Acceptance of a Digital Paper-based Diabetes Diary

Understanding Precursors of Acceptance of Digitally Assisted Diabetes Care

André Calero Valdez and Martina Ziefle
Human-Computer Interaction Center, RWTH Aachen University, Campus Boulevard 57, 52074 Aachen, Germany

Keywords: Usability Testing, Novel Interaction Paradigms, Pen Computing, User Centered Design, Empirical Studies in Interaction Design.

Abstract: Diabetes prevalence has steadily been increasing over the last decades and is expected to continue in this trajectory. Diary keeping is considered the central part of any successful therapy. Since paper-based diary keeping is often poor and electronic diaries often challenging for elderly users, we developed Diabetto. Using focus groups and interviews we analyzed the requirements for Diabetto. It uses a LiveScribe Echo pen as a replacement for a regular pen in traditional paper-based diary keeping, while at the same time adds interactive therapy support and access to a nutrition database through pen-input and text-to-speech. By enhancing an already familiar process we managed to achieve a high acceptance of the device. In an experimental evaluation using an extended TAM model we analyzed the influence of user-diversity factors on the acceptance. Interestingly typical predictors of acceptance, such as efficiency during use, did not influence ease of use, indicating that the digital pen might not be perceived as computerized technology.

1 INTRODUCTION

With the rise of the "Internet of Things" (Atzori et al., 2010) and omni-present computing (Davis, 2002) almost every device will allow us to connect to the Internet and exchange data anytime and anywhere. The question is, whether this will allow overcoming some of the barriers people perceive in regard to medical technology, with their best interest in mind.

Although some may claim that usefulness (Scheermesser et al., 2008) is the most determining factor for pervasive medical technology, technology acceptance is neither static nor predictable using standardized technology acceptance modeling (Ziefle and Wilkowska, 2010). Illnesses influence the user over time, as does aging. Salience of technology may lead to increased perception of barriers in users, who already feel disconnected from modern technology, and reject the conscious use of modern technology. Nonetheless almost all people use ICT everyday, when they enter elevators, drive cars, or use their telephone. There's something about visibility of technology that influences technology acceptance. Acceptance of ubiquitous computing and medical assistance increases when technology is integrated into the familiar environment of the patient (Klack et al., 2011).

This article describes an approach to digital diabetes management system that tries to hide technology as a computer from the user and tries to fit computing into their already existing behavioral patterns. It puts the assistance into the pen that they use to fill out their diabetes diary.

1.1 Obstacles in Diabetes Therapy

Diabetes, if badly treated, causes a plethora of damages to the body (e.g. neural, cardial, etc.) resulting in extremely high treatment costs in comparison to a well-adjusted therapy (Hien and Böhm, 1997).

The biggest problems in diabetes therapy constitute therapy adherence and customization of the therapy especially in regard to designing an assistant (Hien and Böhm, 1997; Chen, 2010). In order to address the individual aspects of each patients illness, patients are required to keep a diabetes diary. In this diary vital parameters such as blood glucose levels are recorded along with insulin dosages, activity data, and food intake. The doctor must now understand the idiosyncratic reactions of the patients body to medication and lifestyle and adjust therapy accordingly.

Quality of therapy is highly dependent on the quality of the diary keeping (Hader et al., 2004). Users often refrain from keeping their diary in a precise fashion, but often fill their diary at the end of the day from memory (Stone et al., 2002). It is necessary for a diabetic to calculate the insulin doses using

Calero Valdez, A. and Ziefle, M.
Acceptance of a Digital Paper-based Diabetes Diary.
DOI: 10.5220/0006734800740086
In *Proceedings of the 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2018)*, pages 74-86
ISBN: 978-989-758-299-8

cross-multiplication, as the dosage is both dependent on current glucose levels and a correctional factor that is individually established for different times of the day. Also, a persisting diabetes can also lead to cognitive deficits (Yeung et al., 2009; Brands et al., 2005; Brands et al., 2007), which in turn hampers therapy.

