors) and longitudinal experiment design will allow for comparative analysis in the further research cycles.

Table of contents

1. Introduction

During the end of the 20th and beginning of the 21st century, the number of children with type 1 diabetes mellitus has steadily increased (Gale, 2002). This disease has a great impact on both individual as societal level. For a patient, it can lead to severe short- and long term health complications such as nerve-, eye- and foot damage or cardiovascular diseases (Tsukuma et al., 1993). Furthermore, the onset of the disease can induce prolonged stress not only for the patient, but also for their parents and siblings as the recent study of Streisand, Swift, Wickmark, Chen and Holmes (2005) showed. For the society, an increase in patients who need intensive and chronic patient support and medication means an increased amount of healthcare costs, and need for healthcare professionals. Accurate self-management and continuous self-care is required of both the patient and parents in order to reduce risks on serious health care complications (Shrivastava, Shrivastava, & Ramasamy, 2013; Diabetes Control and Complications Trial Research Group, 1993) and to increase the patients and parents overall quality of life (Jaser et al., 2012). Fewer health problems due to poor self-management in diabetes patients may then also reduce societal health costs (Boren, Fitzner, Panhalkar, & Specker, 2009).

Although diabetes self-management skills play a large role in preventing health complications, acquiring these skills can be particularly hard for children as they are still developing physically, mentally and emotionally. They need to control their carbohydrate intake, physical activity, monitor blood sugar levels and manage the insulin that needs to be injected. Calculating carbohydrate intake and the needed insulin dosage, combined with self-care and other activities such as school and social life can be complex and overwhelming for (recently) diagnosed children. Even though they are supported by healthcare professionals, these appointments are periodic and can't provide tailored help at all times. As a result, taking over the diabetes management causes many parents to experience paediatric parenting stress. This may increase risk of mental health problems (Streisand, Swift, Wickmark, Chen, & Holmes, 2005), and negatively impact the child's self-management skills by reducing their autonomy (Streisand, Swift, Wickmark, Chen, & Holmes, 2005).

In order for diabetes education to be as effective and efficient as possible it should be available throughout the daily life of the child, and take the independent development of the child and his self-management skills into consideration. The PAL system, which consists of a NAO robot, mobile avatar of the robot, mobile quiz and mobile timeline, gives an opportunity for the child to learn about diabetes and diabetes management by play and social interaction. Children can get more insight in their disease and may improve their self-management by adding their activities, meals, glucose values and emotions on a timeline. Also, their knowledge about diabetes may be improved by playing the quiz and answering diabetes related questions. Parents and health care professionals are supported by giving them up to date information of the child's health status and goals progress.

The Personal Assistant for a healthy Lifestyle project (PAL) of the European Horizon 2020 research program strives to induce active user engagement through personalization. The system will use the user's preferences, characteristics and learning goals to personalise communication and education through a robot and application and ensure a long term relationship between the robot (avatar) and the child (Janssen, van der Wal, Neerincx, & Looije, 2011). By providing a learning style that keeps adapting to the child's needs, and following the zone of proximal development, the child is challenged in his learning but not overwhelmed (Kozulin, Gindis, Ageyev, & Miller 2003). By automatically adapting to the child's preferences and needs, the robot might support the intrinsic motivation by elements of the self-determination theory: autonomy, competence and relatedness (Ryan & Deci, 2000). This theory is the base for many of the PAL project design choices to stimulate motivation. The children are supported in their autonomy, as they are presented with and asked how they would solve complex (diabetes related) situations in the quiz. Competence is stimulated by offering new tasks and goals for the child to reach. The relatedness to the social robot is created as it both behaves towards the preferences of the user, and engages in mutual self-disclosure to create a personal bond. It important to note that the current system includes a minimal amount of personalisation. This includes the incorporation of personal goals but excludes any personalized communication or feedback.

In the PAL system, a robot for initial play and support was chosen as a physical conversational agent as it was shown to have a considerable impact on initial motivation, feelings of relatedness and learning (Blanson Henkemans et al., 2013). As it is not practically feasible to supply a great number of children with a robot over a long period of time, a digital version of the robot and interaction was chosen in the form of an application (with a quiz and timeline) and an avatar of the robot. This combination ensures the motivational benefits of a physical robot, with the prolonged interaction possibilities of the application. The quiz was chosen as a first game within the application as this was identified as one of the most positively rated games during a previous PAL pilot study, while also providing a platform for validation and development of diabetes knowledge in the participating children (Blanson Henkemans et al., 2013). Other apps like MySugr (http://mysugr.com) have been using a similar construct of a quiz and timeline but have not incorporated personalised communication and learning.

Two main challenges arise in the development of the PAL system which need to be addressed for the system to be developed as effective as possible in further stages. The first challenge in the development of the PAL system is that it is unclear to what extend the current system contributes to actual knowledge development about diabetes and self-management in children. Also, there might be factors that directly influence differences in knowledge development. Some of these may be addressed by personalised communication to foster the most effective knowledge development for different kinds of users. For example, cognitive differences like preferred learning style have been shown to relate to differences in learning performances (Lynch, Woelfl, Steele, & Hanssen, 1998). One of the main goals of the PAL system is therefore: to improve diabetes management and knowledge by providing a personalised learning environment with tailored communication and learning challenges. The second challenge is that recent human computer interaction (HCI) research has revealed that the actual continuous usage of digital (health) interventions is scarce. Many patients show a severe decrease in usage as time passes or even show an initial low use of health interventions. For most intervention products to be beneficial, prolonged optimal use is deemed necessary for it to have a long-term effect. However, whether this prolonged use is actually performed by the patients is often not researched in a longitudinal study due to practical reasons like costs (Gerken, 2011). Any trends in system usage of the PAL system over time need to be identified to see if the system can accurately motivate children to keep using the system over a prolonged period. Also, in order to improve the personalization of the system communication and stimulate a prolonged use in children, possible predictors for system usage need to be identified. For example, perceived fear has been identified as a factor that promotes positive health behaviours (Venkatesh, Morris, Davis, & Davis, 2003). The perceived threat of their diabetes and possible complications may influence the motivation of a patient to perform positive health related behaviour, which in this case is using the PAL system.

To address these problems, this thesis strives to answer two main questions: "What kind of system usage and knowledge development patterns arise during a prolonged use of the current PAL system?" and "Which factors may contribute to individual differences in system usage and knowledge development during a longer period of use of the PAL system?". At the start and end of the experiment, the children will interact with the PAL robot and application at the hospital. During the experiment, children will be free to interact with the PAL application for four weeks at their own home. Any patterns in system usage and knowledge development are researched through repeated measuring of the system data during the experiment. This use of repeated measures will allow for inspection of individual system use and knowledge developments. In order to identify specific factors which may contribute to differences in system usage and knowledge development, this study will first perform a literature research to get a global overview of possible factors. Main criteria for these factors are that they may be (indirectly) addressed by the personalised system, or are general mediator variables to take into account. Including additional factors could be interesting in a research point of view but will result in an incomprehensible amount of factors and may most likely not be useful for the practical further development of the system.

In conclusion, this study will answer the following question: "What kind of patterns of system usage and knowledge development arise during prolonged voluntary use of the current PAL system, and which factors might contribute to individual variances in system use and knowledge development?". We answer this question by combining a literature research on possible predictors for system usage and knowledge development, with a longitudinal experiment with the current PAL system for the identification of patterns possible predictors for system usage and knowledge development. With this knowledge, the system may be further personalised to optimize the usage and learning experience for different users.

Master Thesis

Healthcare professionals gain control of children's diabetes self-management

Designing a healthcare management tool for healthcare professionals' assessment of T1DM knowledge, goals and development.

Submitted by
Jet Shin Hong
101 86271

In partial fulfillment
for the degree of:
MSc Information Studies
Track Human Centered Multimedia
University of Amsterdam
Faculty of Science

Thesis Supervisor:
Dr. Tobias Kuhn

Second Examiner:
Dr. Victor de Boer

Company Supervisor:
Dr. Olivier Blanson Henkemans

Company Supervisor:
Dr. Sylvia van der Pal

June 28, 2016

University of Amsterdam · TNO innovation for life · PAL · VU Vrije Universiteit Amsterdam

Healthcare professionals gain control of children's diabetes self-management

Designing a healthcare management tool for healthcare professionals' assessment of T1DM knowledge, goals and development.

Jet Shin Hong
University of Amsterdam
MSc Human Centered Multimedia

Abstract – Children aged 8 to 12 with diabetes type I are motivated to get involved in their diabetes management to reduce the impact of their illness on their short- and long-term health. Self-management of diabetes is an active and proactive process and it involves shifting and sharing responsibility for diabetes care tasks and decision-making in frequent collaboration with healthcare professionals. The research question this study sought to answer is: 'How can a healthcare management tool support healthcare professionals in guiding children with diabetes self-management involving a social actor (robot/avatar)?'. To answer this question, a prototype of a healthcare management tool was developed and evaluated with end users (diabetes nurses) and an important stakeholder (diabetes doctor), following the situated Cognitive Engineering approach. Overall, this prototype of a redesigned PAL Control was perceived positively by the healthcare professionals and the findings suggested that a combination of an assessment with a robot or its avatar, setting goals, selecting actions and the progression page, is a suitable and effective approach to healthcare professionals in guiding children with diabetes self-management. Healthcare professionals mentioned that this system has provided them support in making the consult with children and parents more meaningful due to the fact that they can understand their needs better on forehand. However, evaluations for a longer period of time is needed in order to validate if the needs are completely fulfilled. Nonetheless, useful suggestions were found during the evaluation of the prototype and provided important pointers for further development.

Keywords - Healthcare management, diabetes self-management, social robotics, usability.

1. Introduction

1.1 Diabetes self-management

Type 1 diabetes mellitus (T1DM) is the most common type of diabetes affecting children globally (Scott, 2013). With T1DM the body mistakenly destroys its own cells in the pancreas that produce insulin. Unfulfilled blood glucose targets are still a crucial concern (McKnight, 2015). Children aged 8 to 12 with diabetes type I are motivated to get involved in their diabetes management to reduce the impact of their illness on their short- and long-term health (Dedding, 2009). Self-management of diabetes is an active and proactive process and it involves shifting and sharing responsibility for diabetes care tasks and decision-making in frequent collaboration with healthcare professionals. Self-management of diabetes consists of many different activities related to e.g. monitoring metabolic control, dosing insulin, and regulating diet and exercise (Schilling et al., 2002).

When children grow older their self-management skills and knowledge will develop. Although their self-management performance improves accordingly, they are still developing cognitively and emotionally at this age. Therefore, children are not always able to apply their skills and knowledge optimally. Activities related to giving the correct amount of insulin or handling low glucose levels in unfamiliar situations can be demanding. Furthermore, to keep a good quality of life, children will have to obtain a balance between self-management activities and experiences in important aspects of life, such as school and social life (Dedding et al., 2004). From a medical perspective, this may cause mismanagement of diabetes (Snoek and Skinner, 2007).

1.2 PAL & PAL Control for healthcare professionals

In order to support children with diabetes self-management, the European PAL (Personal Assistant for healthy Lifestyle) project started in 2015. The goal of PAL is to assist the child, healthcare professional and parents, to jointly perform diabetes management. As such, the child aged between 7 and 14 learns to be more self-reliant before adolescence. Figure 1 provides an overview of the components in the PAL system: PAL Control, PAL Inform, MyPAL, PAL robot and its avatar.

