

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 Neerinx

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.....	5
2. Background.....	8
2.1 PAL project.....	8
2.1.1 Pal Prototype	8
2.1.2 Pal evaluation cycle.....	11
2.2 Diabetes self-management.....	12
2.2.1 Definition	12
2.2.2 Knowledge development as self-management determinant	12
2.3 System usage	13
2.4 Longitudinal research	13
2.4.1 Definition and characteristics.....	14
2.4.2 Longitudinal study designs.....	14
2.4.3 Necessity of longitudinal research in HCI	15
2.5 Individual difference in system usage and knowledge development.....	16
2.5.1 System usage	16
2.5.2 Knowledge development.....	18
3. Methods	20
3.1 Participants	20
3.2 Materials	21
3.2.1 Questionnaires.....	22
3.2.2 Semi-structured interviews.....	23
3.2.3 Information letters and informed consent	23
3.2.4 Audio and video recording.....	24
3.2.5 Statistical computing programs	24
3.3 Design.....	24
3.3.1 Experiment overview	24
3.3.2 Experiment design.....	25

3.3.3 Procedure.....	25
3.4 Measures.....	26
3.4.1 Explanatory variables.....	26
3.4.2 System usage - timeline.....	27
3.4.3 System usage – Quiz	27
3.4.4 Knowledge development.....	28
3.5 Analysis	28
3.5.1 Data preparation	28
3.5.2 System usage – timeline data	29
3.5.3 System usage – quiz data	29
3.5.4 Knowledge development.....	29
4. Results.....	30
4.1 System usage	30
4.1.1 System usage –timeline	30
4.1.2 System usage – quiz	34
4.2 Knowledge development	36
5. Discussion.....	37
References.....	41
Appendix A Diabetes know and do goals.....	45
Appendix B Outcome variables	48
Appendix C Individual scatterplots.....	51

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 is 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 extent 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.

2. Background

2.1 PAL project

The PAL project strives to develop a learning system for children with diabetes to learn about and assist them in the management of their diabetes. The system needs to adapt its communication to the user in order to create an adaptive learning situation for different kind of users. The following section describes the total PAL system and consequently the total PAL four year project analysis cycle. This current explorative study takes place during the first project year and is a formative evaluation.

2.1.1 Pal Prototype

The current PAL prototype consists of a physical NAO robot, its avatar, MyPal, Pal Control and Pal Inform. For this study and experiment the main focus lies on the NAO robot and MyPal application. Children can play a quiz on both diabetes related and general questions with the NAO robot, and interact with him through speech and touch. The MyPal app consists of multiple sections: a quiz and the timeline section. Both of these PAL elements will be explained in more detail throughout this chapter. The avatar in the MyPal application is a digital representation of the physical robot and communicates in the same way. PalControl enables the diabetes healthcare professional to enter and adapt self-management learning goals for their patients. These goals will be visible in the MyPal and Pal Monitor apps. PalMonitor allows parents to see an overview of the child's condition and progress on their learning goals. PalControl and PalMonitor will be left out of the scope in this study, as they are not fully implemented in this year's experiment, and lie beyond the goals and interests of this particular study. The so called PALactor is the underlying reasoning mechanism which manages all behaviours of the robot in both its physical and virtual form in the MyPal app. It also manages the goals of the child and quiz. Interaction with the system will thus be consistent throughout the experiment and across platforms, as the avatar and robot share the same underlying reasoning mechanism. Further versions of the system will manage the personalised communication and content based upon the continuous information it receives on the child's preferences, goals and status through the quiz, timeline and communication. Figure 1 shows an overview of the system and its components.

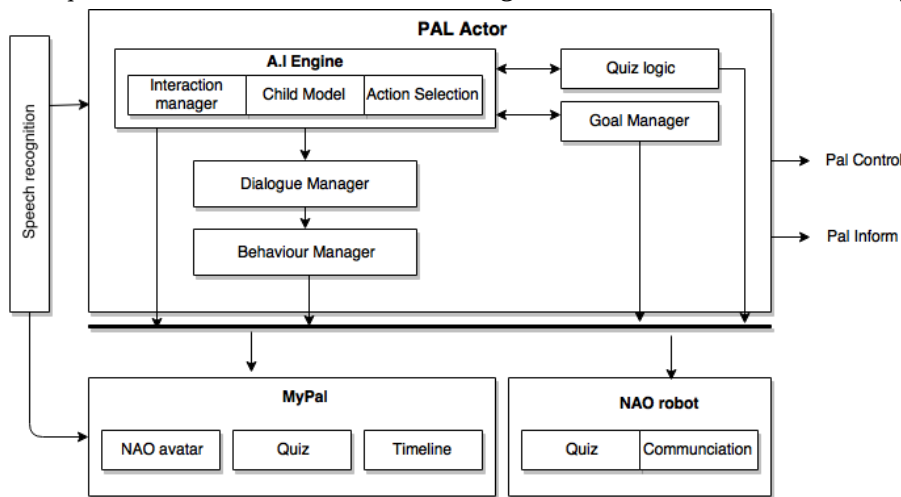


Figure 1. Overview of the PAL system components.

Nao Robot

The NAO robot, as illustrated in figure 2, is a small humanoid robot that has been used for multiple purposes like health research and education. Many examples of studies that used a NAO can be found in therapy. Shamsuddin et al. (2012) used the NAO for research on robot interactions with autistic children, and how robotics may play a role in future interaction therapy. Kose, Yorganci, Algan and Syrdal (2012) described how a NAO robot can be of assistance in the learning of sign language.

The robot can move around freely, lay down, get up and make hand gestures. The sensors in his limbs and head, as well as the cameras and microphones allow him to perceive and act on his environment or users. Communication with others around him can be done by speech and movement of arms, hands and head. The NAO derives what actions he has to take from an action selection module. This module assesses the quality of proposed action suggestions given by a Natural Multimodal Interaction Module (NMIM). It does this based upon the current state of the child and its preferences through speech recognition and learning goals as set by the health care professional.

For this study, the NAO interacts with the patients and plays a quiz with them about diabetes and general topics. The diabetes related questions for the quiz were constructed on the base of the know- and do goals as set by the Expertisegroup Paediatric Diabetes Nurses (EPDN). The questions were validated and altered by diabetes healthcare professionals in the cooperating hospitals, and sorted on difficulty level. Categories on which the avatar may present diabetes related questions are: blood glucose, insulin, food, physical activity, hypo and general. An overview of these goals are situated in appendix A. The general non-diabetes related questions were derived from the prior robot interaction research ALIZ-E by TNO. Any interactions outside the quiz range from small talk to walking together. All interactions will be personalised towards the child in further system versions, based on his goals, preferences and the situation to encourage a long-term relationship with the system and stimulate motivation to learn diabetes self-management behaviours.



Figure 2. NAO, humanoid robot used during the PAL project.

MyPal

The MyPal application consists of three elements: the NAO avatar, timeline and quiz. The NAO avatar is a digital representation of the physical NAO and is connected to the same underlying reasoning system and database. Its general interaction is therefore the same, apart from physical restrictions such as the ability to walk together. To foster a long-term relationship with the system (and NAO) and motivate the child to use the application, the avatar responds to entries the child makes in the timeline and answers he gives in the quiz. Communications may be encouragements to the child to tell more about a certain activity, or casual conversations that fit the situation. These interactions will be personalised to the child's goals, preferences, MyPal entries and situation to create both an adaptive learning environment and relatedness through human-like communication. However, in the current version the personalisation has not been fully implemented. While there is still feedback on interactions, this is not personalised to the users characteristics or his development. The application further includes a virtual quiz. Just as with the physical robot, the avatar will present both diabetes and non-diabetes related questions based on the child status, progress and learning goals.

The timeline in MyPal provides the child with the possibility to enter daily events with or without the help of the NAO avatar. In the timeline one can enter data on food, activity and glycaemia values. When adding a meal, the child can enter how much and what he or she ate. It is optional to enter whether the glycaemia levels were measured, and if an insulin injection was administered. In the activity section, the child can enter what activity he or she performed, and for how long. The child can write notes and disclose any emotions. In the glycaemia section, the child can enter measured values, how he or she corrected them, and optionally write additional notes. Figure 3 shows two sections of the timeline.

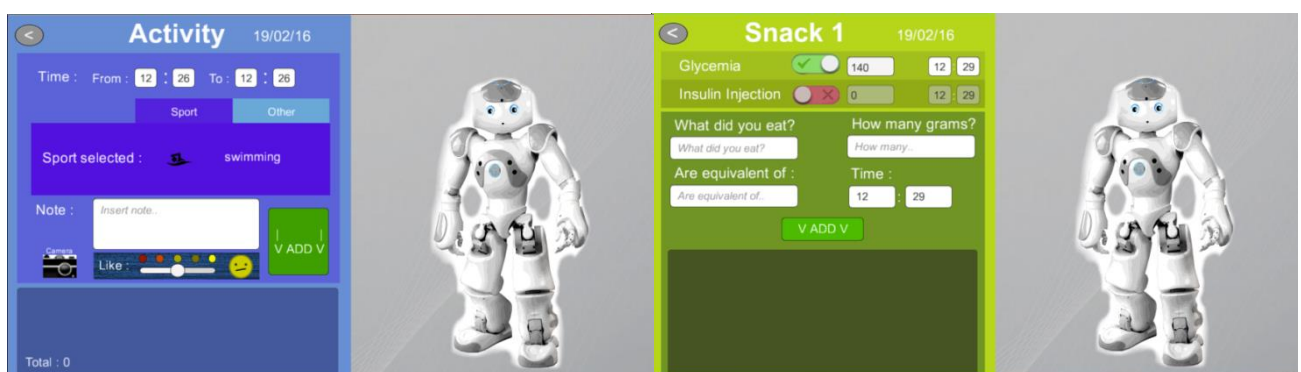


Figure 3. Two sections of MyPal timeline: activities and meals.

2.1.2 Pal evaluation cycle

The total project evaluation of the PAL system will be done in three cycles of one year. The development takes an iterative design approach as after each cycle, the system is tested, and adjusted according to results. An overview of the three development cycles can be found in table1. Every cycle increases the scope of evaluation by increasing the usage duration period, the age range of the participating children and focus of self-management development. The project is currently in cycle one, which focusses on the development of five determinants of diabetes self-management in children. These determinants are knowledge, awareness, attitude, self-efficacy and skills. Acceptance and trust of healthcare professionals and parents in the system are also included during the evaluation. During this stage, the project focusses on the age group from 7-10 years old. The elements that are tested in the MyPal application are the quiz, avatar and timeline for the duration of one month. For this thesis, the emphasis will remain on the evaluation of development of diabetes related knowledge and system usage. An overview of the experiment is described in chapter 3.

Table 1.

Overview of the three design cycles of the PAL system.

	Initial Phase	Intermediate Phase	Final Phase
Determinants	Knowledge & Awareness	Diabetes Regimen Adherence	Shared Child-Caregiver Responsibility
Children - age dependent claims	<u>7-10 yr.:</u> + knowledge + awareness + attitude + self-efficacy + skills	<u>7-12 yr.:</u> + regime adherence + glucose monitoring	<u>7-14 yr.:</u> + shared responsibility + coping with anomalies - hypos/ hypers + glycaemic control
Usage	1 month	4 months	9 months
Caregivers - claims	<i>Professionals:</i> + trust + acceptance <i>Parents:</i> + attitude + knowledge	<i>Professionals:</i> + awareness <i>Parents:</i> + trust + skills	<i>Professionals:</i> + tailoring <i>Parents:</i> + 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

2.2 Diabetes self-management

As the main purpose of the PAL system is to stimulate and support the diabetes management of diabetic children, it is important to understand what this self-management exactly entails. The following section describes diabetes self-management and self-care behaviours that accompany this. Then, the knowledge determinant for performing diabetes self-management, and why this thesis chooses to investigate knowledge development is described.

2.2.1 Definition

Self-management of one's health is generally accepted to be not just any kind of self-care behaviours such as personal hygiene maintenance. It refers to behaviours and processes which are specifically required to manage (outcomes of) a (chronic) disease (Lorig & Holman, 2003). Although there have been multiple definitions in healthcare research, this thesis assumes the definition by Modi et al. (2012) which defines self-management as “the interaction of health behaviours and related processes that patients and families engage in to care for a chronic condition (p. 475)”. The reason for this is that they do not only take the individual patient, but also the close social circle into consideration. The self-management of a disease often exceeds ‘just’ the patient himself. The family and close social circle often perform similar behaviours and processes to manage the disease. In the context of the PAL project all patients are children, which makes this involvement of family an important aspect to include. For diabetes patients, self-management includes the management of blood glucose, adherence to medical plans, making dietary adaptations, (planning) physical activity and making psychosocial adaptations (Mulcahy et al., 2003). For parents, support of the child's self-management mainly entails the supervision and support on all these aspects, while providing access to healthy nutrition, exercise and healthcare (Modi et al., 2012).

2.2.2 Knowledge development as self-management determinant

For performing actual self-management patients will experience driving factors (determinants) which encourage the behaviour such as attitude, and barriers which inhibit the behaviour such as high costs. The total PAL project aims to assess five behavioural determinants during the first cycle evaluation: knowledge, awareness, attitude, self-efficacy and skills.