2 RELATED WORK

In the following we review the literature on digital diabetes diaries and management applications. From the literature we see different forms of application and input methods. Generally electronic and paper based diaries have been investigated with differing results. We then introduce the the LiveScribe Echo Pen, which was used in an approach to combine the best of both worlds—a digital pen paper-based diabetes diary, that provides a multitude of the features that typical electronic diaries provide.

2.1 Diabetes Management

As early as 2003 Kerkenbush conducted a study on the efficacy of PDA-based diabetes diaries against paper-based diabetes diaries. Kerkenbush and Lasome (2003) identified the importance of PDA based solutions for diabetes care. They identified the importance of systematic and regular tracking in a diabetes diary for patients health as well as the need to teach potential users how to use a PDA based diary system.

The use of PDA based diabetes diaries is in particularly interesting as Stone et al. (2002) found a high unwillingness to paper diaries in many diabetes patients. This is not necessarily caused by a lack of therapy adherence but might have been caused by simple forgetting. An electronic device could have a reminder function, leading to a possible increase in diary keeping.

Broderick and Stone (2006) argue though, that comparability in regard to the device type is not yet possible. Reasons for compliance rates are not assessable through the conducted studies and could be conflicted by errors of self-reporting, motivation, and lack of actual medical compliance analysis.

Nonetheless, Forjuoh et al. (2008) investigated the medical impact of using PDA based diaries between high and low usage of the PDA software “Diabetes Pilot”. Over a course of six months high usage resulted in an significantly better increase in vital parameters than in low usage users. Still both groups showed improved of their vital parameters.

Duke et al. (2008) developed the Intelligent Diabetes Assistant (IDA). IDA uses machine learning algorithms to improve therapy by measuring lifestyle, nutrition and glucose levels. The measurements are instantly shared with a physician to allow shorter feedbacks loops on the diary evaluation. A similar solution has been developed by (Tani et al., 2010), which was also evaluated positively.

Burke et al. (2005) investigated the differences in monitoring food intake on a PDA after training hemodialysis and weight-loss patients in using a digital food diary. They found a higher loss of weight and higher self-monitoring adherence in users of a PDA based diary than users of paper-based diaries (Burke et al., 2011). If daily feedback was given through the PDA improvements were even larger.

When looking at PDAs (Arning and Ziefle ,2007) and mobile phones (Ziefle, 2002) acceptance of these devices might depend on prior experience and in particular be important when the users are part of the target demographic—-elderly users.

In a time when the technical challenges are no longer the largest barriers, it becomes important the regard the users emotive, hedonic (Alagöz et al., 2010) and cultural needs(Alagöz et al., 2011), especially when developing medical technology. Users might have a preference of using paper over PDAs or mobile phones for sheer haptic reasons (Gregory et al., 2008). When specifically looking at diabetes devices age shows a particular negative influence on usage performance (Calero Valdez et al., 2009). Additionally the screen size of the device increases usability in particular for older users (Calero Valdez et al., 2010b; Calero Valdez et al., 2011). Since some users are not familiar with the hierarchical menu structures of mobile phones or PDAs, other forms of navigation through a device should be considered (Calero Valdez et al., 2010a). This is particularly the case in mobile devices, where screen real-estate is scarce and can not be using excessively for user guidance.

2.2 Technology Acceptance

Since actual usage of devices depends largely on user acceptance it is necessary to incorporate acceptance modeling in the development process of such a device. For our case we have chosen to rely on the Technology Acceptance Model (TAM) by (Davis, 1989). TAM based the intention to use a system on two explaining variables. The perceived ease of use (PEU) and perceived usefulness (PU) determine largely the behavioral intention.

Other models such as UTAUT (Venkatesh et al., 2003), TAM2 (Venkatesh and Davis, 2000) and TAM3 (Venkatesh and Bala, 2008) were also considered but are not used due to the amount of factors that they introduce to the method. The TAM model has successfully been used to model the acceptance of medical technology in a health care setting (Venkatesh et al., 2011).

3 REQUIREMENT ANALYSIS

In order to develop a Livescribe-based diabetes-management-assistant that is both usable and accepted the user must be taken into account. To get insights into the life of a diabetic a focus group was conducted, the results of which are presented next.