This study focuses on the component PAL Control, which is the functionality designed for healthcare professionals in guiding the children with T1DM self-management (blue circle in Figure 1). From the healthcare professional perspective, PAL Control is intended to be 1) a gatekeeper for information on the young patients to personalise their healthcare, 2) a tool to author and control the PAL robot and its avatar and aid the child in their self-management, and 3) a tool to provide explanations to the informal caregivers (e.g. parents) on the desired activities of the children (PAL 2016). The current state of PAL Control (version: 17-02-16) enable healthcare professionals to set learning goals for children during consults, whereby the goals are visualised in a 'goal tree' structure categorised by different levels. Furthermore, it enables the healthcare professional to enter data of the child including personal data and preferences such as sports and hobbies. Next, education materials can be added manually. Print screens of this current state and the 'goal tree' can be found in appendix A.

The current state of MyPAL App (version: 17-02-16) on the tablet for children consists of three functionalities: playing a quiz, filling in a diary and seeing the progress of quiz. For the parents a monitor & inform tool will be developed, which is also referred to as PAL Inform. Furthermore, the PAL system is composed of a social robot (NAO) and its (mobile) avatar, which all connect to a common knowledge-base and reasoning mechanism. This provides the possibility for children to make use of PAL in different settings, for example in the hospital and diabetes camps. Furthermore, children can interact with the social robot at home and/or at school through the virtual avatar of the robot. This project will take in total four years and involves the research partners TNO (coordinator), DFKI (Deutsches Forschungszentrum Für Künstliche Intelligenz), FCSR (Fondazione Centro San Raffaele), Imperial College, Delft University of Technology, next to end-users (the hospitals Gelderse Vallei, Meander, the Diabetics Associations of Netherlands and Italy), and SME's (Mixel and Produxi) (PAL, 2016).

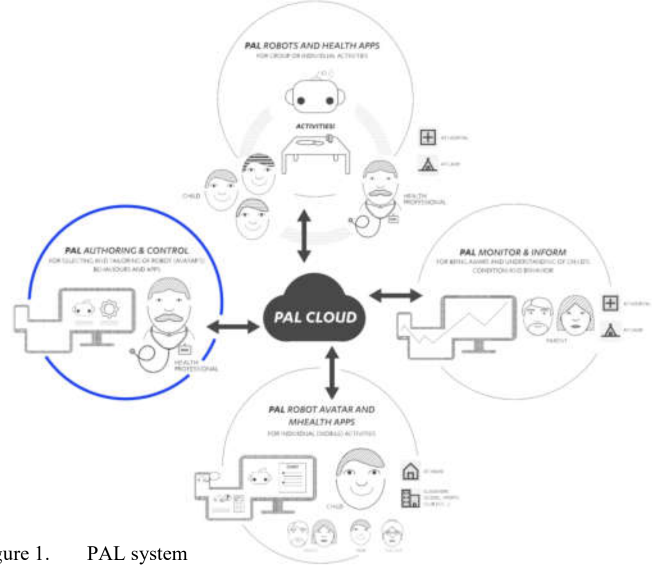


Figure 1. PAL system

1.3 Research question

This study aims to improve the current version of PAL Control through a needs assessment amongst healthcare professional, setting design specifications, prototyping and evaluating a clickable high-fidelity prototype of a healthcare management tool. This tool aims to support the healthcare professional in guiding children with diabetes self-management and authors the PAL robot and its avatar. The development is based on requirements collected during previous research in the PAL project (PAL, 2016). It will further contribute to fulfil the needs of children with T1DM, stimulate self-management, and enlarge their health and wellbeing through personal, pleasurable and social interaction under the guidance of healthcare professionals in PAL Control. The issue this study wished to address was: *'How can a healthcare management tool support healthcare professionals in guiding children with diabetes self-management involving a social actor (robot/avatar)?'.*

2. Background

2.1 Situated Cognitive Engineering

The design and evaluation of PAL Control follows the situated Cognitive Engineering (sCE) approach (Neerincx and Lindenberg, 2008), which aims to establish and test theories from the domain for which the application is developed. Furthermore, this approach provides quick, incremental, and iterative generate-and-test cycles. It has been previously applied in a wide variety of application domains and in multiple projects, such as a support system for human-robot team performance in complicated tasks (Mioch et al., 2014). Moreover, sCE maintains the sharing and reusing of design knowledge by a heterogeneous, multidisciplinary development group. Crucial is the generation, refinement, validation, maintenance, and reuse of consistent and brief design specifications. Such design specifications outline what the technology should do and the underlying design rationale, which is the 'why' and 'when'. Three main sections are distinguished: foundation, specification, and evaluation. As can be seen in figure 2, each of these sections has a small set of obligatory components that must be specified (Neerincx and Lindenberg, 2008).

The foundation section in the sCE methodology outlines the design rationale with regards to operational demands, relevant human factors knowledge, and envisioned technologies. The operational demands describe the current practice as it is. The human factors knowledge component in sCE describes available knowledge elicited from previous research about how to solve the problems that have been identified in the problem analysis. The component envisioned technology outlines the available possibilities of using existing technology and/or the need to elaborate new technology in order to achieve a system solution. Together, these three components describes the problem to be solved, the existing knowledge on ways to solve the problem and the technology needed to implement that solution (Neerincx and Lindenberg, 2008).

The section specification consists of design scenarios, use cases and requirements. It outlines the solution to the problem in the form of a system design that is based on the described relevant human factors knowledge and the envisioned technology. Design scenarios are short stories that gives a clear description of how the user will use the technology. Next, these scenarios are used to create more specific descriptions of step-by-step interactions between the technology and its users in the form of use cases. Thereafter, use cases are used to acquire functional requirements, which are specific functionalities the technology should give to its user (Neerincx and Lindenberg, 2008).

The last part of the sCE method is the evaluation, which aims to test and validate the system's design in order to improve the current design. It consists of the artefact, the evaluation method, and the evaluation results. The artefact is a prototype that integrates a given set of requirements, technological means and interaction design patterns. The evaluation method can be done in many different ways, such as an expert review. The evaluation results outline the results of the test. Due to the iterative and rapid research cycles, the evaluation does not necessarily integrate all requirements and use cases described in the system specification (Neerincx and Lindenberg, 2008).

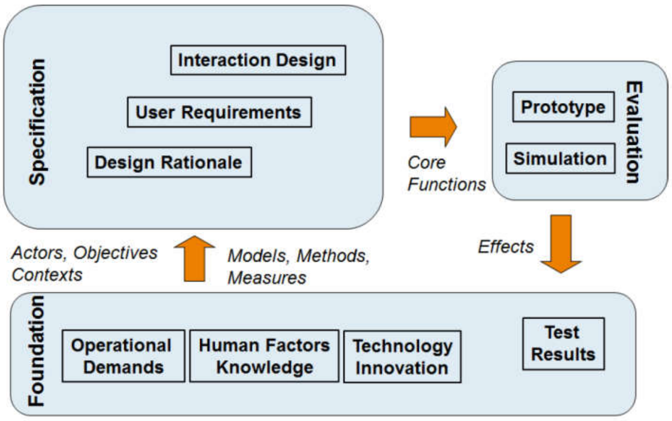


Figure 2. Situated Cognitive Engineering Method

2.2 Diabetes care

Optimal diabetes management begins with establishing the foundations of care. The healthcare professional preferably takes a holistic approach in providing care, considering all aspects of the patient's life circumstances. A team approach to diabetes management gives a thorough assessment and a development plan that describes the patient's values and circumstances. The components for the comprehensive medical evaluation that consist of aspects related to medical history, physical examination and laboratory evaluation, provides the healthcare team with information to optimally guide a patient with diabetes. Furthermore, information related to nutrition and psychosocial assessments should also be acquired (ADA, 2016).

The foundation of successful diabetes management consists of behavioral changes, an ongoing individual lifestyle, engagement of the patient, and assessment of the patient's level of understanding about the disease and level of preparedness for self-management. The core concepts of diabetes management are: diabetes self-management education (DSME), diabetes self-management support, counseling on stopping with smoking, medical nutrition therapy, education on physical activity, support on routine immunisations, and psychosocial care. Patients should also receive recommended preventive care services, such as immunizations. Furthermore, podiatric, ophthalmological, and dental referrals should also be provided. Healthcare providers should make sure that individuals with diabetes are screened for complications and comorbidities (ADA, 2016).

Patients with diabetes should receive medical care from a integrated team that may consist of physicians, nurse practitioners, physician assistants, dietitians, exercise specialists, nurses, pharmacists, dentists, podiatrists, and mental health professionals. The patient, family and healthcare professionals should work out the management plan. Different strategies and techniques should be used to enable patients to self-manage diabetes, which includes giving education on problem-solving skills for all aspects of diabetes management. Treatment goals and plans should be individualised, while considering patient preferences. In developing the plan, healthcare professionals should take the following aspects into account: the patient's age, physical activity, school/work schedule and conditions, social situation, eating patterns, cultural factors, diabetes complications, health priorities, other medical conditions, preferences for care and self-management, and life expectancy (ADA, 2016).

Current best practice of DSME is an approach based on skills that focuses on supporting those with diabetes to make informed self-management choices (Jensen et al., 2009; Charron-Prochownik et al., 2013). DSME has changed from a didactic approach, which is focused on giving information, to an empowering approach, which is focused on supporting those with diabetes to make informed self-management decisions (Charron-Prochownik et al., 2013). Diabetes care has moved to an approach that is more patient-centered and places the patient with diabetes and his or her family at the focus of the care model, working in collaboration with healthcare professionals.

2.3 Needs in children with diabetes

Children living with a chronic disease, such as diabetes, need additional support to enable them to achieve the main outcomes described by 'Every Child Matters', which is one of the most important policy initiative and development programs in relation to children and children's services of the last decade. These outcomes are: be healthy (enjoying good mental and physical health), stay safe (being protected from suffering), enjoy and achieve (getting the most out of life), make a positive contribution (being involved with the society and community) and achieve economic well-being (not being prevented from achieving prospects due to disadvantage. Nurses need to take these outcomes into account when working with children with diabetes (Owen, 2006).

The American Diabetes Association (ADA) (2016) announced that children have characteristics and needs that require different standards of care. However, the literature on self-management seldom makes a difference between self-management in children and in adults (Schilling et al., 2002). Research has shown that children with diabetes have a double greater prevalence of depression, and adolescents up to triple greater than youth who do not have diabetes (Grey et al., 2002). For patients diagnosed at a very young age with T1DM, having a strong support network is important for managing their diabetes effectively. It was found that often contact with certified healthcare professionals could improve the HbA1c value (glycated hemoglobin) and decrease hospitalisation rates (Howells et al., 2002; Svoren et al., 2003).

The goal of managing diabetes in childhood is to support the child in becoming an emotionally mature and physically healthy adult, free from difficulties associated with diabetes (Clarke, 2011). Children should be supported to manage their diabetes as part of their daily lives and most importantly, children must be involved in the decision-making of their management (NCC-WCH, 2004). A combination of specialist pediatric and diabetes care is needed for children with diabetes to be well-adapted and lead long and healthy lives in families which have come to accept the condition. Progressive accomplishment of self-management happens as developmental changes develop during childhood. It is crucial to be aware of the cognitive, psychosocial, and emotional skills needed to support the child with diabetes in accomplishing age-appropriate self-management. Appropriate expectations for the child with diabetes can be established through an understanding of normal growth and development as it relates to diabetes self-management tasks. Appropriate expectations and support in achieving self-management skills can support the child with diabetes move to self-management (Scott, 2013).

Self-monitoring of blood glucose (SMBG) has a critical role in the management of insulin-dependent diabetes mellitus. Patients have to test their blood glucose levels at least four times a day and set insulin doses for high and low blood glucose levels. Furthermore, frequent measurements, accurate record keeping of SMBG results is necessary for the patient and health care practitioner to evaluate diabetes control and make appropriate treatment adjustments on the long term. However, there is proof that adherence to both of these requirements is problematic for many patients (Gonder-Frederick et al., 1988).