This thesis focusses on the development of the determinant ‘diabetes related knowledge’. Although no conclusive evidence has been found that diabetes related knowledge has a direct impact on health outcomes such as glucose values, some studies have stated a relation between knowledge and the number of self-management behaviours (Persell et al., 2004). Most studies agree that although it may not be a direct and only determinant of actual self-management it does play an important part in it (Coates & Boore, 1996). Furthermore, a base level of diabetes related knowledge is assumed to be necessary for basic self-management, as a total lack of knowledge appears to relate to diminished prevention of later complications (Pace, Ochoa-Vigo, Caliri, & Fernandes, 2006). This is particularly important in this current study as a base level of knowledge may not yet be present in a lot of our

patients due to their young age and needs to be developed.

The Dutch Expertisegroup Paediatric Diabetes Nurses (EDPN) has constructed a description of so called ‘know and do goals’. This document contains all major knowledge and skills children with diabetes are expected to obtain at a certain age. This document was converted into a list of specific goals for each age group (6-7, 8-9, 10-11 years), and indicated with a difficulty level. Goals that should be obtained between the ages of 6-7 are level 0, goals between the ages of 8 and 9 are level 1 and goals between the ages of 10 and 11 are level 2. Goals for older children are left out of consideration as only children between the ages of 6 and 10 years old will be studied during the evaluation of this development cycle. For this study the focus will remain on the knowledge goals, which are represented in table 1 in appendix A. The knowledge goals are divided into six categories which constitute knowledge of: blood glucose, insulin, food, physical activity, hypo and general diabetes knowledge. These categories represent the needed knowledge of the self-management behaviours the children have to develop. For example, learning how to monitor your blood glucose is related to the self-management act of actually measuring the blood glucose levels. These knowledge goals are represented and aimed to achieve by diabetes related quiz questions in the MyPal application.

In conclusion, as a positive knowledge development may improve the self-management behaviours of children with diabetes, development of diabetes related knowledge is set (besides system usage) as a main parameter to investigate in this research.

2.3 System usage

Given that most healthcare products and interventions are developed for an extended period of use by the patient, one would expect that both the effects and use of interventions are still being evaluated after a longer period of time. Limited resources like money and personnel is however one of the reasons that this does not happen for the majority of developed interventions (Gerken, 2011). Creating an intervention which ends up never being used is of course never the intention of a study. However, by not investigating the system usage over an extended period of time, one can never be sure whether and how it is still being used let alone whether any effects of the treatment remain after a longer time period. Furthermore, the few studies that do take a longitudinal approach in intervention system development, for example that of Van Gemert-Pijnen, Kelders, & Bohlmeijer (2014), report a stalling of continuous and thorough usage over time. This might pose serious issues for the possible effects of the intervention, as it cannot help if it is not used.

2.4 Longitudinal research

For the experiment in this study, a longitudinal design was chosen in order to track both the development of system usage and knowledge over a longer period of time. The following section describes the characteristics, pros and cons and most importantly the absolute necessity of longitudinal experiment designs in HCI research.

2.4.1 Definition and characteristics

While the precise definition varies among researchers, a common ground is found in the basic assumptions of longitudinal research. These assumptions or conditions are that data on a specific variable is measured through multiple distinct periods, and some form of comparison of the data of different periods (Menard, 2002; Ployhart, & Vandenberg, 2010). Although the study of Ployhart and Vandenberg (2010) argue that a longitudinal study should contain at least three data gathering periods, and Menard (2002) stated that two are the bare minimum, most studies agree on the statement of Willet (1989) that more data gathering periods result in more reliable measurement of change. This is however often accompanied by higher costs in both administrating and analysing more data gathering periods.

In longitudinal research designs, as in cross-sectional research, the chosen sample and comparison style such as within participant or between participants comparison are still important. More interesting however is the importance of study duration and data-gathering schedule. Determining the study duration is of great importance regarding the research question. Certain developments may take (more or less) time depending on what is measured. For example, in a study on the relation between smoking and lung cancer, it may be necessary to study participants for more than 10 years to see effects. In the case of the PAL study, the development of knowledge on diabetes self-management may be most relevant to study for several months, to examine initial system usage and the first developments in knowledge and behaviour. However, additional research after a longer period of time may be useful to examine long-term adherence and sustained effects in knowledge and behaviour.

2.4.2 Longitudinal study designs

Although recent HCI studies report a variety of longitudinal study designs as most common, these can be divided into three main study designs: panel, retrospective panel and repeated cross-sectional (Gerken, 2011; Menard, 2008). A repeated cross-sectional design entails repetition of data periods during which each period another comparable sample of participants is observed (Menard, 2008). The main reason for using this design is that in many cases participants don't want to, or simply can't, participate in the study for a considerable time length such as 20 years. Also, by introducing new participants every data gathering period, development processes inherent to the individual will less bias the overall picture of the investigated change process (Gerken, 2011).

Panel study designs studies the same sample of participants over different periods of time by several data gathering moments. In general, four retrospective designs can be distinguished. The repeated sampling method includes just one test period and two data gathering points (at the beginning and the end of the period). We must be aware that by the definition of some studies such as that of Vandenberg (2010), some researchers may not consider this a truly longitudinal study as it contains less than three data gathering moments. A study with at least three data gathering moments is

considered a prospective panel design. For prospective panel designs several cautions need to be mentioned. Self-report data from one measure moment might influence what answer the participant gives the following data gathering moments (Trivellato 1999). It might also be that the participant develops not through the intervention, but mere from participation in the study (Menard 1991). Longitudinal case studies, which are a part of prospective panel design, investigate a smaller sample for a longer period of time with the main goal to further observe and understand particular behaviours. The revolving panel design aims to address these issues which may cause bias in the results by introducing new participants and letting some participants go in the sample every period. This allows the researchers to investigate potential external influences by comparing with the new unbiased participants. The retrospective panel design includes only one data gathering moment. During this gathering, data is collected from multiple times in a particular period in the past. One form may be a retrospective interview. Data gathered from retrospective interviews may be very prone to recall bias, and is therefore often only applied to factual life events such as time of marriage unless the research question related to interests in for example beliefs and perceived experience (Buck et al. 1995). Shortening the period on which data is gathered may diminish recall bias.

2.4.3 Necessity of longitudinal research in HCI

Even until a few years ago, very little practical longitudinal studies with repeated measures were performed in the field of HCI (Ployhart & Vandenberg, 2010). But increasing attention to the topic during conferences and multiple researches advocating the benefits and necessity of this research method have made longitudinal research of increasing interest in practical HCI studies. Still, longitudinal references studies for HCI in healthcare, particularly in the combined setting of a controlled and uncontrolled environment are scarce. Reasons why the method is not yet widely applied are mainly bureaucratic. High costs and an increase in time and effort in conducting the experiment compared to a cross-sectional study, as well as the difficulty of finding participants who are willing to commit for a longer period are perceived as great barriers (Gerken, 2011). More methodological barriers related to longitudinal research are difficulty in obtaining valid results. Issue in obtaining valid results over time can relate to sampling biases, conditioning effects (Cantor, 2008) and maintaining construct validity through repeated measures (Singer & Willet, 2003). For example, participants who are willing to commit to an experiment for a longer period of time may not be representative (in terms of motivation) for the population one tries to examine.

For research that is concerned with the development of an interactive system however, longitudinal research can be considered not only preferable but sometimes even absolutely necessary to increase its usability and validate its effects. First of all, in order to validate any sort- and long-term effects of the system on the development of knowledge or skills, a longitudinal experiment design is necessary. Otherwise, no statements can be made about actual long lasting effects of an intervention. Secondly, if one wants to study any kind of development, a longitudinal design is necessary to

identify trends (Menard, 2002). The experiment of this thesis for example needs a longitudinal design to study trends in knowledge and system usage development when using the PAL system. Knowledge develops over time and may fluctuate, as is the case of system use. Taking a longitudinal approach is necessary in establishing actual usage and usage patterns in daily life of a system to help establishing if, and when a system is more or less used. Lastly, general consensus is that learnability is a great part of usability (Grossman, Fritzmaurice, Attar, 2009) and longitudinal studies in system development may therefore be necessary to account for learnability. Although a new system may at first appear to elicit worse performances from users, this may be due to the fact that users are trained to use another system. Time and repeated use may be needed to attain the same or higher level of performance of users. This is demonstrated in the longitudinal study by MacKenzie and Zhang (1999), where after repeated use the new design of a keyboard achieved higher performances even though performances on initial uses were far lower compared to a widely used generic keyboard model.

In conclusion, a longitudinal research is necessary in evaluation of interventions if we wish to validate its use and effects in the long run, identify processes or developments, and overcome effects of learnability and initial use on long term performance. For the PAL experiment, this means that we not only need to measure over a longer period of time, but also incorporate repeated measures if we want to validly identify developments in knowledge and system usage.

2.5 Individual difference in system usage and knowledge development

The variables which might explain variation across and within participants on system usage and knowledge development are summarised in table 2 (page 19). These variables should be taken into consideration during and after the experiment as possible explanatory variables. The following section further explains these variables.

2.5.1 System usage

We need to create an understanding what user characteristics may contribute to variations in system usage (Van Gemert-Pijnen, Kelders, & Bohlmeijer, 2014). This is needed as the PAL project tries to personalise its communication, hoping to optimize system usage and ultimately knowledge development for different kinds of users. For example, low-active users with a particular trait such as low self-efficacy may need to be reminded or stimulated in another way to use the intervention. Grounded theories in health behavioural change research are used to narrow our field of possible predictors for system usage which will then be used in the experiment.

The first possible predictor for system usage is not a characteristic of the user himself, but of his judgement of the system: his user experience. Multiple theories have been developed to describe the acceptance of technology (Venkatesh, Morris, Davis & Davis, 2003). Elements of user experience are found in the majority of these theories, with perceives usefulness and perceived ease of use as the main recurring predictors for technology acceptance and use (Castenada, Munoz-Leiva & Luque, 2007; Davis, Bagozzi & Warshaw, 1989). Variation in system usage is not only expected across, but

also within participants. While during the initial use, ease of use is considered most important, usefulness becomes more important for the user through the phase of incorporation of the system in their daily lives (Karapanos, Zimmermann, Forlizzi, & Martens, 2009). More recent studies found that the aesthetics of (web-based) interventions and products should be included in a technology acceptance and use model as it was found to influence the perceived ease of use and perceived usefulness (Cyr, Head & Ivanov, 2006; Sonderegger, Sauer, 2010).

While user experience will be examined during the experiment, other possible predictors for usage of a health intervention system may be found in the *health belief model* (Becker et al., 2010). Although this theory does not directly address system usage, it does address direct health promoting behaviours. In the context of the PAL experiment, system usage would be a direct health promoting behaviour in the intervention. The health belief model was developed and refined through multiple studies during the second half of the 20th century and tries to explain how health promoting behaviours (or the lack thereof) come about (Becker et al., 2010 ; Jahanlou, Lotfizade & Karami, 2013). This early model states four variables as influencers of health promoting behaviour: perceived susceptibility and severity, perceived benefits and perceived barriers. The perceived susceptibility and severity means to what degree an individual judges a disease on the severity and likeliness of getting the disease. Any positive change in reducing the (perceived) susceptibility and severity a person expects after a health promoting action is labelled as a perceived benefit, while expected costs and efforts to achieve this benefits are barriers. General demographics, like age and sex, and socio psychological variables, like personality, intrinsic motivation, socio economic status and social environment, are taken into account as mediators. Jahanlou, Lotfizade and Karami extended this model through multiple experiments and mainly added self-efficacy. Their studies mainly used diabetes as the observed disease which is obviously very relevant for the PAL experiment. Figure 4 illustrates the extended model after their final experiment which tested the behaviour maintenance by the outcome blood glucose (HbA1c) values.

As the PAL experiment deals with young children of which the parents still play a large part in their diabetes management (Modi et al., 2012) they might influence the motivation of the child to interact with the system. Possible influencers are most likely to be similar to the health belief model variables, as the parent is confronted with the same elements of fear, benefits and barriers. Fears in this case are for the health of their child and possible complications of their diabetes. Barriers in motivating and supporting the child might be a lack of confidence and experience with similar PAL technology.

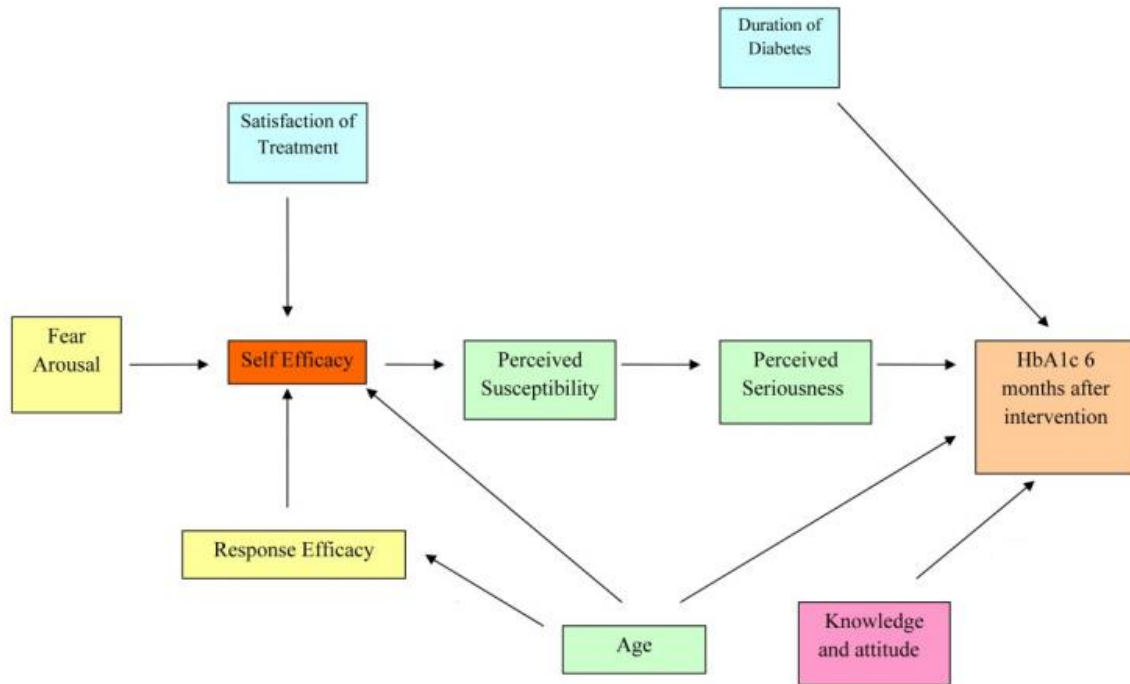


Figure 4. Extended Health Belief Model for Behavioral Maintenance based on HbA1c. Adapted from “A New Behavioral Model (Health Belief Model Combined with Two Fear Models): Design, Evaluation and Path Analysis of the Role of Variables in Maintaining Behavior,” by A.S. Jahanlou, M. Lotfizade, and N.A. Karami, 2013, *Diabetes Mellitus - Insights and Perspectives*, 1, p. 297. Copyright 2013 by InTech.