3.1 Focus Group Results

The focus group was conducted in order to get first insights into using a mobile device for diabetics. In order to guarantee focus group success a list of core questions was initially established, which were then given to the participants. The questions were picked with five topics in mind.

The first topic is "diabetes in general" and questions were selected with special focus on therapy and diabetes management. The second topic is "mobile devices". The questions regarding this topic were selected to understand the impact of the level of technology expertise in using medical technology (i.e. mobile devices for diabetics). Special focus was put on the features of the used devices and their acceptance. The third topic is "diabetes diaries". Questions for this topic were chosen to improve understanding of how potential users keep their diary. The last two topics are usage motives and barriers. Questions for these topics were picked to assess future hurdles to prevent when implementing a device.

Five participants contributed to this setting and shared the following demographic information:

1. Type 1 diabetic, male 26 years, student of Computer Science, athlete
2. Type 1 diabetic, female, 14 years, pupil (accompanied by her mother)
3. Type 1 diabetic, female, 42 years, civil servant, Insulin-pump
4. Type 2 diabetic, male, 46 years, university degree in Mathematics
5. Type 2 diabetic, male, 63 years, retiree,

3.1.1 Results

In the focus group five participants were asked to elaborate on the experience with diabetes, mobile devices, diabetes management and possible motives and barriers for using a diabetes management assistant. The participants were selected from a proactive group of diabetics (i.e. visitors of a diabetes congress), they were self-selected, and had no known secondary disorders. They were mostly well-educated and all on insulin therapy.

Important findings that would need consideration from this focus group are (in order of occurrence):

1. Diabetes in general
   (a) Diabetes strongly influences the daily routine.
   (b) Physical activity needs special attention.
   (c) Willingness to adhere is high, but mood dependent.
   (d) Domain knowledge of diabetes is high, but varies with severity of the disease.
   (e) Informedness varies with understanding of diabetes.
   (f) Different parameters are known differently well depending on (perceived) domain knowledge.
   (g) Important parameters are blood glucose level, HbA_1c, body-fat percentage & blood pressure.
2. Mobile devices
   (a) Insulin pumps are perceived as very helpful, but sometimes as foreign.
   (b) General satisfaction with devices is high.
   (c) Advanced functions are not used.
3. Diabetes diaries
   (a) Paper-based diary keeping is usual but perceived as cumbersome.
   (b) Shortcuts are used to simplify logging.
4. Motives and barriers
   (a) Therapy is cumbersome and device implementation a necessity.
   (b) Operating expense (i.e. time) must be minimal.
   (c) Some people prefer paper-based diaries over digital ones.
   (d) Data safety is important.
   (e) Low level of perceived usefulness might be a barrier.
   (f) Benefits for health are imaginable if the device integrates well into the daily routine.
   (g) Insulin-dosing calculation would be a helpful feature.
   (h) Deal-breakers are bad usability and public visibility.
   (i) Mobile phone integration is only attractive for some users.

4 AN INTERACTIVE ANOTO-PAPER BASED DIABETES DIARY

From the requirements gathered, we conceptualized a diabetes diary using the *Livescribe Echo* digital pen. The idea behind the software named *Diabetto* was to let users keep their paper-based diaries but offer interactive support using the *Livescribe Echo*. This pen uses an infrared-camera to detect what is being written with the pen and has a microphone and loudspeaker included. Both can be used from software on the pen. Usage feels like using a normal pen on regular paper, but with digital support.

All versions of the *Livescribe Echo* are based on an ARM9 processor. The pen is larger (158x19mm) and heavier (36g) than an average pen. Data transfer happens using a micro-USB connector. The pen also has a 96x18 pixel OLED-Display, an integrated microphone and and integrated loudspeaker. It also has a headphone jack, which further allows recording of stereo-data.

The software Diabetto does not focus on the evaluation of generic diabetes diary keeping, but on assisting the user by verbal feedback on diary entries. Furthermore it should allow self-defined shortcuts the simplify diary keeping in order to reduce barriers. The following section introduces how the the Diabetto software was developed, which is then evaluated in the next section.