Freeborn et al. (2013) conducted focus groups with 16 children in order to discover challenges and to understand experiences in the school setting of children and youth with T1DM from their own perspectives. They discovered that the three main challenges for children with T1DM are: self-care activities, low blood glucose and feeling lonely and/or different.

In a recent study it was concluded that living a good life with diabetes is demanding for the individual. However, experienced barriers can be relieved by aid from others in the personal sphere, and by professional support from diabetes care. Diabetes care was an important resource to develop the individual's ability and knowledge to manage diabetes, and to ease life with diabetes by giving support, guidance, medical treatment and technical devices tailored to individual needs (Engström et al., 2016). In the next paragraph technical tools used by healthcare professional in the guidance with diabetes self-management are discussed.

2.4 Diabetes management tools

Robots are becoming more inherent into everyday life. Time-critical domains, such as emergency response (Murphy, 2014) will include human-robot teams into missions, and robots will function as peer team members (Scholtz, 2003). One of the projects that provided valuable insights about how human-robot teams act in the healthcare domain is the ALIZ-E project, which started in 2010 to build the artificial intelligence (AI) for small social robots and to study how children would react to these robots (Aliz-e, 2014).

Another related project is REACTION (REACTION, 2014). Their goal is to develop an integrated ICT platform that facilitates improved long term diabetes self-management based on wearable, continuous blood glucose monitoring sensors and automated closed-loop delivery of insulin. This platform provides professional, integrated remote monitoring and therapy management to diabetes patients. Looking further at the functionalities within REACTION for healthcare professional, this platform can connect legacy healthcare systems, medical expert knowledge systems and provide closed-loop feedback in hospital environments to healthcare professional. Furthermore, it can aim at the outpatient regime and offer personalised feedback to healthcare professional.

A recent new monitoring system is the FreeStyle Libre Flash Glucose Monitoring System, which is a small sensor that automatically measures and continuously saves glucose readings day and night and is intended as a replacement for blood glucose meters (FreeStyle Libre, 2016). From the healthcare professional perspective it is useful that this monitoring system consists real-time glucose values, trend information and comprehensive reports, such as a summary of glucose data into percentiles throughout the day. Another monitoring system is COMMODITY12 (Commodity, 2015). It consists of ambient, wearable and portable devices. This project built a multi-layered multi-parametric infrastructure for continuous monitoring of diabetes type 1 and 2. The system used multi-parametric data to provide healthcare professional and patients, with clinical measurements for diabetes treatment. Therefore, it integrates software interoperation, state-of-the-art networks and artificial intelligence approaches in order to achieve a Personal Health System.

According to Catherine et. al (2012), who reviewed web-based diabetes guidance for patients and healthcare professionals, the quality of the support varied and showed a moderate impact on psychological and health outcomes. Another research showed that healthcare professionals have a positive perception towards web-based support for adolescents with T1DM (Nordfeldt, 2012).

3. Approach

Figure 3 provides an overview of the approach of this study which is based on the situated Cognitive Engineering (sCE) approach. In accordance with the sCE approach, the study was divided along three parts.

The first part is the foundation and consisted of a literature research and needs assessment (paragraph 2 and 3.1). Literature were researched about the diabetes care as usual and state of the art of technical tools used by healthcare providers in guiding children with diabetes self-management. The findings are described in the background section. A needs assessment of children was based on literature and the needs assessment of healthcare professionals were conducted through semi-structured interviews with four diabetes nurses.

In the second part the system redesign was specified (paragraph 3.2 and 3.3). Problem and design scenarios of how healthcare professionals interact with PAL Control were created based on the insights of the needs assessment, literature research and observations. Thereafter, the scenarios were visualised into storyboards, which were used to create more specific descriptions of step-by-step interactions in the form of use cases. Next, requirements were obtained from these use cases. The insights acquired from the 'specification' phase were used to formulate the core functionalities of a redesign of PAL Control.

In the last part of the sCE approach, the evaluation, the specified core functions were prototyped and evaluated with six healthcare professionals (paragraph 4). The perceived usability of the prototype was evaluated qualitatively, with think-aloud method and semi-structured interviews, and quantitatively with the System Usability Scale (SUS).

Figure 3. Overview approach

3.1 Needs assessment

A needs assessment of children is based on literature. The databases PubMed and Google Scholar were used and the keywords were: chronic diseases challenges, children' needs and diabetes self-management. Also, hospital documents related to children's diabetes self-management were analysed. A needs assessment of healthcare professionals was conducted face-to-face in a semi-structured interview. Four healthcare professionals (diabetes nurses) were interviewed with the aim of understanding the values and needs of healthcare professionals in guiding children with T1DM with the help of a healthcare management tool. Furthermore, based on observations a first pilot with PAL Control (version: 17-02-16) was analysed. Diabetes nurses were asked to fill in the basic profile page of the child and to set goals in the 'goaltree'structure in PAL Control. The observations consisted of screen recordings and videos, whereby the interaction between child and the diabetes nurse and parents was filmed while using PAL Control.

Participants were recruited from the two Dutch hospitals Gelderse Vallei and Meander. Eligibility criteria were participation in the experiments and experience of using PAL Control with a minumum of five children. Each interview was between 30 to 60 minutes and the topics for the semi-structured interviews were: functions, expectations, satisfaction, usability, setting goals and future improvements. The interviews were audio taped and transcribed for further analysis. The transcriptions were thematically summarised and analysed with codes, which gave an overview of the values and needs of the healthcare professionals. Figure 4 illustrates the most often mentioned problem scenario, which the nurses experienced when using the current version of PAL Control.

3.2 Scenarios and Storyboards

The values and needs (see table 1) elicited from the interviews, observations and literature are used as the foundation of the design specification, i.e., scenarios, storyboards, use cases and requirements. Based on the interviews and observations design scenarios could be created, which are described briefly in table 1. A more detailed design scenario (Fig. 5) illustrates how the new version of PAL Control can help the healthcare professional in selecting the right goals and the corresponding actions for the child, which resulted from the problem scenario described earlier in figure 4.

Furthermore, a visual representation of the scenarios was designed in storyboards in order to make the functionalities of PAL Control easy to understand for healthcare professional in the evaluation stage. The complete storyboards can be found in appendix B. The storyboards are based on the scenarios of how healthcare professional can use PAL Control in the situation before the first consult, during the consult and at the next consult.

3.3 Use Cases and Requirements

The scenarios and storyboards are used to create more specific descriptions of step-by-step interactions between PAL Control and healthcare professional in the form of use cases. Figure 7 illustrates the use case of the interaction between the healthcare professional and the child in setting goals and selecting actions after the child has played the assessment with the robot in the hospital. The other use cases can be found in appendix C. Thereafter, use cases are used to obtain requirements (see table 1).

Problem scenario

'Sarah (nurse) has difficulties in selecting the appropriate goals for John (child).'

'Sarah experiences difficulties in deciding which goals she should select in the goaltree of PAL Control for John (age 9), while talking to John at the same time. Furthermore, she finds the description of the goals not clear and some goals are missing. She tries to ask John several open questions in order to understand the needs of John, but he only answers in short sentences. Sarah tries to find a balance in making it a fun conversation for John, while testing his knowledge. Due to time limitation of the consult she can only ask a few questions. Therefore, she is still not certain if she has chosen the right goals for John to work on.'

Figure 4. Problem Scenario

Design scenario

'John (child) plays assessment with robot at the hospital, which can support Sarah (nurse) to decide which goals John should be working on.'

'Sarah explains to John that he is going to play a quiz on the tablet with the robot in the hospital. After playing the assessment John will see his scores of the topics (icons) and the robot would ask him if there are topics that he would like to work on. John can select the icons. If not, he can go to Sarah to discuss it. Sarah can discuss together with John by looking at the scores which goals he should be working on. After selecting the icons a preset of actions are presented and Sarah can personalise the actions or add new ones by discussing it with John. When both Sarah and John agree on the goals and actions, she can click on save and the goals with the actions will appear in MyPal.'

Figure 5. Design Scenario

4. Evaluation

4.1 Prototype

The scenarios, storyboards, use cases and requirements derived from the interviews with healthcare professional, observations and literature provided input and focus for the prototype development. The selected requirements (see table 1) were being prototyped as core functionalities in PAL Control.

It was decided to create a prototype consisting of four main components: PAL Control on desktop; assessment with PAL Robot at hospital; assessment with PAL avatar at home; and MyPAL environment with the new functions 'information' and redesign of 'goals'. The interface was designed to allow healthcare professional to develop a rapid understanding of the value of new functionalities of PAL Control in their practices. Furthermore, in order to get a complete impression of functionalities that the healthcare professional might find useful, it was decided to prototype a complete PAL Control including redesigning features such as select child, sign out/sign in page, language option and a redesign of the profile page where healthcare professional can enter basic information of the child. These redesign requirements were elicited from the observations and requirements of previous research in PAL project. The goal of evaluating a complete PAL Control was to facilitate a more consistent and realistic system for the healthcare professional.

A paper prototype was created first to rapid sketch out the main ideas. Furthermore, by creating a paper prototype that implements the scenario an early user test could be conducted. Next, a mockup was created with the tool Balsamiq Mockups[1], which is a rapid wireframing tool that produces the experience of sketching on a whiteboard using a computer. The mockup was useful in discussing the core ideas and helpful in the decision process of which functionalities to evaluate. Appendix D shows the lo-fi prototype and the elements they contain. After the main concepts were decided, a high fidelity prototype was created.

[1] https://balsamiq.com/

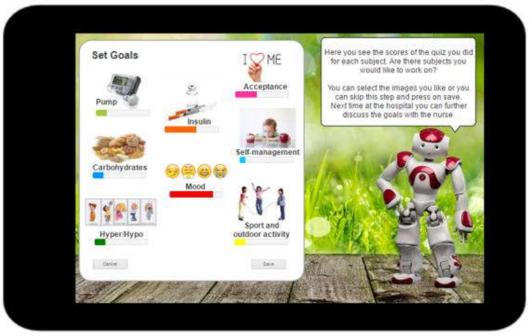


Figure 6. Prototype print screen: set goals with PAL Avatar

First, static screens were designed in photo editing software Adobe Photoshop[2] and graphic design software Adobe Illustrator[3]. Next, Axure RP Pro[4], which is a wireframing and rapid prototyping tool aimed at web and desktop applications, was used to make the static screens clickable. The complete screens of PAL Control prototype can be found in appendix E. The prototype consisted a variety of functionalities, which were categorised into 11 core functions (see table 1 and Fig. 6).

An important design choice within this prototype is by facilitating 'Nurse View', which is where healthcare professional can sign in/out within the MyPAL environment of the child in order to start the diabetes assessment with PAL robot, set goals, select actions and edit progress. The aim is to improve the interaction between the child and healthcare professional and to involve the children more in the decision process of goals during the consult by creating this authoring function in MyPAL on the tablet. This can enable the child to understand how the goals are set and how the activities appear in MyPAL. Furthermore, it can also invite the child to touch the screen together with healthcare professional to set goals and select actions. Another design choice was to minimise the volume of text in the screens through the use of graphical representation in the form of pictograms in order to stimulate the interaction between the child and healthcare professional.

This prototype is evaluated to assess to which extent the core functions of the redesigned PAL Control, which aims at fulfilling the needs as described in table 1, can support the healthcare professional in guiding children with diabetes self-management. Furthermore, the prototype evaluation aims to determine whether there are specific aspects more useful than others and what needs to be improved.

4.2 Evaluation Method

Participants were recruited between April and June 2016. In total six healthcare professionals from the hospitals Gelderse Vallei (3) and Meander (3) participated in the evaluation. These healthcare professionals consisted of five diabetes nurses and one doctor (a pediatrician and who is also a medical director), whereby four of them were also interviewed earlier to elicit their needs (paragraph 3.1). All diabetes nurses that participated had experience with using PAL Control with a minimum of five children during the first PAL-pilot. Additionally, we decided to also incorporate a doctor's perspective by interviewing one doctor who worked with two of the diabetes nurses.