2.5.2 Knowledge development

The possible predictors for knowledge development are similar to the predictors for system usage. It is important to understand that the system use over time itself needs to be taken into account for differences in knowledge development. It may be that an increase in system use positively relates to an increase in knowledge (which is one of the main goals of the intervention). This would inherently mean that the factors that drive system use might indirectly influence the knowledge development.

Setting this aside, even if system use would not relate to a higher knowledge development, some of the afore mentioned variables may still account for differences in knowledge development on their own. This is especially the case for demographic and socio-psychological factors. For example, it is clear that knowledge development would vary for children of different ages, as their cognitive abilities are different. But also socio-psychological factors like experiencing a divorce of ones parents (Amato, 2005) and learning style (Lynch, Woelfl, Steele & Hanssen, 1998) have shown to relate to learning and cognitive differences.

Table 2

Possible predictor variables for health intervention system use and knowledge development

Possible predictor concept	Operational variable
Demographics	Nationality Gender Age
Socio-Psychological	Divorced parents Learning style Personality (big 5) Socio-economic status
Perceived threat	Experienced diabetes symptoms Experienced treatment barriers Treatment adherence Worry about diabetes Diabetes communication problems Diabetes communication
Perceived barriers	Self-reported diabetes knowledge Technology confidence
User experience	
Perceived ease of use	Comprehensibility Accessibility Visual appeal
Perceived usefulness	Motivation

3. Methods

During this study, one of the aims was to identify any trends and groups in the use of the system and knowledge development with specific characteristics. With these groups and their main characteristics, the PAL system may be personalised with adaptive strategies. This chapter describes the main experiment- and analysis methods that were used to identify the trends, user groups and their characteristics. The main PAL study was conducted by the total PAL project group and consisted of more areas of interest than relevant for this particular study. This means that in practice there were measures and procedures that fell beyond the scope of this study and analysis. Only the for this thesis study relevant methods of the experiment are described.

3.1 Participants

35 children participated in the first year experiment. Of these children, 25 were Dutch and participated in the Netherlands while 10 children were Italian who participated in Italy. In the Netherlands, each child was accompanied by one of their parents during the experiment appointments totalling to 25 participating parents. In Italy however, with the exception of two children, all children were accompanied by both their father and mother during the experiment appointments making the total count of Italian parents 18. An overview of both the Italian and Dutch sample is presented in table 4.

Due to practical planning reasons, the total experiment duration differed for the participating children ($M = 25.7$, $SD = 6.558$). The total experiment duration mainly differed for the two nationalities. For the Dutch children, the total experiment length ranged from a minimum of 15 days to a maximum of 29 days with an average length of 22.44 days ($SD = 4.119$). For the Italian children, the experiment lasted a minimum of 29 and maximum of 40 days, with a mean of 33.90 days ($SD = 3.780$). The ages of the children ranged from six until twelve years with an average age of 9.29 years ($SD = 1.482$), with a small difference between children of different nationality.

The method of convenience sampling was used for the selection of participants. Participating diabetes healthcare professionals in the cooperating hospitals, Meander MC (NL), Ziekenhuis Gelderse Vallei (NL) and Ospedale San Raffaele (IT), invited their diabetes patients and parents who fitted the criteria through an information letter. The limits were set at a minimum of ten and maximum of fifteen participants per hospital, which would result in a total sample of thirty to forty five patients. A minimum of ten and maximum of fifteen patients per hospital was deemed fit in this case, as this is an observational pilot study, and the main goal was not to make general population statements but to get a preliminary view of the system usage and individual development of diabetes knowledge. The sample size was determined by the project group in a pragmatic manner, based on the number of possible available diabetes patients within the predetermined age range in the three cooperating hospitals.

To enter the study, patients had to be between the ages of six and twelve years old. Another inclusion criterion was that patients needed to be diagnosed with T1DM for longer than six months. This criterion was set to avoid any influences of recent diagnoses. Children could only enter the study if at least one parent was available and willing to participate. The participation of at least one parent was important to ensure guidance and support for attendance at appointments and use of the system at home. Furthermore, they were necessary to be included in the study to get a holistic overview of the child and their family and home environment. An informed consent needed to be filled out by both the parents and children in order to participate, as most of the children were below the age of twelve. The experiment was approved by both the University of Twente ethics committee, and the MEC-U which is the medical ethics committee of a range of Dutch hospitals.

Table 4

Overview of both the nationalities samples and the total sample.

	NL Sample				IT Sample				Total Sample			
	N	% of NL	M	SD	N	% of IT	M	SD	N	% of Total	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
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 with Charlie	13	52.0%			4	40.0%			17	48.57%		

3.2 Materials

In order to identify possible predictors for differences in system usage and knowledge development as determined in chapter 2.5 were operationalized through different materials such as questionnaires and interviews. Some concepts were not possible to include in this experiment, they remain to be tested in further test cycles of the PAL project. Table 5 shows an overview of all used materials, and the possible predictor concepts they operationalized.

Table 5. Possible predictor concepts for health intervention system use and knowledge development that were tested in the experiment, and their operationalization.

Predictor concept	Operational variable	Tested in:
Demographics	Nationality	Family demographics questionnaire (FDQ)
	Gender	FDQ
	Age	FDQ
Socio-Psychological	Divorced parents	FDQ
	Socio-economic status	FDQ
Perceived threat (For parent and child)	Experienced diabetes symptoms	Pedsq TM 3.0 Diabetes module
	Experienced treatment barriers	Pedsq TM 3.0 Diabetes module
	Treatment adherence	Pedsq TM 3.0 Diabetes module
	Worry about diabetes	Pedsq TM 3.0 Diabetes module
	Diabetes communication problems	Pedsq TM 3.0 Diabetes module
Perceived barriers (For parent and child)	Diabetes knowledge	FDQ
	Technology confidence	FDQ
User experience		
Perceived ease of use	Comprehensibility	User experience interview
	Accessibility	User experience interview
	Visual appeal	User experience interview
Perceived usefulness	Motivation	User experience interview

3.2.1 Questionnaires

Perceived threat and susceptibility of the disease was set as a possible determinant for system usage and knowledge development, the PedsQLTM 3.0 diabetes module operationalizes this perceived threat in 5 subcategories (Varni, et al., 2003). The diabetes module addresses general diabetes problems and worries the child experiences: symptoms (11 items), treatment barriers (4 items), treatment adherence problems (7 items), diabetes worries (3 items) and diabetes communication problems (3 items). All items present the patient with a statement and the possibility for him to judge whether this applies to him on a scale of 0 (never) to 4 (almost always). Even though this paediatric self-report has shown to be of high validity and reliability (Varni, Seid, & Kurtin, 2001), the (sometimes different) view of the parent is necessary to provide a thorough overview of the child's health functioning as they are closely intertwined with the treatment (Varni, Limbers, & Burwinkle, 2007). The parental version of the PedsQLTM 3.0 diabetes module questionnaire is therefore included in this study. It measures the same constructs with similar items from the parental view. The parent is, just as the child, provided with statements and the possibility for him to judge whether this applies to his or her child on a scale of 0 (never) until 4 (almost always).

General characteristics which were set as possible mediators or determinants for system usage and knowledge development were addressed in a family demographics questionnaire. These included child demographics, socio-psychological factors, technology experience (confidence) and perceived diabetes knowledge. The parents were asked to fill out this questionnaire.

3.2.2 Semi-structured interviews

During this study, the child participated in three semi-structured interviews with one of the researchers, at the hospital during the first and final appointment, and by phone two weeks after the start of the study. Global structures of discussion topics and questions were made prior to the interviews.

The first interview measured the current state of the self-management determinants, knowledge and skills which were operationalized by a discussion on the know and do goals of diabetes self-management. The motivation before actual system usage of the child to engage with the PAL system was measured through a series of statements which the child can then rate on a scale of 0 to 5. 0 meaning that the statement is very untrue for them, and 5 meaning that the statement is very true for them. The second interview was a quick check-up on the perceived user experience of the MyPal application after two weeks of usage. The interviewer focused on this through several questions on comprehensibility, visual appeal and general feedback. The final interview measured the current state of the self-management determinants, knowledge and skills in the same manner as during the first interview. The user experience after the four weeks of usage was addressed through multiple choice questions in an interview, by the same process of statements which the child had to rate on a scale of very untrue until very true. The user experience was measured by statements and questions on accessibility, comprehensibility, visual appeal, and motivation.

3.2.3 Information letters and informed consent

Every child that fit the inclusion criteria for participation received an information letter of their diabetes healthcare professional. It asked the child if it would like to participate, and explained the procedure in a simple and brief manner. The parents received an adult version which explained the project, its goal, the procedure and privacy of data. If the parents were willing to participate, they would receive a larger and more elaborate information folder, which explained the research, measurements, possible benefits and effects, selection criteria, time scheme of the experiment, privacy and data, financial compensations and options for complaints about the research.

If children and parents started the experiment, both needed to sign an informed consent. This explained that participation was always voluntarily and permission could be retracted at all times without reason. Parents were also asked to give permission for their child to participate, and whether video and audio recordings of their child during the study may be used for research purposes.

3.2.4 Audio and video recording

All interviews were audio recorded in case the parent gave their consent. Interaction sessions at the hospital between the child and physical robot were video recorded for later analysis if written consent for this recording was given.

3.2.5 Statistical computing programs

General descriptive and testing procedures were performed in the IBM program SPSS statistics version 22. For more advanced graph plotting and creating linear mixed effects models, the statistical computing software R version 3.3.0 (R Core Team, 2013) was used within the RStudio environment version 0.99.491 (RStudio Team, 2015).

3.3 Design

3.3.1 Experiment overview

The total experiment lasted four weeks and during this experiment a total of three appointments per participant were arranged. The first appointment was at the start of the experiment at the hospital, the second mid-way after two weeks by phone, and the final one at the end of the experiment in the hospital after four weeks. The parents were also present at every appointment. The appointments were used for interaction between the child and NAO robot, introduction and set-up of the MyPal application, and to gather child data on several topics through interviews and questionnaires. During the whole four week experiment, the children were able to use the MyPal application on a provided tablet at home. In this application, the children could play the quiz, enter data like activities, food and blood glucose on their timeline, and interact with the NAO avatar. They were informed to use the application as often and long as they pleased, with no minimum or maximum usage. The children and parents could determine for themselves if they wished to use the application together, or if the child wanted to use it by himself. Figure 4 provides a visual overview of the total experiment.

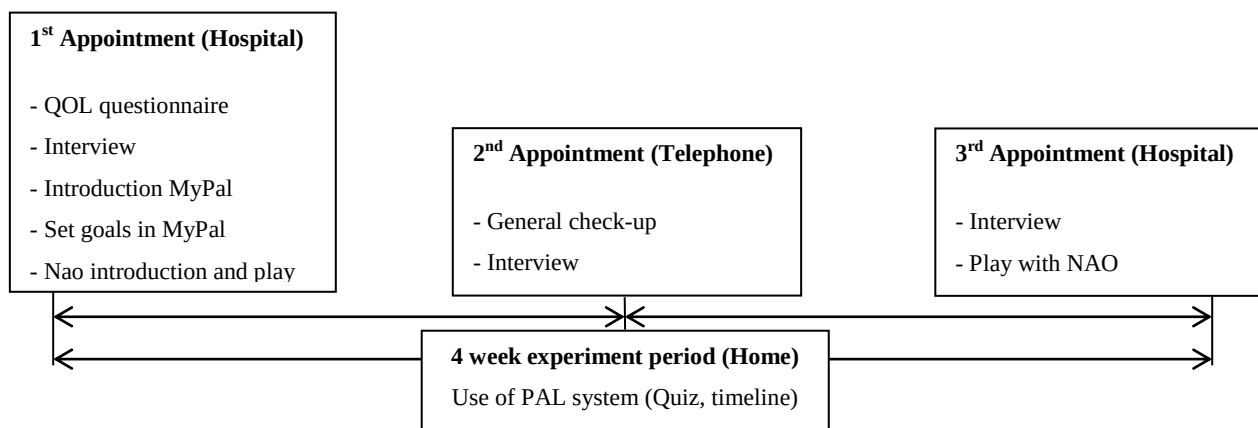


Figure 4. Timeline of the total experiment.