4.1 Diabetto – A Digital Pen based Diabetes Diary

The first essential part of Diabetto is the paper-based diary. The diary consists of three parts. The first part is the actual diary pages that the users uses to log his diary entries. The second part is a set of pages that contain a database of food items that the user can modify as a shortcut to regularly consumed meals. The last part is a set of general abbreviations that allows the user to record and input with a self-defined abbreviation.

4.1.1 The Diary Pages

The first part is the actual diary (see Figure 1). The diary itself consists of lined paged that are empty. Each page refers to a certain day. Users can write on these lines indicating what they have done on a particular day.

When a line starts with a *time-information* (e.g. 9:00) the diary assumes this to be the time of the entry (used when logging forgotten information). When no time is given the current system time is used as the time of entry. Time-information is recognized by a regular expression detecting any four or five set of numerical characters with a colon in between.

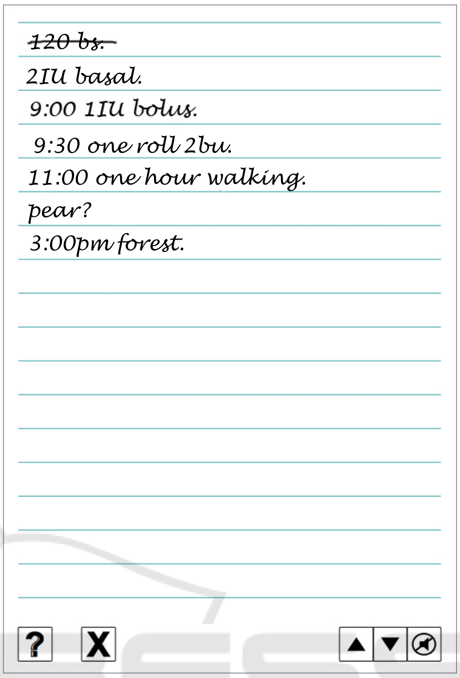

Figure 1: A sample diary page. The diary page contains a list of entries, one per line. Entries that are striked-through are deleted entries. Entries that have no time are recorded using the system time.

Blood glucose readings are detected, when a line ends with the characters “bs”. The pen then tries to recognize the number before these characters and interprets them and gives advice accordingly. When the numbers are on the low end of health blood glucose levels an auditory warning is issued: “Warning, your glucose levels are low. Please verify your measurement or react to this state.” When levels were too high an analogous warning was issued. This message also included a suggestion of how much insulin to administer in order to reduce the amount of blood glucose level to a healthy level. In the case of normal measurements a confirmatory message containing the numerical value written in the diary was played.

Whenever the two characters “BU” are found after numerical characters a *food consumption* entry is assumed. The amount of BU-units is played as auditory feedback and then recorded.

Insulin dosages were recognized when the characters “IU” were found after numerical characters. Auditory feedback on the amount of administered international units of insulin was played.

When a word is written that ends with a question mark, the pen looks into his *nutrition facts* or *abbreviation database*. If an entry in the nutrition facts exists, the amount of BU-units is reported to the user using audio feedback. If the same word is written with a dot at the end, the BU-units are recorded as a diary entry. Found abbreviations are "unfolded" and the reinterpreted.

The rest of the line, which is not interpreted, is used as a *commentary* to the respective diary entry.

If no interpretable data is contained in a single row, the entry is recorded as an *activity* (e.g. "one hour walking").

Entries can be *deleted* either by striking them through or by clicking the X-button on the bottom of the page immediately after input.

Furthermore every page has a *help* button (showing a "?") that starts a short audio-tutorial to the pen, as well as three buttons to *adjust the volume* of the pen (i.e. volume up/down/mute).

The Nutrition Facts Pages

The second part of the diary contains a table with columns (see Figure 2). The left column is used to enter the name of a meal or a food item that should be stored in the diary. The right column (in the same row) takes the amount of BU-units that are associated with this entry. These entries can be used in the diary as a shortcut. Furthermore deletion of entries is possible by striking them through.