The six evaluations were conducted face-to-face at the hospitals Gelderse Vallei or Meander. Prior to each evaluation, an introduction was given to explain the prototype and the goals of the evaluation. Furthermore, participants were asked for permission to record (video & audio) the evaluation. All interviewees were asked to speak freely and were assured that the recordings remained confidential. The evaluation took approximately an hour and consisted of: 1) showing the storyboards that represent the context of PAL Control being used before, during and after the appointment; 2) applying the 'think-aloud-method, whereby the healthcare professionals were asked to use the clickable prototype to test the different functionalities, while continuously thinking out loud; 3) conducting a semi-structured interview in order to understand whether they are interested in these functionalities and what the improvement points were. Interview topics were: need fulfillment, functionalities, expectations, satisfaction, usability and future improvements; 4) completing the System Usability Scale (SUS), which provides a reliable tool for measuring the usability. It consists of a 10 item questionnaire, whereby a response is given on a scale of 1 to 5, according to the user's degree of satisfaction (Borsci et al., 2009).

Data from the semi-structured interviews and think-aloud-method were transcribed during the evaluation, which supported early familiarisation with the data. Certain parts were listened again from the recorded audio and video file to ensure that a correct transcription has been noted. Transcript analysis was an iterative process using thematic analysis, which involved identifying themes and codes across all transcripts and SUS data. Codes were meaningful groups of data that described the core of the data. Atlas.ti[5] was used to facilitate the organisation of codes and themes.

- **Goal:** Child and healthcare professional set goals and select actions together in MyPAL
- **Actor:** PAL robot, MyPAL, Sarah (healthcare professional) and John (Child)
- **Pre Conditions:** The robot, a tablet with the app MyPAL, John and Sarah are in the hospital.
- **Post Conditions:** With the help of the PAL robot, Sarah could easily set the appropriate goals for John and select the corresponding actions. John understood what the goals are and felt fully involved in the decision process of selecting his learning goals.
- **Action Sequences:**

1. Sarah introduces the robot to John and explains that he will play a quiz about diabetes management with the robot.
2. Sarah asks John if she can use his tablet to sign in 'Nurse View' within MyPAL to start the assessment. John gives her the tablet.
3. Sarah star the quiz and goes to another room, while John does the assessment (e.g. 3 questions for each topic) with the robot.
4. After answering the questions, John sees his scores. The scores are presented in a progress bar, an image of the topic and the level he has reached of that specific topic .
5. Robot asks if John has any subjects he would like to learn more about and invites John to select a few images, whereby the progress bar is visualised below the images.
6. If there is no response from John, the robot will suggest John to show the goals to Sarah and discuss this together.
7. John selects a few images and press on continue.
8. Robot asks John to show these goals to Sarah and to ask her opinion of the selected goals.
9. John and Sarah sit together and Sarah explains that with the goals the child can improve on his diabetes management and that these goals can be achieved by playing PAL activities with the PAL Avatar.
10. Sarah agrees with the selected goals after looking at the progress bar that is visualised below the goals.
11. After pressing on 'save', the next screen shows a preset list of actions (e.g. watch a video, upload diabetes values or play the quiz), whereby the robot has already selected a few activites based on the scores of the assessment from John.
12. Sarah asks John if he agrees with these actions and if he has any actions that he would like to do that are not in this list yet.
13. John and Sarah discuss together about the actions and Sarah can deselect/edit the actions or add a new action if it is not in the list yet.
14. Sarah compliments the child for his goals and indicates that she is going to sign out 'Nurse View' now and that John can start doing the selected activities with the PAL avatar when he gets home.

Requirements:

- PAL shall provide a visual overview of icons representing topics related to diabetes management.
- PAL shall invite the child to first p ick icons they want to talk about, followed by suggesting the child to discuss about the selected icons with the healthcare professional.
- PAL shall provide an overview of different activities that are related to the topic and suggest some activities based on the child preferences, whereby the healthcare professional can add and

Figure 7. Use case: Set goals and selecting actions

[2] http://www.adobe.com/products/photoshop.html

[3] http://www.adobe.com/products/illustrator.html

[4] https://axure.com/

[5] http://atlasti.com/qualitative-analysis-software/

Values & Needs	Scenarios	Requirements	Core Functions
Access - The need of having access into MyPAL in order to understand the activities that the child is doing in MyPAL.	Before the consult the healthcare professional can already start experiencing the functionalities in MyPAl in order to get a **better understanding** when the child talks about certain activities **in MyPAL**.	PAL Control shall provide a functionality whereby the healthcare professional can **see how MyPAL looks** like for the child anytime.	MyPAL View
Information - The need of getting insights about the data that the child has typed in MyPAL.	The healthcare professional can already start looking at the **statistics**, which are elicited from the information that is **typed by the child** in MyPAL. This will help the healthcare professional to better understand the needs of the child during the consult.	PAL Control shall provide an overview of child-user's **progression on PAL-objectives**, **emotion, daily activities and system usage**, whereby important/critical situations can be filtered.	MyPAL Statistics
Information - The need of getting insights of the knowledge and concerns of parents.	Healthcare professional can already start reading the **concerns** that parents has sent before the consult or during the consult and use these concerns as **topics to talk about with parents**. Furthermore, parents can play an **educational quiz** to test their knowledge about diabetes, whereby the healthcare professional can see the results as well to better fulfil the needs of the parents.	PAL Control shall provide an **overview of concerns that the parents** have send. PAL Control shall provide an **educational quiz for parents** and an overview of the results of the quiz for the healthcare professional.	PAL Inform
Information - The need of having a clear overview of the progress of each goal of the child.	Healthcare professional can see in the goal tree on **which level the child** is. When the child has fulfilled the actions the goal will switch from orange to green. Furthermore, the healthcare professional can also manually turn goals green if preferred.	PAL Control shall provide an **overview of goals, a clear description, ability to add more goals and the progression** of child's performances.	Goals
Education - The need of selecting (digital) educational material for children.	At the clinic before/during/after the consult, healthcare professional can contribute to child's autonomy in managing diabetes by **selecting appropriate educational materials** of the database or **add new digital materials** (e.g. scanning a flyer) in the database, whereby this source can be shared with other healthcare professional as well. The selected material for the child will be sent to MyPAL app of the child.	PAL Control shall provide **a portfolio of educational materials (for children)** from which the healthcare professional can select/add, whereby **PAL Actor will motivate the child** to read these materials in MyPAL.	Education [Child]
Education - The need of selecting (digital) educational material for parents.	At the clinic before/during/after the consult, the healthcare professional can fulfil the needs of **parents** by **selecting educational materials** of the database or add new digital material (e.g. scanning a flyer) in the database.	PAL Control shall provide a **portfolio of educational materials (for parents)** from which the healthcare professional can select/add and send as an attachment (under 'report').	Education [Parent]
Information - The need of having a guideline tool of topics where healthcare professional can talk about during the consult with parents/child.	Before consultation healthcare professional can already **type a checklist** in PAL Control and add notes that can be used as a **guideline of topics to talk about during the consult** with parents. Furthermore, an overview of selected goals and actions of the child can be **sent to the parents** and be used as a guideline to talk about in the next consult.	PAL Control shall provide an option to **send a summary of notes** to the parents, an overview of the **selected goals and the corresponding actions** in order to give the parents a better understanding of what their child is learning and the **educational materials** as an attachment (if any material was selected in 'education').	Report
Development - The need of a fun assessment in order to test the diabetes knowledge of the child in order to select the appropriate goals for him.	The healthcare professional can decide to let the child do the **assessment at the hospital with the robot** or healthcare professional can activate the assessment in Pal Control in order to let the child play the **assessment at home with the robot-avatar** in their own time.	**PAL Robot/ PAL avatar shall motivate the children in finishing the assessment** and praise children in ways that acknowledge their efforts.	Assessment in hospital/home
Equality - The need of making children more involved of their own learning goals and actions. ***Development*** - The need of easily selecting goals, whereby the child has a basic understanding what the goals represent.	During consultation, child and healthcare professional are at the same level. healthcare professional performs **shared decision-making about the learning goals** with child. The **child understands** what the goals are by looking at the images that he can select.	PAL shall provide a **visual overview of icons representing topics** related to diabetes management. PAL shall invite the **child to first pick icons they want to talk about**, followed by suggesting the child to discuss about the selected icons with the healthcare professional.	Set Goals
Equality - The need of having a possibility to add actions by the healthcare professional.	Healthcare professional performs **shared decision-making about the actions** that helps in achieving the learning goals with child and if the child has any suggestions of actions he would like to do, the healthcare professional can add these together with the child to the list.	PAL shall **provide an overview of different activities** that are related to the topic and suggest some activities based on the child preferences, whereby the healthcare professional can add and personalise the actions if preferred.	Select Actions
Achievement - The need of having a clear overview of the child's progression in achieving his goals and the possibility to edit it for the healthcare professional. ***Information*** - The need of having a clear overview of the progression of achieving goals for the child in MyPAL.	At the beginning of the next consult the healthcare professional can have a look at the progress together with the child to see which goals the child has achieved and which actions he found difficult to fulfil. The healthcare professional can give **motivational feedback** and together with the child they can decide to add more actions or move on to achieve another goal. The healthcare professional can decide based on the progress of the child to enable the child to do another assessment with the robot at the hospital or robot-avatar at home.	PAL Control shall provide an **overview of child-user's progression** on goals and achievements **for the healthcare professional**. PAL Control shall provide an **overview of child-user's progression** on goals and achievements in a way **that motivate the child** to continue achieving the goals.	Progress for healthcare professional & Child

Table 1. Overview values, needs, scenarios, requirements and core functions

4.3 Evaluation Results

4.3.1 Usage of Tablet

The participants appreciated that the shared decision making process of goals and actions is offered on a tablet instead of on a desktop. The healthcare professional mentioned that the tablet created a unconstrained atmosphere and is more inviting for the child to participate actively in setting goals/selecting actions. One diabetes nurse shared her observation of the previous experiment that she noticed that some children found it easier to share their emotions to the healthcare professional by looking at the tablet instead of looking at the healthcare professional directly. Furthermore, the nurse said that she also found it more comfortable to look at the tablet while talking to the child.

4.3.2 Role of PAL robot/avatar

Overall, the participants were positive about the assessment functionality with the robot/avatar. The healthcare professional mentioned that it is appreciated that the robot is involved in setting goals and selecting actions at the beginning. They believed that the presence of PAL robot or its avatar motivates the child. Regarding the preference between PAL robot or its avatar, the healthcare professional indicated that a combination of both would give the best outcome. Their expectation is that the child would like the robot the most. One healthcare professional mentioned: *'You are making the consult more fun for the child'*. However, you do not want the child to travel to the hospital every time and therefore, doing the assessment with PAL avatar was perceived as a good solution to give the child the option to do it at home in their own time. She also liked the idea that you can see what the child has achieved or has difficulties upfront in order to have a more meaningful consult.

4.3.3 Usability

Overall, the usability of the prototype was perceived positively. The healthcare professional appreciated the colour use and pictograms. Furthermore, the healthcare professional mentioned that it is nice that a picture of the child could be added as a profile picture in the system. A healthcare professional said: *'My first impression is that it is nice to use and easy to understand. After using it in practice, it will become clear if this impression stays the same.'* The mean SUS score attributed by the five diabetes nurses also confirms this (see table 2). The score was above average, at 79. This score can be interpreted as a grade of B in a range of A to F, whereby A is the highest score achievable (Borsci et al., 2009). Explanation of the calculation of the user scores and SUS scores can be found in appendix F.