3.3.2 Experiment design

All participants were observed on their use of the PAL system through longitudinal repeated measures. Although the main PAL project took a mainly qualitative approach, this thesis took both a qualitative and quantitative standpoint. All participants received the MyPal application, interacted with the robot and were asked to attend their personal appointments with their researcher and health care professional. The study contained no control group, as the main goal of the study was not yet to make inferences about causality or a general population. During the experiment, data was registered at three set appointments for qualitative interview and questionnaire data, and continuously during the four week usage of the system by the children at home. This research design had the advantage that it allowed to track changes over a period of time, instead of just a pre- and post-measurement. This provided detailed information on the progress rather than depending on averages. For this study, this was useful as we wanted to see how the system usage and knowledge development would progress over the time of one month. Also, this progress was explored both between as with-in participants.

3.3.3 Procedure

Several weeks before the start of the experiment, parents and children of cooperating hospitals were informed about the project through their own diabetes healthcare professional. The parents were provided with an information letter which contained quick information about the experiment. The children were handed a simplified child friendly version. Furthermore, a detailed folder about the experiment and all the participant's rights, along with an informed consent form for both the child and parents were given to the parents. A few days after receiving the information, the parents would be contacted by phone with the question whether they and their child would like to participate in the experiment. If they did, a first appointment for the experiment was made.

For the initial appointment, the diabetes healthcare professional, researcher, parents and child would meet at the hospital. After the possibility to ask questions, the parents and children were asked to sign the informed consent forms before starting the experiment. They could then both start the PedsQLTM diabetes questionnaire. When the child finished his PedsQLTM diabetes questionnaire, he was interviewed by the researcher. Parents were asked to fill out the family demographics questionnaire, after finishing their quality of life questionnaire. Both the parents and child were then introduced to the MyPal application and could set up the child's profile together with the healthcare professional. After further explanation of the application and intended use at home, the diabetes self-management learning goals were then set up in consultation with consideration of the child's current knowledge and motivation as determined during the initial interview. After all formalities, the child was introduced to the NAO robot and could play and interact with him for about twenty minutes. This included playing the quiz (and/or sorting game) together. After the appointment, the child was able to use the MyPal application at a provided tablet at their home during two weeks. They could play the quiz, interact with the NAO avatar or enter activities on the timeline. Parents were allowed to play or

help the children if they pleased. In case of problems, they could call the technicians of the PAL team. Number of times and length of usage were up for the children and parents to decide as there was no minimum or maximum usage prescribed.

After two weeks, the children were contacted by a researcher by phone for a quick interview. The researcher would check on their goals and engagement of the child with the MyPal application through a semi-structured interview. There was also room for the child to give direct feedback on the application. The parents were asked to fill out an online questionnaire about their perceived usability and experience of the application. The next two weeks were then again attributed to home use of the MyPal application by the child.

During the final appointment, the child could play with the NAO robot while the parents fill out a questionnaire on user experience and feedback on the MyPal application. Then, the researcher engages the parents in an interview which follows through on some of the questionnaire questions to get a more thorough answer. Subsequently, the child was also interviewed. This interview mainly focused on self-management knowledge and skills and the user experience of both the NAO robot and the MyPal application. At the end of the appointment, the children and parents were thanked for their participation, while the children received a small memento of the study in the form of a picture of them with the robot.

3.4 Measures

This section describes how the variables of interest were measured or calculated during the experiment. First the measures of the possible explanatory variables are discussed and then the measures of the outcome variables for system usage and knowledge development are described.

3.4.1 Explanatory variables

Any demographic variables were registered through open questions. This was also the case for the socio-psychological variables: socio-economic status and whether the parents were divorced. The perceived diabetes threat of the children was measured through 28 written statements, which the child could then judge to be true or untrue for them on a scale of 0 (never true) to 4 (almost always true). The perceived threat was divided into five categories: symptoms (11 items), treatment barriers (4 items), treatment adherence problems (7 items), diabetes worries (3 items) and diabetes communication problems (3 items). Each category was scored by taking the average of the scores on all corresponding categorical items. A higher score indicated a higher level of problems or perceived threat in the specific category. The perceived threat scores for the parents were calculated in the exact same way. The perceived barriers in performing health promoting behaviour were measured through multiple choice questions. Both children and parents were asked to judge their diabetes knowledge and technology confidence on a scale of 1 (insufficient) to 5 (excellent). The user experience elements (comprehensibility, accessibility, visual appeal and motivation) were measured through a range of

statements which the participant could judge on a scale of 1(very untrue) until 5(very true). A higher score indicated a more positive user experience. Any scores on negative questions were inverted.

3.4.2 System usage - timeline

The system usage was examined through frequency and quality of use. The *frequency of use* is in this experiment referred to as the quantitative measure of all added content and the consistency with which this was added. The total amount of added content was calculated by adding all entered activities and event notes. A timeline entry was counted every time a participant added an event (activity glycemic value, meal or mood) to the timeline. Each timeline entry could contain one optional added note.

$$total\ added\ content = n_{activities} + n_{notes} \quad (1)$$

Measuring the consistency of added content was done through an equation which was previously used in a PAL pilot study. This way we could ensure compatibility and comparability of the data in the pilot, this experiment and coming experiments. The consistency with which the content was added is expressed in equation 2 (Ligthart, 2016). In the consistency equation, nonzero represents all days on which content was added and d_{ij} is the relative experiment day on which content was added.

$$consistency_{added\ content} = 1 / \frac{\sum_{j=1}^{n_{nonzero}} d_{ij} - d_{i(j-1)}}{n_{nonzero}} \quad (2)$$

The *quality of use* consisted of a qualitative analysis. For this, a categorizing scheme that was used in the first PAL pilot study by Ligthart (2016) was utilized. By classifying all added notes into one of the four following categories, the data was comparable in respect to the pilot study.

- (1) None. Factual description of the activity.
- (2) Minor. 1+ Minor indication of thoughts and feelings.
- (3) Moderate. 2+ Information about friends and/or family,
- (4) Major. 3+ Major indications of thoughts/feelings about the activity and/or friends and family.

3.4.3 System usage – Quiz

The quiz analysis entailed the same as in the timeline analysis: frequency and quality. For the frequency analysis of the quiz, every started play round was logged, as well as all answered questions. *Frequency* in this context entailed the quantitative measure of the number of played questions and consistency of started quiz rounds. The calculation of this consistency is similar to the timeline equation 2 as created by Ligthart (2016) in a former pilot study of the MyPal project. In equation 3, nonzero represents all days on which a quiz question was answered and d_{ij} is the relative experiment day on which a quiz question was answered.

$$consistency_{quiz\ questions} = 1 / \frac{\sum_{j=1}^{n_{nonzero}} d_{ij} - d_{i(j-1)}}{n_{nonzero}} \quad (3)$$

3.4.4 Knowledge development

The knowledge development for each participant was operationalized by calculating the percentage of correctly answered quiz questions per day. This percentage only includes diabetes related questions, and excluded any non-diabetes related quiz questions. By calculating the percentage each day, a trend may be identified over the course of the experiment.

3.5 Analysis

The analysis started with descriptive statistics on the demographic characteristics of both the children and their parents (table 4) to get a main overview of the sample. The explorative data analysis of the system usage was then divided in the use of the MyPal Quiz and the MyPal Timeline. For each of these MyPal parts, analysis of the system usage took place on two main aspects: frequency, and quality. The knowledge development was mainly explored through analysis of the quiz. This paragraph contains a description of the data preparation and analysis procedure for both the system use and knowledge development.

3.5.1 Data preparation

All data from the interviews and questionnaires were translated into English and entered in a digital format, after the final appointments took place. This was done in the same way for both the Italian and Dutch data to ensure that the data can be used together in the analysis. Data from the MyPal app on quiz data and timeline data were derived from different parts of the PAL database and restructured into one Microsoft Excel data file for analysis. Again, both data from Italian and Dutch participants were combined in one file.

For all interactions with the system, the corresponding date was converted to the relative experiment day of the child. This was done to prevent a distorted picture of frequency data as the experiment did not start and end on the same day for each child. This strategy allowed for a more realistic and clear view of the data relative to the experiment period of each child.

For the quiz, each started quiz round was stored with a time and date stamp. However, after close inspection of the data it became clear that a small part of the started rounds appeared invalid. This was recognized by the start of multiple quiz rounds by a single participant within the timeframe of zero to two seconds. Explanations for this invalid data may be that the app crashed or the child repeatedly tapped the 'start quiz' button. Also, some started rounds lasted for less than 10 seconds without any played quiz questions. As these are not actual played quiz round, any sequences of started quiz round with zero answered questions by the same participant within zero to two seconds were deleted. Started quiz rounds that lasted for less than 10 seconds without any answered questions were deleted for the same reason. Not only the started quiz rounds, but also all answered quiz questions were logged. Of all answered quiz questions, 3.8% of the answers were neither labelled as answered correct nor as incorrect. An explanation for this may be that these questions were not answered as the

child closed the app, or as the app crashed. The questions with answers labelled as ‘unknown’ were therefore deleted to ensure only actually played questions were analysed.

3.5.2 System usage – timeline data

In the MyPal Timeline, it was possible for the participants to enter an event within one of four categories: activity, glycaemic values, meal and mood. Within these categories, the participant was free to add notes besides the required data such as glucose value. For the analysis of the use of these Mypal Timeline functions, several aspects needed to be explored. The main aspects of this system usage analysis concerned: frequency (amount added content & consistency) and quality of use.

To get a global understanding of the timeline data, besides descriptive analysis, scatterplots and conditional boxplots were created for the total added content. Individual scatterplots were visually inspected for main trends. To identify different groups of system usage, an interaction plot was created for the mean added content per day and the consistency with which the participant added this content. A k-means analysis then allowed for the clustering of the participants on these variables. The quality of the system use was visualised by the four categories of Ligthart (2016), as described in chapter 3.4.2.

3.5.3 System usage – quiz data

For the frequency analysis of the quiz, every started play round was logged, as well as all answered questions. The quiz contained both diabetes related and general questions for the child to answer. The quiz analysis entailed the same as in the timeline analysis: frequency and quality.

First, a scatterplot and conditional boxplot were created for a first insight in the number of answered questions over the course of the experiment per participant. An interaction plot on the mean number of answered questions per day per participant and the consistency with which these were answered was made for identification of different user groups. A k-means analysis allowed for the grouping of users on these two levels.

3.5.4 Knowledge development

The knowledge development is operationalized by the percentage correctly answered diabetes questions per day for each participant. As with the timeline analysis individual and total scatterplots were created to get a better understanding of the data over time. A linear regression line was fitted for each participant, and combined into one graph. This showed the individual trends of quiz performance over time. A combined graph allowed for visual comparison among individuals and identify if a general trend line would be appropriate.

4. Results

4.1 System usage

The following section describes the results from the system use analysis, looking at both the timeline and quiz data consecutively. A general description of the usage is presented for both the timeline and quiz. Then, both the identified trends and user groups are described.

4.1.1 System usage –timeline

The timeline part of the MyPal application consisted of two parts: entries and notes. Entries entailed added activities, blood-glucose-values, meals and emotions while the notes were voluntary open descriptions that could be placed with a timeline entry. During the experiment period, a total of 1538 events were added in the MyPal application. On average, a participant filled in a total of 45.24 events ($SD=50.048$) during the whole experiment period with an average of two events per day ($M=1.84$, $SD=1.913$). The large standard deviations already indicated a large difference in the added timeline entries per participant. Indeed, while some participants never added a timeline entry, some participants added up to 188 entries. A total of 269 notes were added by the participating children with an average of 9 notes per participant ($M=9.37$, $SD=10.730$), and 0.4 notes per day per participant ($SD=0.398$). As with the timeline activity entries, the total number of added notes has a wide range over the participants. The minimum amount of total notes per participant was 0 while the maximum was 35 notes. In the further analysis, a combination of both timeline entries and notes was used.

Figure 7 shows the longitudinal development of the total added content per day per participant. Longitudinal individual scatterplots of total added content per day are situated in figure 8. Through close inspection of the individual scatterplots results a clear observation can be made. There appear to be three main trends over time in the timeline usage. Out of the 35 participants, the majority ($n=20$) shows a nearly flat line indicating a low or non-use over a longer period of time. This trend shows exceptions at the very beginning and/or end of the experiment. Any high use on the initial and last experiment day must be regarded carefully as participants were required to use the application in the hospital at the first and last experiment day and may thus not use by intrinsic motivation. Two participants show a consistent line above 0 entries, while the rest of the participants ($N=13$) show a clear decrease in use over time. The fact that the two main trends are consistent low usage and fast decreasing use prevents us from analysing any predictors for the system usage. As a large majority of the users do not use the system over a longer period of time, it is statistically not valid to investigate predictors for system use, as the data of this group is too sparse.