The Abbreviations Pages

The third and last part of the diary is very similar to the second part. It also has two columns. The left column takes the abbreviation, while the right column takes the expanded full text. Whenever a user writes a word from the left column of the abbreviations pages the text of right column of the same row is virtually inserted and interpreted. This allows for complex abbreviations like "9:00 Breakfast, 2 cheese rolls, 6BU." that may occur regularly in a users life.

4.2 The Applications

Two applications are part of Diabetto. The first application is the penlet that runs within the pen to provide the functionality of the diary. Since any of the features that were described in the previous section are actually a functionality of the pen and not the paper, no detailed further information about the penlet application is necessary. This already indicates that the pen itself is not considered as the interface, but the paper is, although technically the opposite applies. The

Figure 2: Example page of the Diabetto diary showing a nutrition facts sheet. A user may write food items he consumes often in the left column and its BU-units in the right column.

application automatically starts, when the pen is used on Diabetto-diary paper. The second application is desktop-application to synchronize data. This application is not part of our evaluation here.

5 QUESTIONS ADDRESSED

Classical technology acceptance using a TAM model requires to assess the usefulness and ease of use of a technology to estimate the behavioral intention to use the technology. In our case we wanted further information to be available for analysis.

We were interested in finding out whether the auditory-feedback and the writing-based input were usable and helpful to the patients. We assumed that the integration into the already known process of keeping a paper-based diary would be easier than teaching a PDA or mobile phone application. Additionally we were interested in seeing how age, expertise, domain knowledge, and computer self-efficacy influenced the evaluation of the device, as previous studies had shown a strong influence of these factors (Calero Valdez et al., 2009; Calero Valdez et al., 2011) in PDA based devices. In particular the effectiveness of use has shown to be good predictor of perceived ease of use in previous exploratory studies

(Calero Valdez et al., 2009). Under the assumption that a pen might be less likely seen as a technological device, influences off effectiveness during use could differ from previous studies.

6 EXPERIMENTAL EVALUATION

In order to understand, whether possible users would accept a digital paper-based diary and how well they could use such a device, a user study was conducted. Focus of the study was to let users use the pen in a scenario-based task, assess their performance and furthermore assess their acceptance of the device.

6.1 Method

The experiments were conducted at the RWTH Aachen in a laboratory environment. Additional experiments were conducted at the office of Dr. Lätzsch in Aachen in order to remove some effect of the self-selection that occurs when participants have to travel to the experimental destination. Participants were directly contacted in the physician's office or for the control group taken from the immediate social network of the author.

6.2 Experimental Procedure

The experiment was conducted in several steps. First participants were informed about the experiment and that their data was being recorded. In order to measure performance, videos of the experiments were also recorded. The participant was informed that the interaction would be recorded on video, not exposing the face (only the diary and the hand were in frame). Then participants had to answer a questionnaire.

In order to get accustomed to the digital pen, first four tasks were instructed, in which the participant would also receive assistance from the experimenter. These tasks could also be repeated as often as desired. Once a participant stated that he understood the concept behind the digital pen, participants were instructed to perform three performance tasks. After finishing the task sets, a post-experimental questionnaire had to be answered. After this last questionnaire participants were instructed about how they could contact the experimenter in case they wanted any of the data deleted at any time afterwards.

6.3 The Questionnaire

The questionnaire was separated in two parts. Participants were asked demographic data such as age, gender and work. Furthermore handedness was assessed because handedness could influence accuracy of the text recognition.

Additionally health status was assessed. This means type of diabetes, diabetes duration and type of therapy were assessed. Technical expertise was assessed with both household and expertise in computerized technology (ECT) as well as expertise in health technology (EHT). For descriptive purposes coping scales and domain knowledge scales were assessed (i.e. DP, DCI, TDC, IFG, IFO).

Lastly the computer self-efficacy (CSE) (Beier, 2004) and behavioral intention to use the diary were assessed. Computer self-efficacy was measured using eight items, while behavioral intention was measured using three items.