Main improvements from a usability perspective are the way how the data is visualised. One healthcare professional indicated that she did not like graphics and that she preferred a list or a circle diagram. Furthermore, a healthcare professional mentioned that it also should be evaluated with children to see how they perceive graphics as a way to see their progress or that they perhaps prefer another way of information visualisation.

4.3.4 Functionalities

Overall, the functionalities were perceived positively by the participants. One healthcare professional mentioned: *'It is a nice way to make the outpatient visit more child friendly and this approach fits well in these modern times.'* Another healthcare professional said that there are too many functionalities, but she also mentioned that the functionalities sounded useful and perhaps some of them could be moved to other components of PAL system (e.g. PAL Inform/MyPAL). Another participant mentioned that the system was coherent and that it has a logical structure. A nurse said: *'It seems that there has been a lot of thought put into all these different components. There are different possibilities to use the system, which makes it attractive'*. It was for all participants difficult to differentiate which functionalities they preferred. A diabetes nurse said: *'The way I see this system, without experience using it in practice, is that all functionalities are needed and complete each other. My opinion might change when I have worked a few weeks with this system.'*

Profile page – A redesign of the current profile page was created and one healthcare professional mentioned that the profile page is not necessary, because they already have that in another system. However, other healthcare professionals indicated that it was useful to look into it when they are in PAL Control, because otherwise you have to open two systems at the same time. Another healthcare professional suggested to add more information in this profile page about for example the information of the school the child is going, which can be interesting information for the healthcare professional when the child goes to secondary school.

Report – In the 'report function', a healthcare professional suggested for an option that the things you have written as notes/checklist can sometimes already be sent to parents to give them an outline of what they can expect during the next meeting. Another healthcare professional mentioned that it would be valuable to synchronise certain information/functionalities of EMR within PAL Control to avoid manually filling in things for the second time. One healthcare professional mentioned that only the sub functionality of sending an overview of goals & actions of the child after the consult is enough for her. Another participant liked the idea to send a summary of the consult to the parents, such as important notes/advise. She stated: *'Parents can read the summary again whenever they want and they will not lose the information.'*

		HCP 1		HCP 2		HCP 3		HCP 4		HCP 5		Average	
		User Score	SUS Score	User Score	SUS Score	User Score	SUS Score	User Score	SUS Score	User Score	SUS Score	User Score	SUS Score
1	I think that I would like to use this system frequently.	5	4	4	3	5	4	3	2	5	4	4,4	3,4
2	I found the system unnecessarily complex.	2	3	1	4	1	4	1	4	1	4	1,2	3,8
3	I thought the system was easy to use.	2	1	4	3	4	3	4	3	4	3	3,6	2,6
4	I think that I would need the support of a technical person to be able to use this system.	1	4	2	3	3	2	3	2	3	2	2,4	2,6
5	I found the various functions in this system were well integrated.	4	3	5	4	4	3	3	2	5	4	4,2	3,2
6	I thought there was too much inconsistency in this system.	1	4	1	4	2	3	2	3	1	4	1,4	3,6
7	I would imagine that most people would learn to use this system very quickly.	5	4	4	3	4	3	3	2	5	4	4,2	3,2
8	I found the system very cumbersome to use.	4	1	1	4	1	4	3	2	1	4	3,25	3
9	I felt very confident using the system.	4	3	3	2	4	3	3	2	5	4	3,8	2,8
10	I needed to learn a lot of things before I could get going with this system.	2	3	2	3	1	4	2	3	1	4	1,6	3,4
	SUS Score		**75**		**82.5**		**82.5**		**62.5**		**92.5**		**79**

Table 2. SUS Score

Education Child and Parent – The idea of a database with educational material, which can be send to MyPAL for the child was perceived as positive. Additionally, this same functionality for parents was also perceived as useful. One healthcare professional said: *'It's very nice to be able to supply information in a digital way, because papers can be lost quite easily.'* Furthermore, healthcare professional liked the idea that they can add more material if they want and that this database is shared with other healthcare professionals, but it should remain clear who (hospital Gelderse Vallei or Meander) has uploaded the educational material. A healthcare professional mentioned an important challenge within this function is to make sure that the database overview stays clear and that it contains qualitative good material. Furthermore, it should be easy to find material related to certain topics in the database. Another participant mentioned a critical limitation was that practical aspects, such as how to inject insulin could not be so easily learned by only watching a video.

Assessment – The assessment starts with two questions about the age and the device (pump or pen) that the child is using. Some nurses mentioned that these two questions might be too difficult for a young child and perhaps these two steps could be done together with the healthcare professional. Another healthcare professional mentioned that these two questions are unnecessary due to the fact that this information is already in the system. Overall, healthcare professionals liked the idea of the child doing an assessment with the robot or its avatar, because the score overview can help healthcare professionals to determine more easily which topics the child needs to work on.

MyPAL view – The functionality of MyPAL view was perceived as useful by all participants. One healthcare professional said: *'I find it useful to have access to see how the activities in MyPAL look like for the child, so I have a better understanding when the child talks about certain activities in MyPAL during the consult.'*

PAL Inform – Overall, healthcare professionals mentioned that PAL Inform was useful. They indicated that these data help them to better understand the needs of the children and their parents. Within PAL inform there is a functionality whereby parents can already send their concerns about their child to healthcare professional when something comes into their mind. A nurse said: *'Often parents have already forgotten their questions during the consult, especially after the meeting with the doctor first (parents talk to doctor first before talking with the diabetes nurse). Now they can easily ask at home already.'* The idea of parents doing a diabetes quiz was perceived diversely. One healthcare professional mentioned that she does not expect that parent have the need to play the game, while another healthcare provider mentioned that it is very useful to educate parents as well and that by playing a quiz can be a good way to find out the gaps in parents' diabetes knowledge.

PAL Statistics –Interviews with the doctor and nurses indicated that there are main differences in their needs. Most of the nurses suggested that the data about insulin and glucose was not necessary in PAL Statistics, while the doctor would like to see this data in the system. The nurses mentioned that the data related to insulin, glucose and carbohydrates could be found in other existing systems already. A nurse indicated that data of food diary is more valuable for the dietician. However, the nurse mentioned that the function can be interesting as an activity for the child to keep a food diary for a certain amount of time. Then it might be interesting to see what the child has filled in the food diary. Both the nurses and doctor do not want the child to fill in the data manually. Furthermore, another nurse mentioned that it would be useful to include the standard settings of aspects that they have agreed on, such as the base stand and the carbohydrates ratio.

Goals & Progress (PAL Control) – Improving the visualisation of the 'goal tree' structure was not within the scope of this study. The additional visibility of an overview of actions and progression when the goal was selected in the 'goal tree' was perceived as very useful. In the previous version of PAL Control, healthcare professionals could not see which goals the child has been working on. In this prototype their need of having a clear overview of the progress of each goal of the child has been fulfilled. And also the option that the healthcare professional can add more goals based on their preferences was appreciated.

Goals & Actions (in MyPAL through 'nurse view') – Overall, the process of set goals and select actions was perceived as very positive by the participants. One healthcare professional stated: *'It is for the child a fun way to set goals and think about what he/she would like to learn. They really like the robot.'* Furthermore, healthcare professional especially appreciated the option to add their own actions if it is not in the list yet. One healthcare professional said: *'For example, when you know the child around 13 or 14 years smokes, you can add the action searching info what smoking does to your health for example .'*

New functionality – When asked if the participant would add more functionalities in this system, a participant suggested the functionality for children to ask questions towards healthcare professional when something comes into their mind, which they would like to talk about in the next appointment. A similar idea as in PAL inform, whereby parents can send their concerns to healthcare professional before the consult already.

4.3.5 Doctor/management perspective

The doctor did not have any prior experience with PAL Control and therefore a demonstration of the prototype was showed to him directly without conducting the think-aloud method and he did not fill in the SUS. From his experiences working with healthcare professional and in the management of the hospital he provided useful insights during the semi-structured interview. He mentioned that the hospital is doing a pilot project with a monitoring system called Gluconline, which is a virtual coach app for diabetes patients. Furthermore, they are experimenting with wearables. Therefore, he suggested that monitoring data should be uploaded through automatic sensors to make it easier for children. However, he mentioned that an important aspect whether new technologies are being used is the aspect whether it is reimbursed by health insurances. Moreover, he raised the question how data of EMR (Electronic Medical Record) and PAL data can be synchronised. Technically it might not be achievable according to him. Furthermore, he stated that if the information (e.g. medical items) in the 'profile' page has to be updated in two separate systems that it could lead to dangerous errors in case of serious conditions, such as peanut allergy. Last, he asked what the 'rewards' for children to adhere to the system are and he suggested some gamification aspects.

4.3.6 General insights

The lack of an automatic sensor that can upload the glucose and insulin values for children automatically was perceived as the biggest bottleneck according to all participants in this study. Furthermore, healthcare professionals would like to know if it is possible to synchronise EMR and PAL data and how this healthcare management tool could integrate in their current workflow and existing systems. In addition, it was found that if information cannot be synchronised, a safe and efficient way of information transfer needs to be reassured towards the healthcare professionals.

In general the healthcare professionals reacted positive towards the redesign of PAL Control. They perceived the role of the robot and its avatar very useful in motivating the child. Furthermore, they stated that the tablet could create an unconstrained atmosphere compared with the desktop and that it could influence the shared decision making process with child of setting goals and selecting actions positively. They appreciated the functionality of the child doing an assessment with the robot/avatar in order to help the healthcare professional determine which goals the child should be working on. In addition, they appreciated the colour use and pictograms. However, the progression page was perceived less positively than expected. The participants mentioned that they would like to see less graphs for the child and more gamification aspects.

Although participants were positive, they found it very difficult during the evaluation to rank their answer from 1 till 5 (with 1 strongly agree and 5 strongly disagree) to what extent their need has been fulfilled as described in table 1. They all mentioned that they needed more time to really use the system in practice in order to understand these functionalities and to determine their answer. Furthermore, it remains unclear how many functionalities the healthcare professionals would like to have in PAL Control.

5. Discussion and conclusion

The research question this study sought to answer is: *'How can a healthcare management tool support healthcare professionals in guiding children with diabetes self-management involving a social actor (robot/avatar)?'*. To answer this question, a prototype of a healthcare management tool was developed and evaluated with end users (diabetes nurses) and an important stakeholder (diabetes doctor), following the sCE approach. The healthcare professional needs and system requirements, as described in table 1, and usability were the main points of evaluation. Overall, this prototype of a redesigned PAL Control was well perceived by the healthcare professionals and the findings demonstrated that there is potential for the proposed functionalities: an assessment with a robot or its avatar, set goals and selecting actions, create reports, an educational database for children and parents, monitoring solutions (MyPAL view, PAL Inform and PAL Statistics) and a progression page.

All healthcare professionals expressed that they see an added value in the robot/avatar and they especially liked the idea of the child playing an assessment with the robot, which can help them determine the gap in diabetes knowledge of children. Furthermore, healthcare professionals mentioned that this system has provided them support in making the consult with children and parents more meaningful due to the fact that they can understand their needs better on forehand through the progress page and the results of the assessment with the robot or its avatar. Also, healthcare professionals in this evaluation appreciated that the functionalities of setting goals and selecting actions with the child are run on the tablet. Next, they found the progress page useful to establish where the child has worked on and which topics need more attention, but they would like the progression page for the child to be more tailored to the child's age. Therefore, graphs are perhaps not suitable. However, they did like the usage of colours and pictograms.

During the evaluation, it was found that there were important differences in needs among the nurses and between the nurse and the doctor with regards to the number of functionalities and the perceived usefulness of certain data in the system. Many important individual differences need to be taken into consideration. Moreover, we need to discuss more about the role of PAL, is it an addition or will it replace the existing systems that healthcare professionals already use and how does this influence the current workflow of healthcare professionals? Will PAL Control be used after or before the normal consult or replacing the consult?