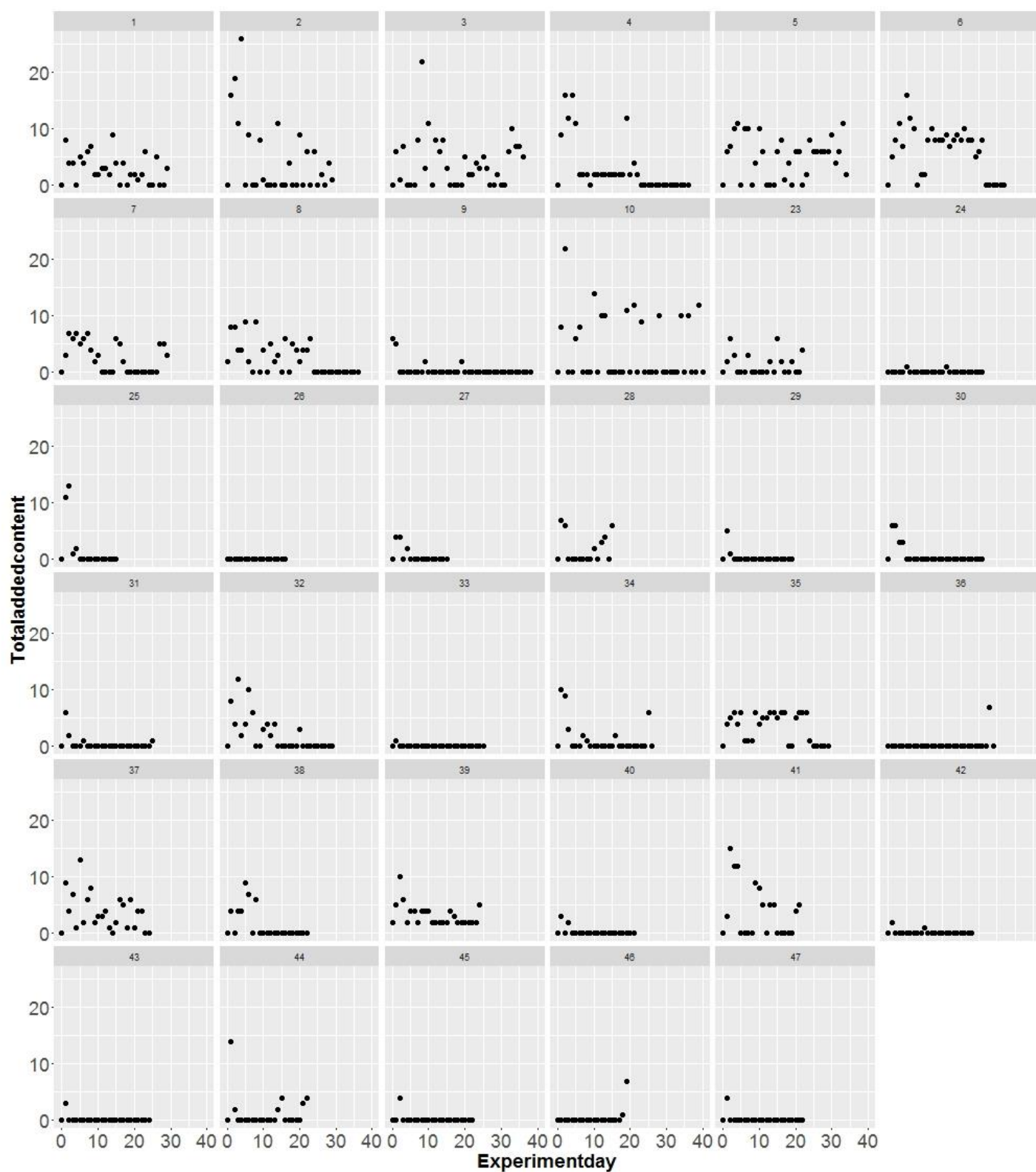


Figure 7. Individual scatterplots of the total added content in the MyPal timeline per day. Each block represents the amount of added timeline data per day of one child.

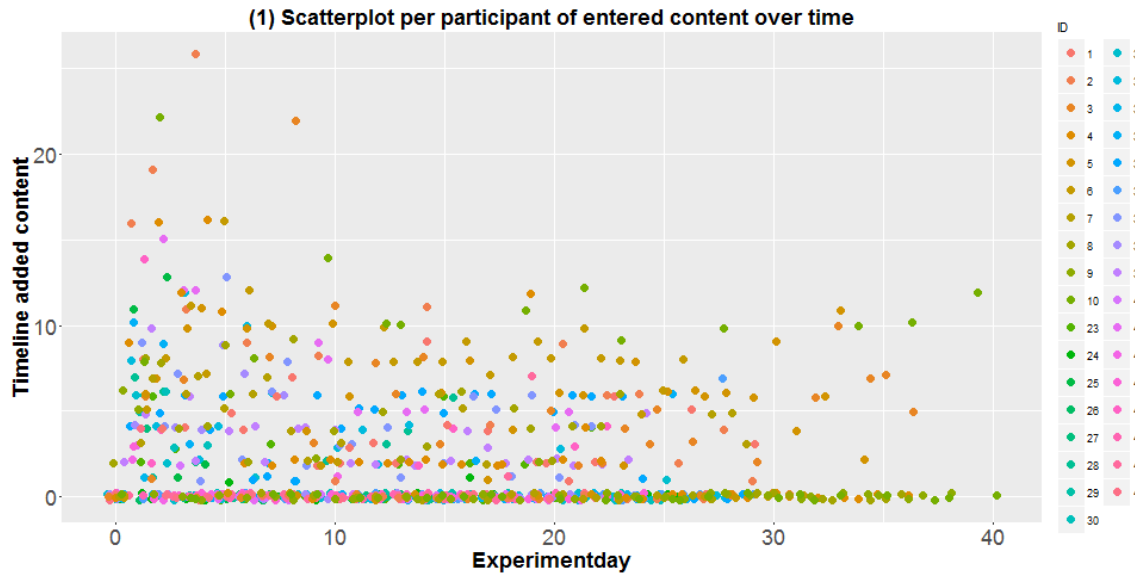


Figure 8. Scatterplot of the number of total added content on the timeline per participant for each experiment day.

Participant clustering. Consistency vs. total added content

In order to get a full overview of the participating children and their system use, it was not only necessary to look at the amount of content added but also at the consistency with which this happened. For every child, a consistency score was calculated for their added content. This score could range from 0 to 1, a higher score represented a higher consistency in the adding of content on the MyPal timeline. (The exact calculation and formula is presented in chapter 5.5.2). Figure 9 shows the consistency and mean added content per day for each participant. It was clear that the participants differ extremely on both parameters. There was no indication for a linear relationship between consistency and mean added content per day per participant ($p=0.435$), so being more consistent in the timeline would not imply a higher mean added content per day and vice versa.

For a further investigation of the system usage per participant, a cluster analysis was performed. This allowed us to categorize participants on both consistency and mean added content per day. The results of a hierarchical cluster analysis showed that an appropriate number of clusters was either 5 or 3. Looking at the scatterplot of consistency vs. mean total content per day, 5 clusters seemed like a more realistic division. However it is important to note that this grouping does not take the trend of usage into consideration.

A K-means cluster analysis of the standardized values indicated the standardized cluster means. Both variables were of significant influence on these means (both $p<0.001$). For the mean added content per day, the clusters differed significantly ($F(4,30)=32.859$, $p<0.001$). Further analysis shows that cluster 1 significantly differed from cluster 3 ($p<0.001$), 4 ($p<0.001$) and 5 ($p<0.001$) but not 2. Cluster 3 did not significantly differ from cluster 4 on mean added content per day. For consistency, cluster 1 differed significantly to all cluster ($p<0.001$). However, the consistency scores of cluster 4 did not differ from clusters 2 and 5.

The clusters could be described as following:

- Cluster 1: Overall low performance (very low)
- Cluster 2: Low content, high consistency (low)
- Cluster 3: Overall medium performance (medium)
- Cluster 4: Medium content, high consistency (high)
- Cluster 5: Overall high performance (very high)

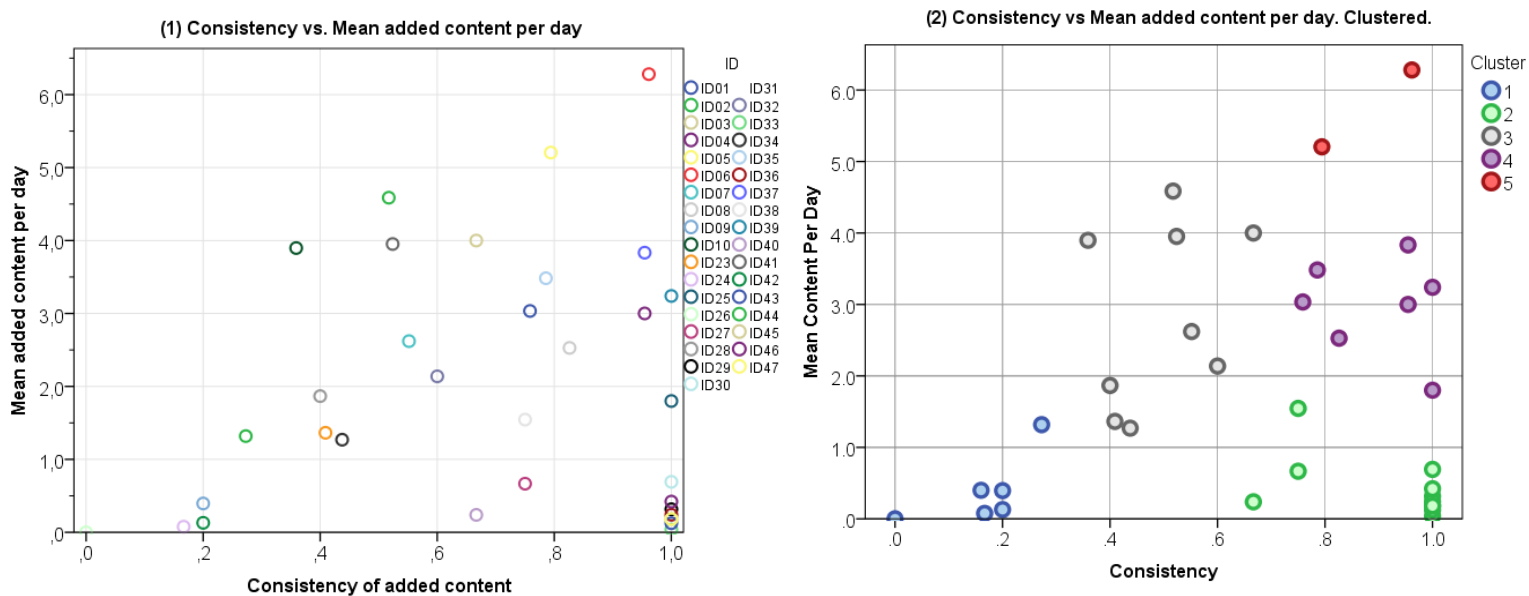


Figure 9. (1) Scatterplot of consistency vs. mean added content per day. Each dot represents a participant and their corresponding consistency score and mean added content per day. (2) Participants are clustered through a k-means analysis.

Quality of added content

Out of the 35 participating children, 30 children added between 1 and 34 notes on the timeline. All notes were categorized along the predefined scale by the amount of personal disclosure. The categories were defined as following:

- (1) None. Factual description of the activity.
- (2) Minor. 1+ Minor indication of thoughts and feelings.
- (3) Moderate. 2+ Information about friends and/or family,
- (4) Major. 3+ Major indications of thoughts/feelings about the activity and/or friends and family.

The added notes were skewed to the left on this index, with a high number of pure factual notes and a low number of elaborate personal disclosures (figure 10). As it is generally accepted that the willingness and ability to self-disclose changes as children grow older (Mandler, 2000), it was investigated whether this may be of influence on the differences in personal disclosure in the experiment. No significant relation was however found.

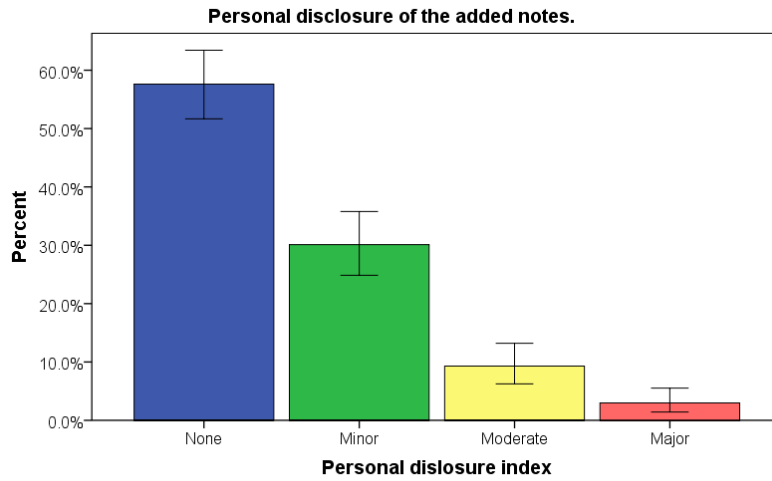


Figure 10. Distribution of the added notes in MyPal on their level of personal disclosure.

4.1.2 System usage – quiz

During the first year experiment, a total of 2477 questions were answered in the quiz by the participating children. The total number of answered questions differs greatly for each child with a minimum of 22 and maximum of 175 answered questions ($M=70.77$, $SD=33.939$). On average, a child answered 2.844 questions per day ($SD=1.480$). Figure 11 illustrates the longitudinal development of the number of answered quiz questions per participant per day, while figure 12 shows all individual scatterplots on the played questions per day. In figure 11, an overall decline in the number of played quiz questions per day per participant was observed. The rise after 34 days could be explained by the low number of remaining participants, as the experiment length differed for the participants.

The scatterplots in figure 12 indicated the same problem as the system usage in the timeline. A large part of the users showed a flat line of 0 usages per day, with small exceptions at the start and end of the experiment. These exceptions probably arose during the initial and final appointment in the hospital where the children asked to play some quiz questions. The users who did not show a flat line of zero played questions per day showed a clear sharp decreasing trend. Their use greatly declined and hit zero well before the end of the experiment. Again, this made it impossible to validly examine predictor variables as the data of actual users was too scarce and the group too homogeneous.