6.4 The Task Set

In order to assure equal conditions between participants a fixed task set was designed, which included all features of the Diabetto diary. Instructions were available in paper-based format (i.e. printed between the pre- and post-questionnaire). Questions to the experimenter were allowed during the first four tasks. The following four tasks were used as introductory tasks:

1. **Opening the Help Function.** Users were asked to activate the help function. This task required them to understand that they could "click" on buttons on the paper, which activates behavior of the pen. Furthermore participants were encouraged to set the audio level output to a comfortable level.

2. **Write a Simple Sentence.** Users were asked to write in a line of the diary "I am taking part in an experiment". To this entry the pen would give audio feedback: "An activity for <date of experiment> was logged". This introduced the participant to the idea that the diary recognizes input as a diary and stores the information.

3. **Log a Glucose Measurement.** Users were asked to add an entry to the diary saying that they had a blood glucose reading of 100mg/dl. The pen would give audio feedback and comment on the healthy value of 100. This introduced participants to the idea that the diary could be used to record glucose measurements, and that the pen would comment on the value.

4. **Log Insulin Dosing.** Users were asked to log two entries of insulin dosing. This introduced the participants to how insulin dosing can be logged in Diabetto.

After completing the introductory tasks, the participant was informed that the following tasks were the actual experiment. Then the performance task sets were presented. The task set consisted of the following complex tasks:

5. **Log a Whole Day.** The user was asked to record typical measurements and doses and activities of a whole day in the diary.
6. **Usage of Nutrition Facts.** The users were asked to retrieve the BU-unit value for an already stored food item from the database. Furthermore they were asked to create their own entry and retrieve that value as well.
7. **Usage of Abbreviations.** Users were asked to create an abbreviation and use it in the diary.

After the completion of these tasks, the pen was connected to a Laptop and the Desktop application was shown, as well as a printout of the diary entries was handed to the participants.

In some cases participants strayed from doing the actual task, because the data given in the task did not coincide with their personal life. In those cases, participants were gently reminded and asked whether they could, after entering the own personal data, complete the actual task. This behavior deemed successful because participant's curiosity, how the pen behaved in their own lives was satisfied, but nonetheless all tasks were completed.

In cases were the pen did not recognize readings, and when it was obvious that recognition failure was caused by the limited dictionary provided by the prototype, participants were reminded of the prototypic nature of the experiment, and asked to retry the task using the words given in the task description. The amount of additional help was noted for additional evaluation.

6.5 Perfomance Measurment

In order to analyze performance video data is analyzed using a form sheet. The sheet has predefined actions that when recognized in the video recording are counted for evaluation. Approx. 14 hours of video data was analyzed for the data.

Items that are counted are the participant's security in using the device (i.e. "asks for help" and "corrects mistakes on his own") and user satisfaction (i.e. "positive mentions of the device", "negative mentions of the device", "laughter"). Additionally it is recorded whether the participant ignores visual and auditory feedback of the device.

Furthermore it is recorded when technical problems (e.g. OCR fails) occur, what participants say about the device and a general impression of the experimenter of the participant.

Timing is measured by entering the time-codec of the frame when a user clearly starts reading a task, when he starts writing, and when he finishes writing (i.e. the pen is lifted). Furthermore it is noted whether the task was effectively completed.

Time on task was calculated as the difference of finishing time codec from start time codec.

7 HYPOTHESES

As independent variables demography and health values were used. Age, gender and diabetes related factors were assessed. Furthermore expertise with technology and domain knowledge, as well as computer self-efficacy were assessed. As intermediary variables task performance was measures in regard to effectiveness and efficiency. Lastly as dependent variables perceived ease of use and perceived usefulness were assessed in order to assess influence on behavioral intention (see Figure 3).