An ethical aspect within this prototype is how much information you can show the healthcare professional without invading the privacy of the child. Perhaps the child wants the healthcare professional to know what he/she has filled in. In the functionality 'PAL statistics' the texts related to emotions can be filtered out for the healthcare professional to read in order to get a better understanding how the child was feeling at that moment. Furthermore, the function 'MyPAL view' allows the healthcare professional to see the interface of MyPAL, which contains data filled in by the child. Although these two functionalities were perceived as useful by the healthcare professional, it should be taken into consideration that the child should not get the feeling of 'big brother is watching'.

The evaluation results together with the data using the 'thinking aloud method' and SUS score were used to iterate specifications of the functionalities and provided guidance in the improvement of the prototype. These improvements are:

Recommendation system – The functionality of the database with educational material was perceived as useful, but a critical concern was that this database should be kept organised in a clear and user friendly overview. A possible redesign of this functionality is to add a recommendation system, whereby healthcare professionals and children can give a rating of the material after they have seen it. Healthcare professionals could also add their keywords when they upload their own educational material. This could make the process of finding suitable educational information for the child more easily.

Recording messages – Based on the suggestion of one healthcare professional about the new functionality of giving children the option to send questions to healthcare professionals, a possible solution is that the child can record a message to the robot. For example, the child has played the quiz and does not understand why the answer he picked was incorrect. He would like to ask the healthcare professional next time about it. In order to not forget this question, the child can already send his question to healthcare professional through the robot, e.g. by clicking a button with 'message to healthcare professional' the voice recorder starts. In other words, the robot functions as a messenger from the child to the healthcare professional. During the consult healthcare professional and child can listen to the recorded questions together and discuss them, followed by optionally adjusting the goals and actions.

Finally, healthcare professionals stated that they would like this prototype to be built in a real system in order to test for a longer period of time and during multiple pilot consults in a setting with children, parents and the robot. Therefore, evaluations for a longer period of time is needed in order to validate if the needs are completely fulfilled. Nonetheless, useful suggestions were found during the evaluation of the prototype and provided important pointers in besides the development of PAL Control, also in PAL Inform and MyPAL. The sCE approach supports the credibility of the findings and ongoing reflective practice enhances the methodological thoroughness. End-users got an opportunity to express their needs in an early stage, which provided a strong foundation for the development of the prototype. Based on these preliminary results, we expect that this healthcare management tool can support the healthcare professionals in the assessment of children's T1DM knowledge, goals and development. In addition, the evaluation results of this healthcare management tool prototype are also relevant for healthcare professionals guiding children with other illnesses requiring self-management such as asthma or issues that negatively influence children's well-being such as bullying.

In summary, our results suggest that a combination of an assessment with robot or its avatar, setting goals, selecting actions and the progression page, is a suitable and effective approach to healthcare professionals in guiding children with diabetes self-management.

5.1 Limitations

Mostly, qualitative data were collected from six individuals, which makes the findings difficult to be generalised to a larger population. However, the qualitative data collected in this study provided useful and detailed information. Also, due to a limitation in time during the evaluation the thinking aloud method was sometimes partly replaced with a demonstration followed by a discussion and a semi-structured interview. Nonetheless, a range of confirming insights seemed to be captured.

5.2 Future work

Although the study found evidence of preliminary positive effects of the new functionalities in the prototype of PAL Control, further research by developing a real system based on the core functions of this prototype that allow healthcare professional to experience working with the tool for a longer period of time in a real pilot setting, are necessary to determine the effects. This will be an important step in the process of establishing a set of guidelines towards creating a useful healthcare management tool for healthcare professional. Also, needs of parents and children with regards to the role of healthcare professional should be elicited through further research. Furthermore, the content of the assessment (quiz, goals and actions) and the 'goal tree' structure should be established together with diabetes experts and reviewed by healthcare professional. Next, this project PAL could elaborate with new technologies, such as wearables that can registrate values automatically and avoid that children have to complete it manually. More research is needed to investigate possible options that allow this healthcare management tool to integrate in the existing systems (e.g. electronic health record and other monitoring programs that the healthcare professional are already using).

Most importantly, the concepts as proposed in this study aims to support children under guidance of a healthcare professional and it should therefore be evaluated in a setting with children to understand how healthcare professionals interact with children, while using PAL Control. Children should be motivated and engaged in using and maintaining the application for it to be a success. Therefore, another future research aspect is how can the healthcare professional/robot reward the child for achieving the goals in order to continuously motivate the child. Finally, more research is needed to study how the PAL avatar should interact with the child in a meaningful way in order to motivate children in achieving their personal goals as set together with the healthcare professional.

5.3 Conclusion

This study shows how a healthcare management tool, PAL Control, can help healthcare professionals guiding children with diabetes. This study focuses on the extent to which this tool fulfills the needs of healthcare professionals and the perceived usability. The results showed an overall positive increase in perceived usability compared with the previous PAL Control version. Healthcare professionals appreciated the usage of colours and pictograms. Overall, the evaluation outcomes are positive and represent a good basis for further development. Our results suggest that a combination of an assessment with a robot or its avatar, setting goals, selecting actions and the progression page, is a suitable and effective approach to healthcare professionals in guiding children with diabetes self-management. Potential improvements in the prototype of a redesign of PAL Control were identified.

6. Acknowledgments

I would like to thank my thesis supervisor Tobias Kuhn of VU University Amsterdam and my internship supervisors Olivier Blanson Henkemans and Sylvia van der Pal of TNO for their insightful comments. Furthermore, I would like to thank the healthcare professionals for their participation in this study.

7. References

Alize-e.org. 2014. *ALIZ-E European project.* Accessed on: May 7, 2016. http://www.aliz-e.org/

American Diabetes Association (ADA). "Standards of medical care for patients with diabetes mellitus." *Diabetes care* 39, no. 1:S23-S35 (2016).

Borsci, Simone, Stefano Federici, and Marco Lauriola. "On the dimensionality of the System Usability Scale: a test of alternative measurement models." *Cognitive processing* 10, no. 3 (2009): 193-197.

Catherine, H. Yu, Robinder Bahniwal, Andreas Laupacis, Eman Leung, Michael S. Orr, and Sharon E. Straus. "Systematic review and evaluation of web -accessible tools for management of diabetes and related cardiovascular risk factors by patients and healthcare providers." *Journal of the American Medical Informatics Association* 19, no. 4 (2012): 514-522.

Charron-Prochownik, Denise, Susan M. Sereika, Dorothy Becker, Neil H. White, Patricia Schmitt, A. Blair Powell, Ana Maria Diaz et al. "Long-term effects of the booster-enhanced READY-Girls preconception counseling program on intentions and behaviors for family planning in teens with diabetes." *Diabetes Care* 36, no. 12 (2013): 3870-3874.

Clarke, William L. "Behavioral challenges in the management of childhood diabetes." *Journal of diabetes science and technology* 5, no. 2 (2011): 225-228.

Commodity12. 2015. *COntinuous Multi-parametric and Multi-layered analysis Of DIabetes TYpe 1 & 2.* Accessed on June 5, 2016. http://www.commodity12.eu/

Dedding, C., R. Reis, C. Ramaker, and B. Wolf. "Hoe denken kinderen over hun diabetes?." *Ned Tijds Diabetologie,* no. 2 (2004):86-89.

Dedding, C. "Delen in macht en onmacht: kindparticipatie in de (alledaagse) diabeteszorg." *Harper & University of Amsterdam*, (2009).

Engström, Maria Svedbo, Janeth Leksell, Unn-Britt Johansson, and Soffia Gudbjörnsdottir. "What is important for you? A qualitative interview study of living with diabetes and experiences of diabetes care to establish a basis for a tailored Patient-Reported Outcome Measure for the Swedish National Diabetes Register." *BMJ open* 6, no. 3 (2016): e010249.

Freeborn, Donna, Carol A. Loucks, Tina Dyches, Susanne Olsen Roper, and Barbara Mandleco. "Addressing school challenges for children and adolescents with type 1 diabetes: The nurse practitioner's role." *The Journal for Nurse Practitioners* 9, no. 1 (2013): 11-16.

Freeborn, Donna, Tina Dyches, Susanne O. Roper, and Barbara Mandleco. "Identifying challenges of living with type 1 diabetes: child and youth perspectives." *Journal of clinical nursing* 22, no. 13-14 (2013): 1890-1898.

Free Style Libre. 2016. *FreeStyle Libre Flash Glucose Monitoring System.* Accessed on June 5, 2016. http://www.freestylelibre.co.uk/

Gonder-Frederick, Linda A., Diana M. Julian, Daniel J. Cox, William L. Clarke, and William R. Carter. "Self-measurement of blood glucose: accuracy of self-reported data and adherence to recommended regimen." *Diabetes Care* 11, no. 7 (1988): 579-585.

Grey, Margaret, Robin Whittemore, and William Tamborlane. "Depression in type 1 diabetes in children: natural history and correlates." *Journal of psychosomatic research* 53, no. 4 (2002): 907-911.

Howells, L., A. C. Wilson, T. C. Skinner, R. Newton, A. D. Morris, and S. A. Greene. "A randomized control trial of the effect of negotiated telephone support on glycaemic control in young people with Type 1 diabetes." *Diabetic Medicine* 19, no. 8 (2002): 643-648.

Jensen, Dorte M., Lars Korsholm, Per Ovesen, Henning Beck-Nielsen, Lars Moelsted-Pedersen, Jes G. Westergaard, Margrethe Moeller, and Peter Damm. "Peri-conceptional A1C and risk of serious adverse pregnancy outcome in 933 women with type 1 diabetes." *Diabetes Care* 32, no. 6 (2009): 1046-1048.

McKnight, J. A., S. H. Wild, M. J. E. Lamb, M. N. Cooper, T. W. Jones, E. A. Davis, S. Hofer et al. "Glycaemic control of Type 1 diabetes in clinical practice early in the 21st century: an international comparison." *Diabetic Medicine* 32, no. 8 (2015): 1036-1050.

Mioch, Tina, Wietse Ledegang, Rosie Paulissen, Mark A. Neerincx, and Jurriaan van Diggelen. "Interaction design patterns for coherent and re-usable shape specifications of human-robot collaboration." In *Proceedings of the 2014 ACM SIGCHI symposium on Engineering interactive computing systems*, pp. 75-83. ACM, 2014.

Murphy, R. R. (2014). Disaster robotics. Cambridge, MA: MIT Press.

National Collaborating Centre for Women's and Children's Health (UK. *Type 1 diabetes: diagnosis and management of Type 1 diabetes in children and young people*. RCOG Press, 2004.

Neerincx, Mark A., and Jasper Lindenberg. *Situated cognitive engineering for complex task environments*. Ashgate Publishing Limited, 2008.

Nordfeldt, Sam, Teresia Ängarne-Lindberg, and Carina Berterö. "To Use or Not to Use–Practitioners' Perceptions of an Open Web Portal for Young Patients With Diabetes." *Journal of medical Internet research* 14, no. 6 (2012): e154.

Owen, Sue. "Every Child Matters." *Early childhood: a guide for students*(2006): 269.

PAL4U.EU. 2016. *Personal assistant for a healthy lifestyle*. Accessed on: January 28, 2016. http://www.PAL4u.eu/

REACTION. 2014. *REACTION European Project.* Accessed on June 5, 2016. http://www.reaction-project.eu/

Schilling, Lynne S., Margaret Grey, and Kathleen A. Knafl. "The concept of self-management of type 1 diabetes in children and adolescents: an evolutionary concept analysis." *Journal of advanced nursing* 37, no. 1 (2002): 87-99.

Scholtz, J. (2003). Theory and evolution of human robot interactions. In *Proceedings of the IEEE 36th International Conference on System Sciences* (Vol. 5, pp. 125-134).