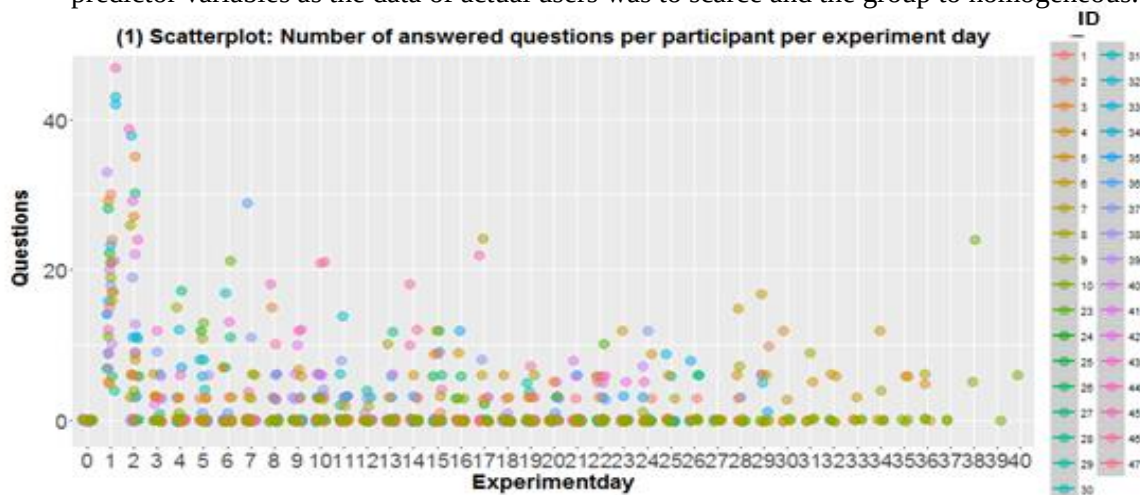


Figure 11. Scatterplot of the number of **answered questions** per participant for each experiment day. The experiment day is relative to each individual participant.

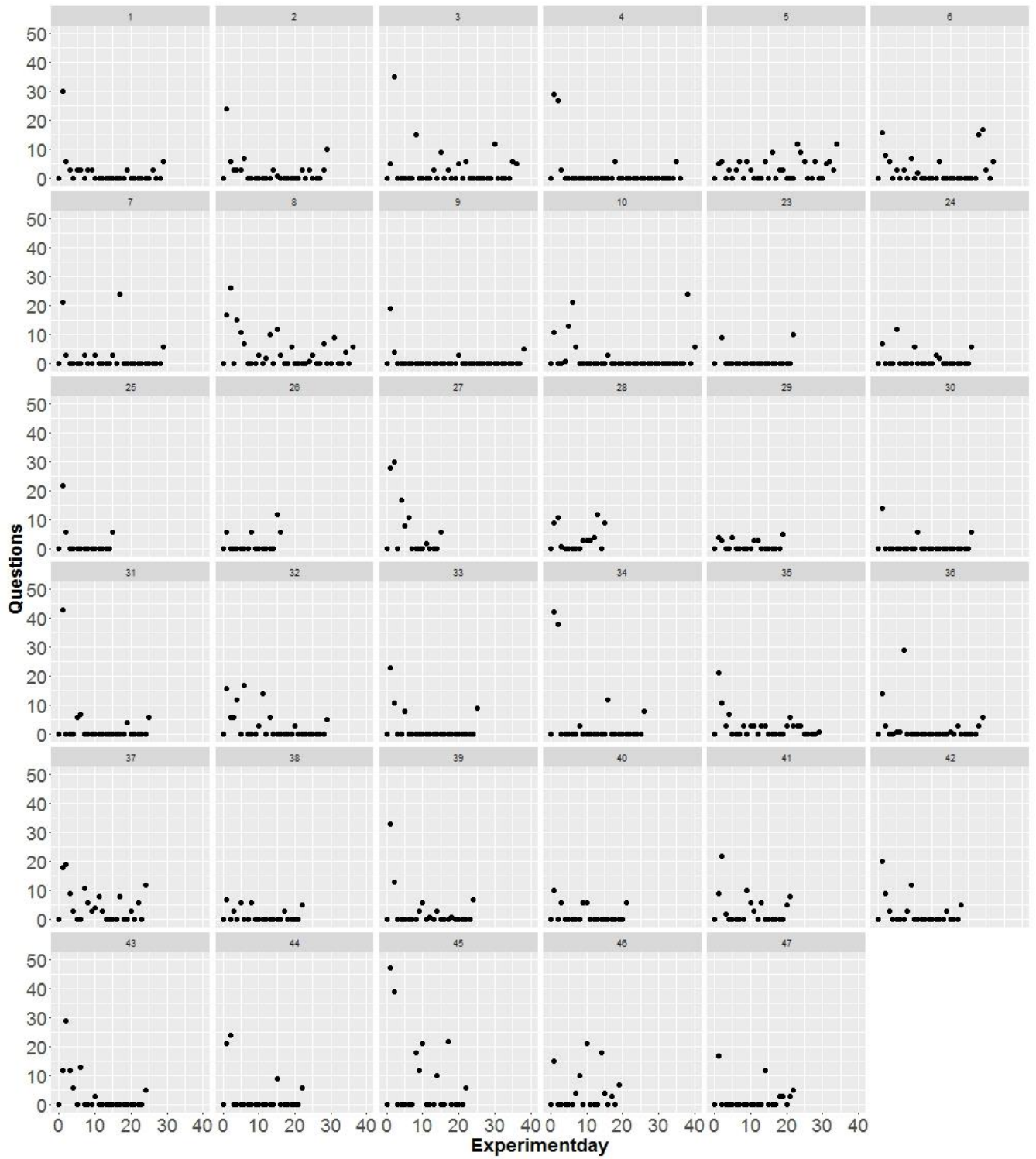


Figure 12. Individual scatterplots of the total played questions in the MyPal quiz per day. Each block represents the number of played quiz questions per day of one child.

Participant clustering. Consistency vs. played questions

Besides the mean number of answered questions, the consistency with which children answered these could tell us more about the system usage. Figure 13 (1) shows the position of each participant on both consistency and mean answered questions per experiment day. Just as with the added content, a hierarchical cluster analysis was performed to determine the number of clusters on consistency and played questions. The elbow method showed that either 3 or 5 clusters would be correct. A clustering of three groups was chosen as this appears more logical in this situation because the differences were not extremely large, and the statistical power was better for 3 clusters. The clustering of 3 groups for the quiz usage as opposed to 5 groups for the timeline usage may be due to the overall lower consistent use of the quiz, as no one exceeds a consistency above 0.6 in the quiz usage but in the timeline usage the consistency ranges all the way up to 1. A visual representation of this clustering is situated in figure 13 (2). The clusters can be described as following:

Cluster 1: Overall low performance (low)

Cluster 2: Medium content and consistency (medium)

Cluster 3: High mean questions, medium consistency (high)

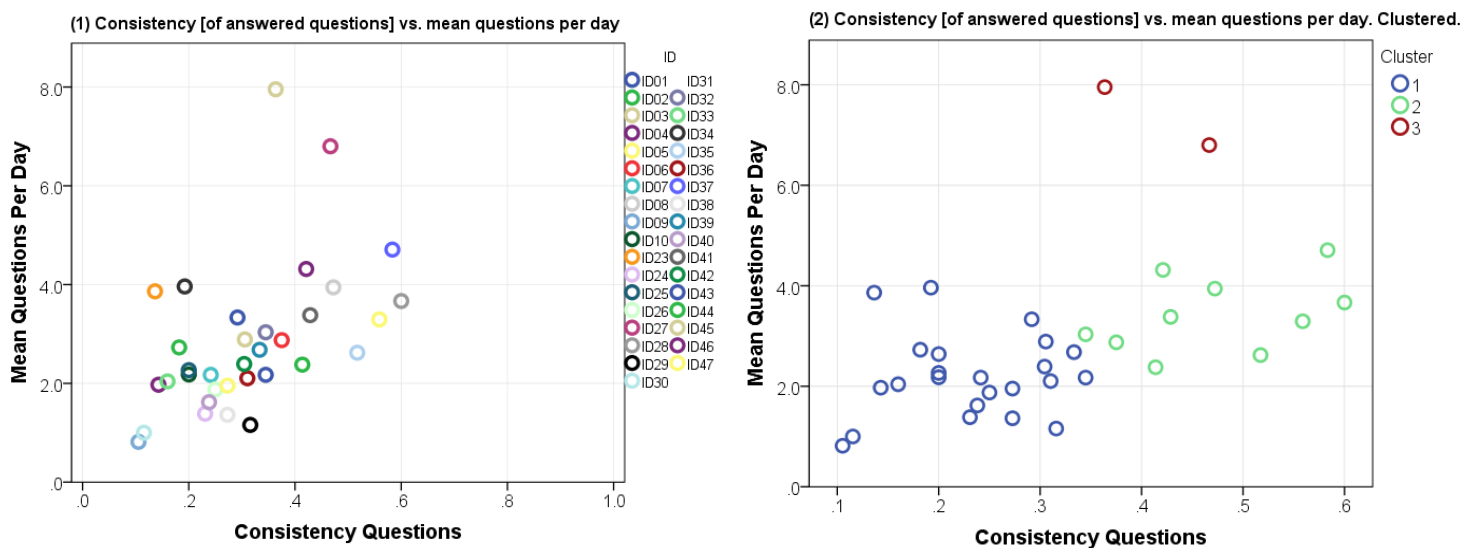


Figure 13. (1) Scatterplot of consistency of the answered questions vs. mean number of answered questions per day. Each dot represents a participant and their corresponding consistency score and mean answered questions per day. (2) Participants are clustered through a k-means analysis.

4.2 Knowledge development

As both the quiz and timeline sections of the system were not used by the majority of the participants, and the users who did use it showed a sharp decline, it was clear that only one conclusion with regard to knowledge development could be made. Children cannot (or very insignificantly) have benefited a positive diabetes knowledge development as a result of the system. Even if large changes in knowledge development were to be observed, these may not be directly attributed to usage of the system.

5. Discussion

The main goal of this thesis was twofold. The first goal was to determine what kinds of system usage and knowledge development patterns arise during a longer period of use of the current MyPal system. Secondly, it was important to take an explorative look at possible factors which may contribute to variation in system usage and knowledge development for the MyPal system. With results of these goals, the system may be adapted in its personalisation to account for these differences in system usage and knowledge development over time. Ultimately eliciting a more equal opportunity for improvement of self-management skills for different users.

In regard to the system usage development, it is quite clear in this study that the total interaction with the system decreases over time. This does not only hold for interaction with the timeline part of the MyPal application but also for the quiz. Three global trends in system usage were discovered. Most of participants only used the application a handful of times during the start and end of the study, but consistently refrain from further interaction throughout the remainder of the study. The second trend is that of users who start off with high usage performances but quickly diminish as time goes by. The last trend is shown by a very small number of users who use the application consistently over the whole course of the experiment and add about the same amount of content every time. This behaviour might be the most beneficial for learning as it takes full advantage of any possible effects of the system. The occurrence of these three trends leads to the following conclusion for the PAL system. The current PAL system is not personalizing its communication and tasks enough to spark the initial attention and interest of a large group of participants and remain the interest to avoid dropouts.

The large number of users that only use the system a handful of times and users who show a rapid decline in the PAL system usage make it both impossible and futile to examine what factors predict system usage and whether the system promotes diabetes knowledge development. In order to prevent the system from keeping to be ineffective in maintaining usage (and consequently knowledge development) it is important to consider what may cause the low usage in the current system. Looking at the most prominent game design and motivation research, it becomes fairly clear that the current MyPal system is missing several important motivational elements.

First of all, the fundamental theory of maintaining attention or ‘flow’ by Csikszentmihalyi (1990) is not implemented (yet). This theory states that there needs to be a balance between the experienced difficulty of the game by the player, and the skills he possesses to play the game. Researchers in the PAL team observed that as the personalisation of difficulty level is not yet implemented in the game, several MyPal players referred to the game as either way too difficult or far too easy. Also, the level of diabetes knowledge of the participating children ranged from very basic to extremely detailed. This lack of balance between skill and task may have a large effect on motivation and attention. A further observation was a limited amount of quiz content resulting in repetitive

questions. This further influenced the unbalanced difficulty. Second, multiple game studies like that of Dickey (2005), Hopson (2001) and Siang and Rao (2003) have emphasized that feedback on performance is a prominent element of successful and motivating games. This has not yet been noticeably addressed in the MyPal system. The timeline section does not provide feedback on entries at all, while the quiz only replies to answered questions sparsely with a small number of predetermined appraisals. Sometimes it was unclear for participants whether their answer was correct or false. Not only is this problematic for the motivation of the players. It also makes it difficult for the player to learn, as reinforcement and confirmation of actions has been generally accepted in learning psychology to foster learning in many learning theories like the fundamental operant conditioning theory of Skinner (Siang & Rao, 2003; Skinner, 1963).

Relating to missing feedback is also the presence of clear short- and long-term goals (Dickey, 2005). The MyPal system does have a global goal overview in regards to the quiz. However, these goals are not present during the quiz gameplay or in the timeline section. Most popular children's games include the continuous striving of small quick goals, and longer large goals. For example, the popular game Clash of Clans allows users to strive for short term goals like unlocking new buildings and long term goals such as defending your newly build town. These goals continuously update and are always in direct sight of the player. This combination of small and large goals that continuously update while showing the progress while playing is a prominent element in most popular games such as Angry Birds and Subway Surf. Shortly after the mass introduction of video games, research of Bowman (1982) and Provenzo (1991) already revealed the importance of a goal orientated approach in successful games that maintained children's attention. Making the MyPal system more goal oriented may therefore contribute to a higher use of the system.