We assume that age influences computer self-efficacy negatively (H_1), similar as age should influence expertise with technology negatively (H_2). Both these factors, as shown in previous experiments, should influence the efficiency during the experiment (H_3). The effectiveness of using a diabetes diary in general should be positively influence (H_4) by either knowing more about diabetes or being more experienced in using diabetes diaries (i.e. longer diabetes duration). Efficiency and Effectiveness should both influence the perception of ease using the diary (H_5), while perceived usefulness should also depend on experience, our previous experiments have not confirmed this finding (Calero Valdez et al., 2009). Therefore this is not a formulated hypothesis. Furthermore we expect to see a positive influence on behavioral intention using both PEU and PU (H_6).

8 DESCRIPTION OF THE SAMPLE

A set of 27 participants took part in the study. The average age was 38 years ($SD = 15.8$, 22-73 years). Eleven participants were male (41%), and 16 were female. Eight participants were diabetics. Three persons were left-handed. Healthy participants were recruited from announcements in a local newspaper, and diabetics via the social network of diabetes patients.

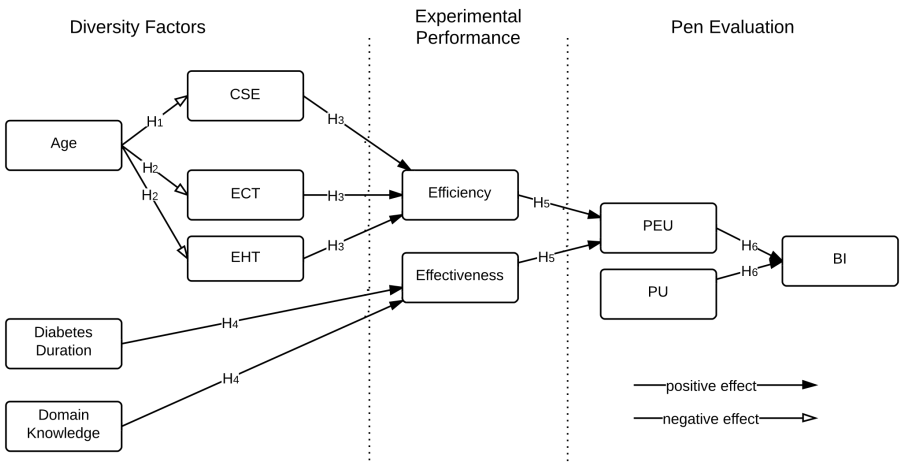

Figure 3: Visual representation of hypotheses.

The participants were split in three groups (tertial-split) according to their age. The young age group showed a mean age of 25 ($SD = 2.7$, $n = 10$), the medium age group showed a mean age of 34 years ($SD = 5.7$, $n = 10$), and the old age group showed a mean age of 62 ($SD = 5.7$, $n = 7$).

Diabetics showed a mean age of 49 years ($SD = 17.7$), while non-diabetics only showed a mean age of 33 years ($SD = 12.7$). Mann-Whitney-U testing reveals that diabetics were indeed older than non-diabetics ($U = 35.5$, $z = -2.157$, $p < .05$, $r = -.42$).

9 RESULTS

In the following sections results are presented. Results are analyzed using bi-variate correlations, univariate analysis of variance (ANOVA) and multiple linear regression analysis.

9.1 Descriptive Results

In order to get a broad overview into the results, first descriptive statistics for the three age groups are presented. When a influence seem plausible, correlation data is also presented. First we look into perception of aging and technical expertise. Then diabetes scales are presented. Lastly the various performance metrics, as well as a pen evaluation is presented.

9.1.1 Technical Expertise

Technical expertise seems to be highly and equally well distributed for the ECT and EHT scale (see Figure 4). Only expertise with medical technology is very low for all age groups. Computer self-efficacy seems to decrease with age. Overall the sample seems to be highly tech-savvy with the typical decrease in self-efficacy with age.

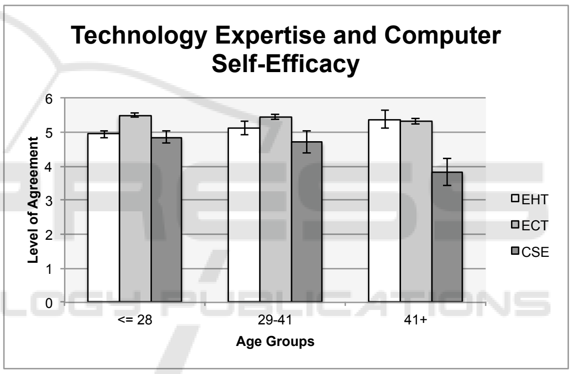

Figure 4: Technical expertise and efficacy scale marginal means over age. EHT=Expertise with household technology, ECT= Expertise with computerized technology, CSE=Computer self-efficacy. Error bars denote standard error.