Scott, Leslie K. "Developmental mastery of diabetes-related tasks in children." *Nursing Clinics of North America* 48, no. 2 (2013): 329-342.

Snoek, Frank J., and T. Chas Skinner, eds. *Psychology in diabetes care*. Vol. 13. John Wiley & Sons, 2007.

Svoren, Britta M., Deborah Butler, Bat-Sheva Levine, Barbara J. Anderson, and Lori MB Laffel. "Reducing acute adverse outcomes in youths with type 1 diabetes: a randomized, controlled trial." *Pediatrics* 112, no. 4 (2003): 914-922

Agents Sharing Secrets

Self-Disclosure in Long-Term Child-Avatar Interaction

A THESIS
in Artificial Intelligence

Submitted to the Faculty of Social Science of the Radboud University Nijmegen
In partial fulfillment of the requirements for the degree of
Master of Science in Artificial Intelligence

Franziska Viktoria Burger

Supervisors
Joost Broekens
Mark Neerincx
Pim Haselager

The Netherlands
5 September, 2016

Franziska Burger
Agents Sharing Secrets: Self-Disclosure in Long-Term Child-Avatar Interaction
Supervisors: dr. Joost Broekens, dr. Mark Neerincx, dr. Pim Haselager
September 5, 2016

Radboud University Nijmegen
Artificial Intelligence
Comeniuslaan 4
6525 HP Nijmegen

Abstract

A key challenge in developing companion agents for children is keeping them interested after novelty effects wear off. Self-Determination Theory posits that motivation is sustained if the human feels related to the agent. According to Social Penetration Theory, such a bond can be welded through the reciprocal disclosure of information about the self. As a result of these considerations, we developed a disclosure dialog module to study the self-disclosing behavior of children in response to that of a virtual agent. The module was integrated into a mobile application with avatar presence for diabetic children and subsequently used by 11 children in an exploratory field study over the course of approximately two weeks at home. It was found that the relative amount of disclosures that children made to the avatar was an indicator for the relatedness children felt towards the agent at the end of the study. Girls were significantly more likely to disclose and children preferred to reciprocate avatar disclosures of lower intimacy. No relationship was found between the intimacy level of avatar disclosures and child disclosures. Particularly the last finding contradicts prior child-peer interaction research and should therefore be further examined in confirmatory research.

Acknowledgements

Theses research is not carried out alone and mine is certainly no exception. I would thus like to seize this opportunity to thank those that have supported me academically, technically, socially, morally, or universally.

I extend my gratitude to my supervisors Joost Broekens, Mark Neerincx, and Pim Haselager. I am particularly thankful to Joost and Mark for the accommodations they made before the project started: to Joost for accepting the role of daily supervisor without having met me and to Mark for making certain that I could carry out the research despite the initial difficulties. I very much appreciate the time that both have taken to discuss my progress, the crucial conceptual input that these discussions produced, and their aid in communicating and negotiating with other members of the PAL consortium. I can confidently say that this thesis would not have been possible without them. I am, of course, also indebted to Pim for ensuring that all organizational matters at Radboud run smoothly. I further thank Khiet Truong for encouraging me to apply for a thesis within the PAL project and for introducing me to Mark and Rosemarijn.

Additionally, various members of the PAL consortium have helped me in the development of the module and in recruiting children for the study. Thanks to Frank and Diego for helping me get started and for always having an open ear for all my system/app related issues; to Silvia and Olivier for aiding me in all matters pertaining to the children; to Uli for explaining ontologies to me, converting my files, and passing some useful life advice in the process; to Bernd for looking into the server problems; to Bert for extracting the Quiz data from the server; to Rifca and Anika for helping me translate disclosures and providing data/pictures from the May-evaluation.

I am most grateful, though, for the aid obtained from unexpected places. Here, I would like to voice my appreciation of the outstanding technical support that I received from Ruud de Jong. I would also like to especially thank Thomas Moerland for his invaluable statistical input and for repeatedly being a friend in need.

It is a bit odd to thank someone for their chronic illness, but it is truly thanks to my brother, Lukas Burger, that I have gotten to experience the incredible feat of diabetic self-management second-hand. In the same breath, I would like to acknowledge the support I have received from my family in general throughout my entire studies. Last but not least, I would like to thank my boyfriend, Volker Strobel, for convincing me that I can do anything by myself and for being my safety net nonetheless.

Contents

List of Figures

List of Tables

1 Introduction

Have you ever done the dishes because your wife asked you to? Taken a class because you thought it would look good on your resume? Worn something not because you liked it, but because you thought your critical mother-in-law would? Or gotten drunk simply because your friends convinced you?

Social relationships often play a large motivational role in our behaviors. But we will obviously not do everything for everyone. How much we like or want to be liked by someone is an important factor. This warrants the assumption that when wanting someone to do something, effort should be invested into the bond with said someone.

Type 1 diabetes mellitus (T1DM) is an autoimmune disease of the pancreas. The illness accompanies diagnosed children and adolescents through various physical and mental stages of development. In the PAL project, a **P**ersonal **A**ssistant for a healthy **L**ifestyle is developed with the aim of increasing the self-management skills of diabetic children (ages 7-14) by supporting them, their caregivers, and health-care professionals in sharing responsibility. The PAL robot and its mobile avatar are intended to function as a pal for the children, helping them accomplish their diabetes-related goals through person- and time-adaptive, engaging interactions. The core of the PAL system includes an embodied conversational agent (ECA) in the form of a robot and its mobile avatar, an Authoring & Control tool for health care professionals, a Monitor & Inform tool for caregivers, and a mobile health application (MyPal) with avatar presence. All these components are intended to connect to a common knowledge base, the *PAL cloud*. The PAL architecture is illustrated in Figure 1.

No child wants to have diabetes mellitus. No child wants to be woken up in the middle of the night to measure blood sugar levels, weigh food every time before eating, or have parents nag that they are not taking their illness seriously enough. Yet, strict adherence to a medical regimen is crucial to prevent many of the health risks associated with diabetes. Ways of increasing the motivation of children to comply with their medication requirements are therefore desirable. Within the Horizon 2020 PAL project, we thus explored the possibilities and limitations of creating a bond between diabetic child (8-12 years) and a virtual companion agent through self-disclosure with the goal of increasing the motivational capacity of the agent.

According to Self Determination Theory (SDT), successful establishment of a social bond between human and agent leads to sustained motivation both to interact with the agent and to engage in activities that the agent proposes. SDT [9] argues that the basic psychological needs for *autonomy*, *competence*, and *relatedness* must

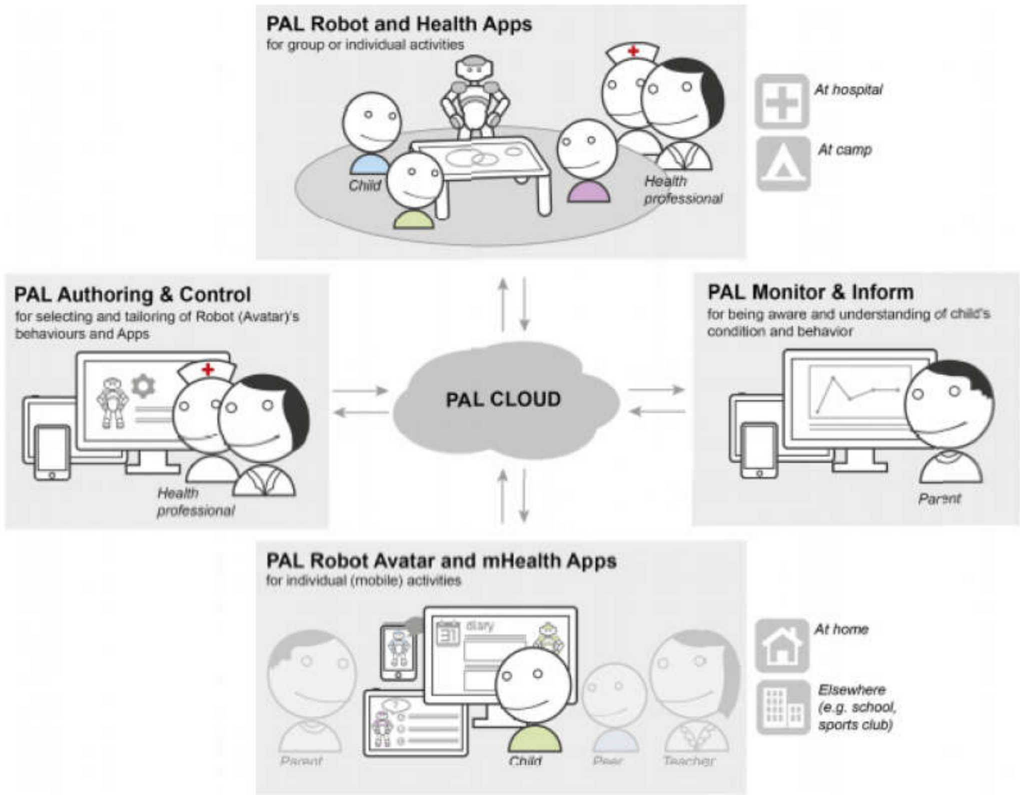


Figure 1: Illustration of the PAL architecture.

be satisfied by the social environment for humans to feel motivated to attempt a task. Relatedness here refers to the feeling that one is accepted and cherished by another individual or community. It comes into play when the intrinsic motivation to engage in an activity is low. More simply put: if we like or want to be liked by someone, we feel more inclined to do what they suggest, even if we are not too fond of the activity itself.

The manner in which such a bond could be established is described by Social Penetration Theory (SPT) [1]. It proposes a directional development of interpersonal relationships whereby the involved individuals first share and explore each others personalities at a superficial level before disclosing more intimate information. Disclosing proceeds along two dimensions: breadth and depth, with *breadth* describing the number of different topics that are disclosed about and *depth* describing the personal value these topics have. Finally, an important determinant of self-disclosure is reciprocity. This describes the tendency to self-disclose as a result of being disclosed to. Reciprocal disclosures in successfully progressing relationships are usually on a similar level of intimacy.

One of the key interests in human-human self-disclosure research has been the close link between disclosure and liking. Specifically, three persistent disclosure-liking effects have been identified [8]: (a) the more someone intimately discloses to us, the more we like that person, (b) the more we like someone at the outset of the interaction, the more we will disclose, and (c) the more intimately we disclose to someone, the more we like that person.

To the best of our knowledge, no study exists that investigates these effects in child-child interaction. However, when children were asked what a friend is and what differentiates a friend from a non-friend, children older than nine indicated that friends take an interest in each others problems and care for their friend's emotional well-being. Additionally, it is argued that cooperation and the insight that each child should contribute equally to the interaction can be expected in this age group [25]. In line with this, 6th grade children's liking of another child was influenced by that child's ability to match the intimacy level of a disclosure while that of 4th graders was not [22].

Support for the disclosure-liking effect has also been found in the domains of human-robot (HRI) and child-robot (cHRI) interaction. In [18], a computer first disclosed some information about itself before asking the user an interview question. As hypothesized, interviewees shared more intimate information with the computer that told personal information about itself but only if this personal information would gradually increase in intimacy throughout the interview. However, the liking for the computer only depended on the sharing of personal information and was not influenced by the intimacy strategy. When a robot was used to elicit self-disclosures from children, those who were prompted to disclose to the robot described the robot significantly more often as a *friend* than children in the control condition [14]. In [13], a two-month study was conducted in an elementary school with a relational robot capable of identifying children and calling them by name, showing more varied behavior with time, and disclosing personal information as a function of a child's interaction time. It was found that children's desire to be friends with the robot at the end of the study was positively correlated with the interaction time.