More recent research in gaming shows another element which may contribute to motivation in educational learning systems. The MyPal system may need a more concise setting and backstory in order to promote relatedness and emotional investment of the players (Dickey, 2005). At this point, the avatar that appears in both the quiz, timeline (and its physical equivalent the NAO robot) does not have a personal story, nor is there a narrative that creates a setting for the common goal to achieve. In other words, why should the child answer the questions and fill in the timeline? Besides the real life practical goals of obtaining a better overview of their diabetes and more diabetes related knowledge, there is no in-game situation or story to justify the need of answering questions and entering timeline data for the children. Creating a convincing setting with characters and backstory that support the gameplay is therefore an important but missing game element (Adams, 2003; Rollings & Adams, 2003).

Some limitations need to be considered in regard to this study. The absence of families with a low socio-economic status, and high number of families with a high socio-economic status raises some concern about the representativeness of the participating families. An explanation for this is likely to be self-selection in accepting to participate in the study. Highly educated parents are more

likely to be familiar with (conducting) research through their studies, and might therefore be more trusting and willing to participate. Another limitation is of a more practical matter: the experiment duration differed greatly between Italy and the Netherlands.

Although no definitive user characteristics which may explain differences in system usage and knowledge development could be determined, it is important to understand the positive value of the found trends and user groups. The current PAL version without personalisation shows similar trends and groups as common diabetes management interventions like a physical nutrition and glycaemia diary (Gerken, 2011; Ligthart, 2016). This provides a data baseline for further testing of the system once alterations on game elements have been made and personalisation is improved. This study has tapped in a fairly new field of longitudinal evaluation. Through the acknowledgement of the importance of this type of study in product development, the human factors field has started to make more contributions of this type. This study however extends this development by using varying number of repeated measures participant. Also, the main part of the experiment took place at the home environments of the participants. This greatly increases the external validity of the analysis.

A number of recommendations regarding the improvement of system use and future PAL experiments can be made. A fair number of fundamental game design elements may need to be implemented in order to peak and maintain users' interest and system usage. These include the continuous visibility and updating of goals and direct feedback, a solid setting for the game with backstory to support the main goal and gameplay and a personalised balance between challenge and skills for the player. Recommended is to only further test the possible impact on diabetes knowledge development of the system and predictors for system usage and knowledge development after the implementation of the most fundamental game design elements. In order to obtain a valid progress on the systems performance during the PAL experiment cycles, it is advised to maintain the current longitudinal method with voluntary data points during the experiment and set measure points at the start and ending of the experiment. This allows for obtaining of critical information about the user characteristics and user experience, while also giving the opportunity to investigate voluntary system use. This voluntary approach should be kept as it gives a realistic view of system use, motivation and possible effects on diabetes knowledge and -management. Also, although the used measure for consistency by Ligthart (2016) may not have been validated by other studies it is advised to maintain this measure throughout the four year cycle in combination with the measure for total added content. Keeping these measures consistent will allow for tracking of user groups and whether for example the user group with low interaction becomes smaller when personalisation is implemented. While keeping the possible predictor concepts for system usage and knowledge development as measured in this study, it is urged to add more psychological characteristics as mentioned in chapter 2.5 as they may be addressed through the personalised system. These might include concepts like the 'big five' personality traits, mental ability or learning style which are linked to differences in intrinsic

motivation and learning differences (Becker et al., 2010; Jahanlou, Lotfizade & Karami, 2013; Lynch, Woelfl, Steele & Hanssen, 1998).

We conclude in regard to the question what patterns arise in the system usage of the PAL system, that the current PAL system is not yet capable of maintaining user's interest, resulting in a pattern of low use over time of the system. This pattern of low use over a longer period of time makes it impossible for the system to contribute to diabetes knowledge development, and consequently impossible to identify predictors for users who do use the system. This is thought to be largely due to the lack of personalisation and use of persuasive game design elements in the current system. It is advised to maintain the voluntary longitudinal use approach in combination with set appointments for the determination of user characteristics and –experiences. It is also recommended to include more psychological user characteristics in the following test cycles of the PAL project, as differences in these characteristics may very well be addressed by the personalised system. Overall the results of this study provide a great baseline for the improvement of MyPal and coming three test cycles of the PAL system.

References

- Adams, E. (2003, April 30). Defining the physical dimension of a game setting [Blog post]. Retrieved from http://www.gamasutra.com/view/feature/131273/the_designers_notebook_defining_.php
- Amato, P. R. (2005). The impact of family formation change on the cognitive, social, and emotional well-being of the next generation. *The future of children*, 15(2), 75–96. doi: 10.1353/foc.2005.0012
- Becker, M. H., Haefner, D. P., Kasl, S. V, Kirscht, J. P., Maiman, L. A, & Rosenstock, I. M. (1977). Selected psychosocial models and correlates of individual health-related behaviors. *Medical care*, 15(5), 27–46. doi: 10.1097/00005650-197705001-00005
- Boren, S., Fitzner, K., Panhalkar, P. S., & Specker, J. E. (2009). Costs and benefits associated with diabetes education: a review of the literature. *The diabetes educator*, 35(1), 72–96. doi:10.1177/0145721708326774
- Bowman, R. F. (1982). A “Pac-Man” theory of motivation: Tactile implications for classroom instruction. *Educational Technology*, 22(9), 14–17.
- Buck, N., Ermisch, J., & Jenkins, S. P. (1995). *Choosing a longitudinal survey design: the issues*. Essex, England: ESRC Research centre on micro-social change.
- Castañeda, J. A., Muñoz-Leiva, F., & Luque, T. (2007). Web Acceptance Model (WAM): Moderating effects of user experience. *Information & management*, 44(4), 384–396.
- Coates, V. E., & Boore, J. R. P. (1995). Knowledge and diabetes self-management. *Patient education and counseling*, 29(1), 99–108. doi: 10.1016/0738-3991(96)00938-X
- Cyr, D., Head, M., & Ivanov, A. (2006). Design aesthetics leading to m-loyalty in mobile commerce. *Information and management*, 43(8), 950–963. <http://doi.org/10.1016/j.im.2006.08.009>
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management science*, 35(8), 982–1003.
- Diabetes control and complications trial research group. (1993). The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *The new journal of medicine*, 328(25), 1800–1801.
- Gerken, J. (2011). *Longitudinal research in human- computer interaction* (Doctoral dissertation, University of Konstanz). Retrieved from: <https://kops.uni-konstanz.de/handle/123456789/17021>
- Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of educational psychology*, 82(3), 525.
- Grossman, T., Fitzmaurice, G., & Attar, R. (2009). A survey of software learnability. *Proceedings of the 27th international conference on human factors in computing systems*, 09, 649–658. doi: 10.1145/1518701.1518803
- Henkemans, O. A. B., Bierman, B. P., Janssen, J., Neerincx, M. A., Looije, R., van der Bosch, H., & van der Giessen, J. A. (2013). Using a robot to personalise health education for children with diabetes type 1: A pilot study. *Patient education and counseling*, 92(2), 174–181.

- Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.) *Handbook of parenting Volume 2: Biology and ecology of parenting* (pp. 231- 252). Mahwah, NJ: Lawrence Erlbaum Associates.
- Jahanlou, A. S., Lotfizade, M., & Karami, N. A. (2013). A new behavioral Model (health belief model combined with two fear models): Design, evaluation and path analysis of the role of variables in maintaining behavior. In O. O. Oguntibeju (Ed.), *Diabetes mellitus- insights and perspectives*(pp.297-312). doi: 10.5772/52966.
- Janssen, J. B., van der Wal, C. C., Neerincx, M. A., & Looije, R. (2011). Motivating children to learn arithmetic with an adaptive robot game. In Mutlu, B., Bartneck, C., Ham, J., Evers, V. & Kanda, T. (Eds.), *International Conference on Social Robotics* (pp. 153-162). Berlin, Germany: Springer.
- Jaser, S., Faulkner, M. S., Whittemore, R., Sangchoon, J., Murphy, K., DelamatA., & Grey, M. (2013). Coping, self-management, and adaptation in adolescents with type 1 diabetes. *Annals of behavioral medicine.*, 43(3), 311–319. doi:10.1007/s12160-012-9343-z.
- Kose, H., Yorganci, R., Algan, E. H., & Syrdal, D. S. (2012). Evaluation of the robot assisted sign language tutoring using video-based studies. *International journal of social robotics*, 4(3), 273–283. doi:10.1007/s12369-012-0142-2
- Kozulin, A., Gindis, B., Ageyev, V. S., & Miller, S. M. (2003). *Vygotsky's educational theory in cultural context*. New York, NY: Cambridge University Press.
- Leventhal, H., Safer, M. A., & Panagis, D. M. (1983). The impact of communications on the self-regulation of health beliefs, decisions, and behavior. *Health education & behavior*, 10(1), 3-29.
- Lorig, K. R., & Holman, H. R. (2003). Self-management education: history, definition, outcomes, and mechanisms. *Annals of behavioural medicine*, 26(1), 1–7.
- Lynch, T. G., Woelfl, N. N., Steele, D. J., & Hanssen, C. S. (1998). Learning style influences student examination performance. *The American journal of surgery*, 176(1), 62-66.
- Maathuis, I. (2014). *Technologies of compliance? Telecare technologies and self-management of chronic patients* (Master's thesis, University of Twente). Retrieved from doc.utwente.nl/94289/1/thesis_I_Maathuis.pdf
- Maehr, M.L. (1984). Meaning and motivation: Towards a theory of personal investment. *Motivation in education: student motivation*, 1, 115-144.
- Mandler, J. M. (2000). Perceptual and conceptual processes in infancy perceptual. *Journal of cognition and development*, 1(918975534), 3–36. doi:10.1207/S15327647JCD0101N
- Menard, S. (2002). *Longitudinal research*. Retrieved from: <http://goo.gl/bLv7xR>
- Menard, S. (Ed.). (2008). *Handbook of longitudinal research: design, measurement, and analysis*. Retrieved from: https://www.researchgate.net/profile/Karl_Mayer/publication/248823545_Retrospective_Longitudinal_Research_The_German_Life_History_Study/links/53d78bf60cf29265323ccf13.pdf

- Kozulin, A., Gindis, B., Ageyev, V.S & Miller, S.M. (2003). Sociocultural theory and education: students, teachers, and knowledge. *Vygotsky 'S educational theory in cultural context*, 1–11. doi: 10.1111/j.0007-1013.2004.397_7.x
- Mulcahy, K., Maryniuk, M., Peeples, M., Peyrot, M., Tomky, D., Weaver, T., & Yarborough, P. (2003). Diabetes self-management education core outcome measures. *The diabetes educator*, 29(5), 768–803.
- Ohkubo, Y., Kishikawa, H., Araki, E., Miyata, T., Isami, S., Motoyoshi, S., Kojima, Y., Furuyoshi, N. & Shichiri, M. (1995). Intensive insulin therapy prevents the progression of diabetic microvascular complications in Japanese patients with non-insulin-dependent diabetes mellitus: a randomized prospective 6-year study. *Diabetes Research and Clinical Practice*, 28(2):103–117.
- Pace, A. E., Ochoa-Vigo, K., Caliri, M. H. L., & Fernandes, A. P. M. (2006). Knowledge on diabetes mellitus in the self care process. *Revista latino-Americana de enfermagem*, 14(5), 728–734. doi:10.1590/S0104-11692006000500014
- Persell, S. D., Keating, N. L., Landrum, M. B., Landon, B. E., Ayanian, J. Z., Borbas, C., & Guadagnoli, E. (2004). Relationship of diabetes-specific knowledge to self-management activities, ambulatory preventive care, and metabolic outcomes. *Preventive medicine*, 39(4), 746–752. doi:10.1016/j.ypmed.2004.02.045
- Ployhart, R., & Vandenberg, R. (2010). Longitudinal research: the theory, design, and analysis of change. *Journal of management*, 36(1), 94–120. <http://doi.org/10.1177/0149206309352110>
- Provenzo, E. F. (1991). *Video kids: Making sense of nintendo*. Cambridge, MA: Harvard University Press
- R Core Team (2013). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3-900051-07-0, <http://www.R-project.org/>.
- RIVM (2014, april 10). *Zorg voor geest kost nog steeds het meest* [press release]. Retrieved from http://rivm.nl/Documenten_en_publicaties/Algemeen_Actueel/Nieuwsberichten/2013/Zorg_voor_geest_kost_nog_steedt_het_meest
- Rollings, A., & Adams, E. (2003). *Andrew Rollings and Ernest Adams on game design*. Indianapolis, IN: New Riders.
- RStudio Team (2015) RStudio: integrated development for R. RStudio, Version 0.99.491
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1), 68.
- Ryan, P., & Sawin, K. J. (2009). The individual and family self-management theory: background and perspectives on context, process, and outcomes. *Nursing Outlook*, 57(4), 217–225. doi:10.1016/j.outlook.2008.10.004
- Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. *The role of interest in learning and development*, 183–212.
- Schmettow, M. (2012). Sample size in usability studies. *Communications of the ACM*, 55(4), 64. doi:10.1145/2133806.2133824

- Shamsuddin, S., Yussof, H., Ismail, L., Hanapiah, F. A., Mohamed, S., Piah, H. A., & Zahari, N. I. (2012). Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO. *Signal processing and its applications*, 188–193. doi:10.1109/CSPA.2012.6194716
- Shrivastava, S., Shrivastava, P., & Ramasamy, J. (2013). Role of self-care in management of diabetes mellitus. *Journal of diabetes & metabolic disorders*, 12(1), 14. doi:10.1186/2251-6581-12-14
- Skinner, B. F. (1963). Operant behavior. *American Psychologist*, 18(8), 503.
- Streisand, R., Swift, E., Wickmark, T., Chen, R., & Holmes, C. S. (2005). Pediatric parenting stress among parents of children with type 1 diabetes: the role of self-efficacy, responsibility, and fear. *Journal of Pediatric Psychology*, 30(6), 513-521.
- UKPDS. (1998). Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes. *Lancet*, 352(9131), 837–853.
- Varni, J. W., Burwinkle, T. M., Jacobs, J. R., Gottschalk, M., Kaufman, F., & Jones, K. L. (2003). The PedsQL™ in type 1 and type 2 diabetes reliability and validity of the pediatric quality of life inventory™ generic core scales and type 1 diabetes module. *Diabetes care*, 26(3), 631-637.
- Varni, J. W., Limbers, C. a, & Burwinkle, T. M. (2007). Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL 4.0 Generic Core Scales. *Health and quality of life outcomes*, 5, 2. doi:10.1186/1477-7525-5-2
- Varni, J. W., Seid, M., & Kurtin, P. S. (2001). 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), 800–12. doi:10.1097/00005650-200108000-00006
- Van der Heijden, H. (2003). Factors influencing the usage of websites: the case of a generic portal in The Netherlands. *Information & management*, 40(6), 541-549.
- Van Gemert-Pijnen, J. E. W. C., Kelders, S. M., & Bohlmeijer, E. T. (2014). Understanding the usage of content in a mental health intervention for depression: An analysis of log data. *Journal of medical internet research*, 16(1). doi:10.2196/jmir.2991
- Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478. Retrieved from <http://www.jstor.org/stable/30036540>
- Willet, J. B. (1989). Some results on reliability for the longitudinal measurement of change: implications for the design of studies of individual growth. *Educational and psychological measurement*, 49, 587–602.
- Witte, K. (1998). *Theory-based interventions and evaluations of outreach efforts*. Retrieved from <https://nnlm.gov/archive/pnr/eval/witte.html>
- Zazworsky, D., Bolin, J. N., & Gaubeca, V. B. (Eds). (2006). *Handbook of diabetes management*. New York, NY: Springer Science & Business Media Inc.

Appendix A – Diabetes know and do goals

Table 1

Overview of the diabetes knowledge goals, and on what age they should be acquired according to the Dutch Children Diabetes Expertise Group. (Ages 6-11 years.)

Diabetes knowledge topic	Diabetes knowledge goal	Needs to be learned at the age of:		
		6-7 yrs.	8-9 yrs.	10-11 yrs.
General diabetes knowledge	Knows that diabetes is a chronic disease and will not go away	✓		
	Knows that diabetes is not contagious	✓		
	Knows that nobody knows why some get diabetes and others do not	✓		
	Knows that not he/she, nor anybody else, is to blame that he/she has diabetes	✓		
	Knows that if he/she uses insulin and eats regular, he/she can feel good and grow up as other children	✓		
	Why glucose measurement is needed		✓	
Blood glucose knowledge	Relation food, activity and insulin		✓	
	Meaning of blood glucose numbers	✓		
	Know to correct hypo with glucose	✓		
	Influence of insulin, carbs and exercise on glucose levels			✓
	When in between measuring is needed			✓
	When regular measurement is needed			✓
	Knows how to interpret glucose values and how to act accordingly			✓

Insulin knowledge	Knows that parents will indicate and control the amount of bolus.	✓		
	Knows that administration goes by pen/pump into leg/butt/belly		✓	
	Knows that the dosage of insulin is the responsibility of the parent.		✓	
	Knows the difference between different kinds of insulin.			✓
	Knows when they need which kind of insulin.			✓
	Knows how to preserve insulin.			✓
Food knowledge	How much to eat	✓		
	When to eat	✓		
	What to eat	✓		
	Knows to mention their diabetes if someone offers candy.	✓		
	Knows to consult an adult if candy is offered, to discuss whether to eat it immediately or take it home.	✓		
	Products with low/high carbs		✓	
	When and how much candy they can eat			✓
	How to vary with different foods and carbohydrates			✓
Psychical activity	Knows the relationship between insulin, nutrition and physical activity, take precautions with adults when exercising. (a1+a2)			✓
Hypo knowledge	Knows to take glucose pills or lemonade when he/she has low blood sugar	✓		
	Symptoms of hypo		✓	
	Knows the amount of glucose pills needed to get bs on			✓

normal level

Knows that he/she can't always depend on the symptoms
of a hypo and therefore should always, if possible, measure
his/her blood sugar.

✓

Knows at which value a blood sugar is too low

✓

Knows the possible causes of a low blood sugar can be

✓

Appendix B Outcome variables

Table 1. Outcome variables – Self-management determinants.

Self-management determinants	Sub-determinant	Measurable variable	Level 0 (6-7yrs) 1 (8-9yrs) 2 (10-11yrs)
Knowledge	General	• Knows that diabetes is a chronic disease and will not go away	• 0
		• Knows that diabetes is not contagious	• 0
		• Knows that nobody knows why some get diabetes and others do not	• 0
		• Knows that not he/she, nor anybody else, is to blame that he/she has diabetes	• 0
		• Why glucose measurement is needed	• 1
		• Knows that if he/she uses insulin and eats regular, he/she can feel good and grow up as other children	• 0
		• Relation food, activity and insulin	• 1
	Blood glucose	• Meaning of blood glucose numbers	• 0
		• Know to correct hypo with glucose	• 0
		• Influence of insulin, carbs and exercise on glucose levels	• 2
		• When in between measuring is needed	• 2
		• When regular measurement is needed	• 2
		• Knows how to interpret glucose values and how to act accordingly	• 2
	Insulin	• Administration by pen/pump into leg/butt/belly	• 1
		• Knows that parents will indicate and control the amount of bolus.	• 0
		• Knows that the dosage of insulin is the responsibility of the parent.	• 1
		• Knows the difference between different kinds of insulin.	• 2
		• Knows when they need which kind of insulin.	• 2
		• Knows how to preserve insulin.	• 2
			• 2
	Food	• How much to eat	• 0
		• When and what to eat	• 0
		• Knows to mention their diabetes if someone offers candy.	• 0
		• Knows to consult an adult if candy is offered, to discuss whether to eat it immediately or take it home.	• 0
		• Products with low/high carbs	• 1
		• When and how much candy they can eat	• 2
		• How to vary with different foods and carbohydrates	• 2
	Activity	• Knows the relationship between insulin, nutrition and physical activity, take precautions with adults when exercising.	• 2
	Hypo	• Symptoms of hypo	• 1
		• Knows to take glucose pills or lemonade when	• 0

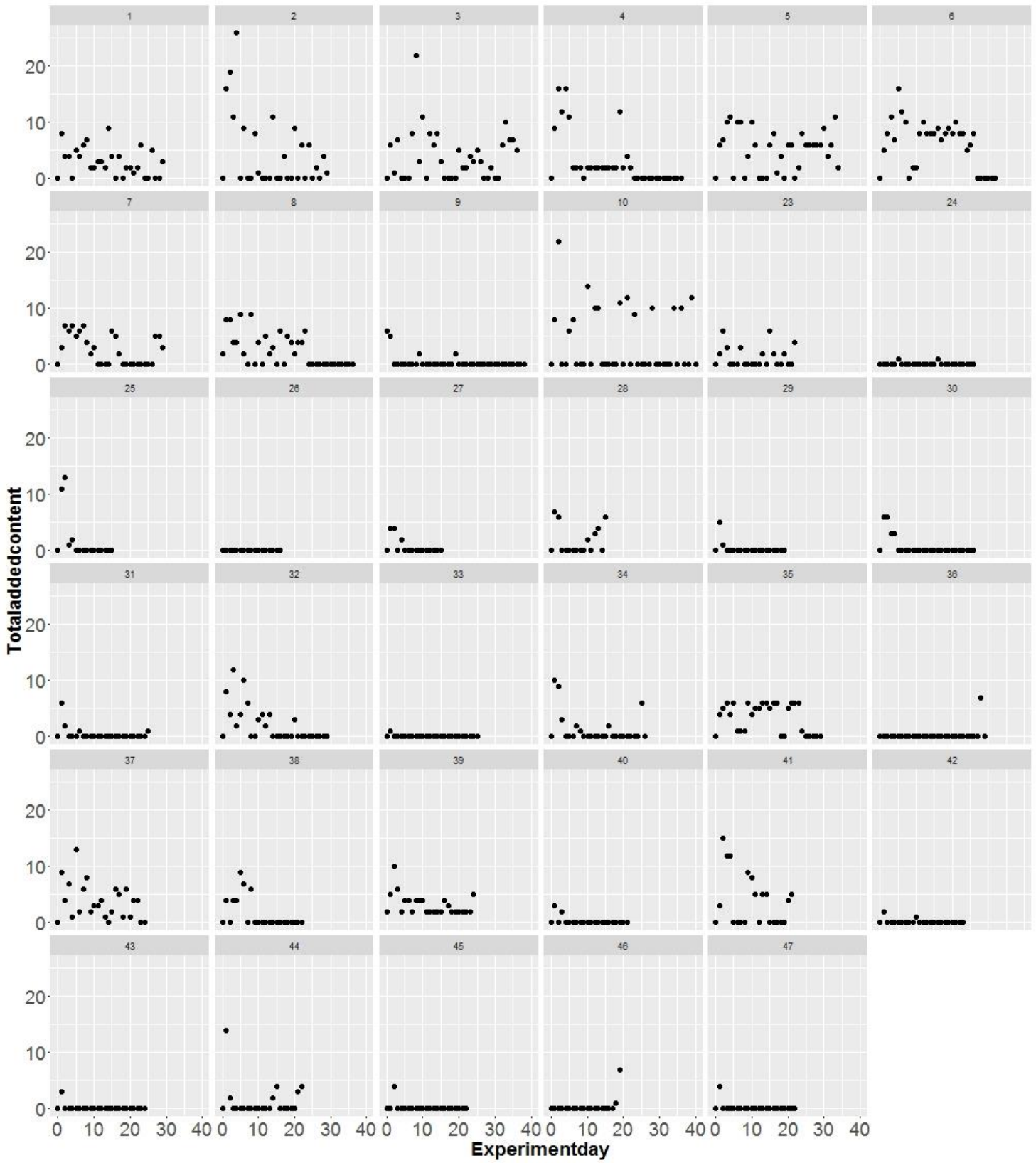
		he/she has low blood sugar	
		- Knows the amount of glucose pills needed to get bs on normal level - Knows that he/she can't always depend on the symptoms of a hypo and therefore should always, if possible, measure his/her blood sugar. - Knows at which value a blood sugar is too low - Knows the possible causes of a low blood sugar can be - Can correct for a low blood sugar by taking dextrose pills or lemonade	2 2 2 2 2
Self-efficacy	Treatment	- take blood glucose tests - take insulin shots - to exercise - keep track of carbohydrates or exchanges - to wear my id bracelet - carry a fast-acting carbohydrate - eat snacks	X
	Communication	- tell the doctors and nurses how I feel - ask the doctors and nurses questions - explain my illness to other people	X
Awareness	Risk of illness	Awareness of child of risk of illness	X
Attitude	Attitude towards diabetes	Attitude towards having to manage diabetes	X
Skills		- Can explain in own words that he/she has diabetes - Can explain in own words that his/her body needs insulin - Can explain in own words that insuline is being given with a insulin pen/pump - Can explain in own words that the insuline reaches his/her body through a insuline pen or pump in his/her belly or butt - Can explain in own words that diabetes is a chronic disease and will not go away - Can explain in own words that diabetes is not contagious - Can explain in own words that nobody knows why some get diabetes and others not - Can explain in own words that not he/she, nor anybody else, is to blame that he/she has diabetes - Can explain in own words why blood glucose needs to be pricked - Can explain what the numbers on glucose meter mean - Can explain what the possible causes of a low blood sugar can be - Can explain the influence of insulin, carbohydrates and exercise on their blood glucose - Can give a simple explanation of the relationship between nutrition, insulin and physical activity. - Prepare meter	0 0 0 1 1 1 1 1 1 0 2 2 1 0

	• Prick glucose • Interpret glucose levels and act • Can self-administer a bolus. • Can self-administer a insulin injection • Can help prepare and insert the new system in stages to learn this their selves. • Can correct for a low blood sugar by taking dextrose pills or lemonade • Alert adult if not feeling well • Can independently have a sleepover with others, whereas an adult is responsible for care taking. • Take precautions with parents for exercise	• 0 • 2 • 2 • 2 • 2 • 2 • 0 • 2 • 2
--	---	---

Table 2. Outcome variables - System usage.

System usage topic	System usage measurement
Frequency	• Times logged in (mean per day per participant) • Times activity/glycemia/meal added (mean per day per participant) • Times content added (mean per day per participant) • Times played quiz (mean per day per participant)
Duration	• Time spent per login • Time spent on Quiz • Time spent on Timeline
Quality	• Personal disclosure in added content • Correct diabetes questions

Appendix C – 1. Individual scatterplots, *added content* per day



Appendix C – 2. Individual scatterplots, *played questions* per day