In summary, one possibility for sustaining motivation is by leveraging relatedness. SPT provides the necessary tool for establishing relatedness: reciprocal self-disclosure with increasingly intimate content. Human-machine interaction studies further indicate that a bond between user and machine can be created through self-disclosure. Two knowledge gaps can be identified from the related literature. For one, there has been no empirical investigation of whether and how the sharing of disclosures between user and system contributes to sustaining user motivation over longer periods of time. For another thing, studies on self-disclosure reciprocity in child-child interaction have been conducted mainly in North America several decades ago (compare [7, 21, 22]). It was therefore uncertain whether insights transfer to today's children in Europe or to child-robot interaction. Furthermore, studies conducted within the framework of the ALIZ-E project[1] also showed differences between healthy and diabetic children with regard to robot interaction.

The here described research presents a first step in closing these knowledge gaps. We developed the initial prototype of a dyadic disclosure dialog module (3DM) to gain insights into how and how readily diabetic children respond to self-disclosures of an ECA and to learn about the possibilities of sustaining children's motivation in this way. A situated approach was taken by integrating the module into a mobile application for diabetic children to be used in an uncontrolled environment for a period of two weeks.

The following two broad research interests guided this exploratory investigation:

1. How do children respond to a self-disclosing avatar?

2. What are the possibilities and limitations of establishing relatedness through self-disclosure and motivation through relatedness in the context of the MyPal application?

The upcoming section, Section 2, briefly describes 3DM and how it was developed using the situated Cognitive Engineering method [19]. Section 3 then details how we used the module within the framework of the MyPal mobile application in an exploratory, long-term field study with diabetic children to obtain answers to the above research questions. In so doing, we found that while children did not match the intimacy of disclosures from the ECA, those children who replied more actively to the disclosures also felt more related to the avatar. Furthermore, children were more likely to reciprocate a disclosure when it was of lower intimacy or when the child was a girl. These findings are further elaborated in Section 4. The extent to which these results can provide answers to the research interests is discussed in Section 5. Finally, Section 7 concludes the article by indicating which findings should be revisited in future confirmatory experiments and how the module can be developed further.

[1] `http://www.aliz-e.org/`

A Disclosure Intimacy Rating Scale for Child-Agent Interaction

Franziska Burger( ), Joost Broekens, and Mark A. Neerincx

TU Delft, Interactive Intelligence, Delft, The Netherlands
f.v.burger@tudelft.nl
http://ii.tudelft.nl

Abstract. Reciprocal self-disclosure is an integral part of social bonding between humans that has received little attention in the field of human-agent interaction. To study how children react to self-disclosures of a virtual agent, we developed a disclosure intimacy rating scale that can be used to assess both the intimacy level of agent disclosures and that of child disclosures. To this end, 72 disclosures were derived from a biography created for the agent and rated by 10 university students for intimacy. A principal component analysis and subsequent k-means clustering of the rated statements resulted in four distinct levels of intimacy based on the risk of a negative appraisal and the impact of betrayal by the listener. This validated rating scale can be readily used with other agents or interfaces.

Keywords: Long-term cHRI · Self-disclosure intimacy · PAL project

1 Motivation

In a focus group conducted with diabetic children in 2012, it was found that children would like a companion robot to share their secrets with and to listen to them when they are sad [1].

According to Self Determination Theory (SDT) [3], successful establishment of a social bond between human and agent leads to sustained motivation both to interact with the agent and to engage in activities that the agent proposes. Such a bond could be established through increasingly intimate, reciprocal self-disclosures [4], that is the exchange of information about the self.

One of the key interests in human-human self-disclosure research has been the close link between disclosure and liking. For example, it was found that 6th grade children's liking of another child was influenced by that child's capacity to match the intimacy level of a disclosure while that of 4th graders was not [5].

To better study children's disclosure behavior when interacting with a virtual agent, we developed the Dyadic Disclosure Dialog Module (3DM) within the framework of the PAL project[1] and using a situated Cognitive Engineering (sCE) [2] approach. This, in turn, necessitated the development of a rating scale for intimacy of self-disclosure.

[1] http://www.pal4u.eu/.

D. Traum et al. (Eds.): IVA 2016, LNAI 10011, pp. 392–396, 2016.
DOI: 10.1007/978-3-319-47665-0_40

2 Intimacy Rating Scale

To design agent disclosure statements at various intimacy levels and to assess the depth of children's disclosures, a rating scale for disclosure intimacy was needed. For this, the following constraints were identified: (a) the scale should discretize the intimacy continuum, (b) each discrete level should have a clear definition, (c) the scale should have a minimum of three levels [6, Chap. 13], (d) the scale should be neither topical nor example-based. Upon reviewing the relevant child and adult literature on self-disclosure, no entirely suitable intimacy scale could be found. We therefore developed and validated the Disclosure Intimacy Rating Scale (DIRS).

As summarized in [7], intimacy of self-disclosure is directly related to vulnerability of the discloser. In a similar vein, it is argued in [8] that the social risk associated with disclosing determines the depth of disclosure. With each self-disclosure, we risk "social rejection [or] betrayal" [8, p. 180].

$$risk(\mathrm{SD}) = risk(\mathrm{SR}) + risk(\mathrm{B}) \tag{1}$$

with SD := self-disclosure, SR := social rejection, and B := betrayal. Betrayal, here, describes the passing on of information by the recipient to third parties.

Risk can be formalized as the product of probability (P) and impact (I). If we further assume that social rejection does not occur at random but only follows if the disclosure is negatively appraised, we can approximate the risk of social rejection through the risk of negative appraisal:

$$risk(\mathrm{SD}) = P(\mathrm{NA}) * I(\mathrm{NA}) + P(\mathrm{B}) * I(\mathrm{B}) \tag{2}$$

with NA := negative appraisal.

The probability of betrayal, $P(\mathrm{B})$, can depend only on characteristics of and prior experiences with the disclosure recipient. It is therefore independent of the content and cannot be considered in the level definitions.

These considerations initially yielded six intimacy levels. Using these, a total of $6(level) \times 3(topic) \times 2(valence) \times 2(repetition) = 72$ statements were fabricated by the first author with the personality and biography of the ECA providing content and style information. To obtain a first validation of the scale, the statements were rated for intimacy by 10 university students (5 female, $M_{age} = 23$, $SD_{age} = 1.612$) on a six-point scale: only levels 0 and 5 were labeled with *not at all intimate* and *extremely intimate* respectively. We decided against asking adult participants to take on the perspective of a child (because results would be questionable in terms of validity) or to rate statements as if coming from a robot (because students are more critical towards the plausibility of a robot expressing emotions and a personality). The biography was hence slightly adapted to fit a 22 year-old student. Before rating, participants were asked to read a persona description of the student and instructions explaining self-disclosure. Intimacy was defined as: "the degree to which a statement reflects information about the self that is sensitive." Further, they were given one example disclosure for each level using a fourth topic. The intimacy levels of the examples was not provided. Participants could thus get an impression of the covered range and the type of statements. Participants found the description of the student and the statements to be believable (the mean believability rating on a 5-point Likert scale was $M_{believability} = 4.3$). The inter-rater reliability

Table 1. The four intimacy levels of the DIRS that resulted from the post-analysis.

Risk	Definition	Example
Low	P(NA), I(NA), and I(B) are low or zero: the discloser cannot be evaluated on the basis of the statement or the statement is very common-place	"I have a lot of brothers and sisters."
Moderate	P(NA) is moderate, because statements are more opinionated, but I(NA) and I(B) are low. Negative appraisal can at best take the form of disagreement. The information cannot really be exploited, so that in the case of betrayal, no loss is to be expected. Includes preferences and opinions on activities and objects	"I like online games in which you have to team up with other players."
High	Either P(NA) is high and both I(NA) and I(B) are low (the content conflicts with the norms of the recipient but does not reflect on the character of the discloser), or P(NA) is low but the content is of great significance to the discloser so that $I(NA)$ and I(B) are high. Disclosures are emotional and may include evaluations of other people	"I'm really disappointed that my sister will not try yoga with me. She already promised it twice but never followed through."
Very high	P(NA), I(NA), and I(B) are high, because the disclosure is at the core of the discloser's self-concept and could easily conflict with the norms of the recipient. In the case of betrayal, great emotional, physical, or material damage may ensue. Social stigmas, self-doubt, deep personal fears and secrets are accumulated on this level	"Whenever I work really hard or I'm nervous, I start sweating like crazy. I can't get close to people then, because I'm really conscious of how I smell."

was assessed using the two-way random intraclass correlation coefficient with the ten raters, yielding $ICC(2, 10) = .947$. Cronbach's alpha using all items was high with $\alpha = .948$. The Pearson correlation coefficient between the level of an item and the average rating it received across participants was determined to be $r = .85$. To check whether we would also find six intimacy levels back in the item pool, a principal component analysis was conducted on the ratings of all items. Using the point of inflexion as a cut-off criterion [9], four principal components explaining at least 10 % of the variance each and 67 % in total were revealed. *Four* was then used as the desired number of clusters in a k-means clustering algorithm. A post-analysis of the resulting item clusters afforded the four intimacy levels of the DIRS detailed in Table 1.

3 Conclusion

3DM is intended to gain insights into how and how readily children in late childhood disclose to an artificial agent. Whether children absolutely, relatively, or do not at all match the intimacy level of a robot's disclosure [10] being a main matter of interest. The DIRS is a supplementary instrument for 3DM to code and compare the intimacy levels of children's disclosures in response to agent disclosures. In an exploratory study using 3DM, 114 child-disclosures were rated by two independent raters using the DIRS. Raters agreed in 67 % of cases and deviated by 1 level in 27 % of cases. However, the disclosures were gathered in the field over the course of two weeks and children were found to mainly disclose on the lower two levels to the ECA (only in 26 % of disclosures was the mean rating of both raters larger than 1).

There are two main limitations to the DIRS. The first is that contextual information of the disclosure is unknown or ignored, and can only be estimated by the rater. As such, raters should have the same cultural background as the discloser. An additional limitation is that the DIRS has only been validated with adults, but is used with children in the PAL project. A next step is therefore to validate it with children of the target age group and using the original biography of the ECA.

In summary, we developed and validated the Disclosure Intimacy Rating Scale to rate statements for intimacy. This scale can be readily used across different human-robot interaction contexts.

References

1. Baroni, I., Nalin, M., Baxter, P., Pozzi, C., Oleari, E., Sanna, A., Belpaeme, T.: What a robotic companion could do for a diabetic child. In: 2014 RO-MAN: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, pp. 936–941. IEEE Press (2014)
2. Neerincx, M.A., Lindenberg, J.: Situated Cognitive Engineering for Complex Task Environments. Ashgate Publishing Limited, Aldershot (2008)

3. Deci, E.L., Ryan, R.M.: Overview of self-determination theory: an organismic dialectical perspective. In: Handbook of Self-determination Research, pp. 3–33. University Rochester Press (2002)
4. Altman, I., Taylor, D.: Social Penetration Theory. Holt, Rinehart & Mnston, New York (1973)
5. Rotenberg, K.J., Mann, L.: The development of the norm of the reciprocity of self-disclosure and its function in children's attraction to peers. Child Dev. **57**, 1349–1357 (1986)
6. Tardy, C.H.: A Handbook for the Study of Human Communication: Methods and Instruments for Observing, Measuring, and Assessing Communication Processes. Greenwood Publishing Group, New York (1988)
7. Moon, Y.: Intimate exchanges: using computers to elicit self-disclosure from consumers. J. Consum. Res. **26**(4), 323–339 (2000)
8. Omarzu, J.: A disclosure decision model: determining how and when individuals will self-disclose. Pers. Soc. Psychol. Rev. **4**(2), 174–185 (2000)
9. Cattell, R.B.: The scree test for the number of factors. Multivar. Behav. Res. **1**(2), 245–276 (1966)
10. Rotenberg, K.J., Chase, N.: Development of the reciprocity of self-disclosure. J. Genet. Psychol. **153**(1), 75–86 (1992)