Correlational analysis confirms that age correlates significantly with computer self-efficacy (CSE). ($r = .-457$, $p < .05$). Interestingly the seemingly equal technology expertise with computerized technology (ECT) correlates highly significantly with the CSE ($r = .577$, $p < .01$).

9.1.2 Diabetes: Coping Scales and Domain Knowledge

When looking at coping scales no differentiation between age groups can be made, because of the small sample of diabetics ($n = 8$). Nonetheless diabetics seem to perceive their diabetes to be as only lightly pervasive (DP: $M = 2.4$, $SD = 1.3$). The diabetics seem to have a high desire for information and control (DCI: $M = 4.7$, $SD = 0.7$) as well as trust in their doctor's competence (TDC: $M = 4.1$, $SD = 0.5$).

Diabetics were very well informed about diabetes and rather fairly informed about their obesity (IFG: $M = 3.9$, $SD = 0.3$, IFO: $M = 2.5$, $SD = 0.2$). This is not unusual because all diabetics were type-1 diabetics.

Correlational analysis showed that both IFO and IFG did not correlate with any other measures, except with themselves ($r = .824$, $p < .05$). The duration of diabetes only interacted also interacted with the informedness of obesity related factors ($r = -.793$, $p < .05$). The longer a participant has had diabetes (early onset) the less he was informed about obesity, which is externally valid, since obesity related information is of more interest to late-onset diabetes. Furthermore we found that diabetes pervasiveness did correlate with perceived cognitive deficits ($r = .812$, $p < .05$). All other measures did not correlate with coping style scales.

Further analyses of these variables are not performed.

9.1.3 Task Performance

Task performance was determined by both task success rate for effectiveness and time on task as efficiency. Effectiveness seems to be very high in the younger age group over all tasks (see Figure 5). The older age group seems in contrast to have more problems with the later tasks. The only tasks that seems to make no difference between age groups are task one and task five. Correlational analysis shows that tasks three ($r = -.473$, $p < .05$) and four ($r = -.450$, $p < .05$) do correlate significantly with age. The older a user is the less effective he is at completing these tasks.

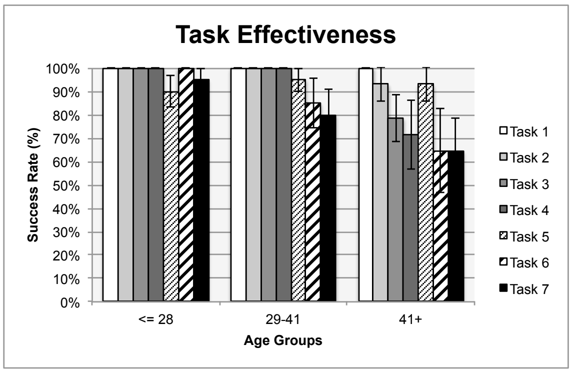

Figure 5: Marginal means of effectiveness for all tasks in percent over age. Error bars denote standard error.

When looking at efficiency for individual task (see Figure 6) individual task length becomes visible. Task five seems to take the most time from all participants, while task two seem to take the least.

As only the last three tasks are measured as a performance experiment, reliability of these three task as a scale was assess with a Cronbach's $\alpha = .734$. This indicates that the last three tasks can be used as one scale. Because these three individual measurements ($\mathrm{ToT}_5, \mathrm{ToT}_6, \mathrm{ToT}_7$) have different maxima and minima, the geometric mean is chosen as a function to unite these variables to a total time on task (ToT) as follow:

$$ToT = \sqrt[3]{\mathrm{ToT}_5 * \mathrm{ToT}_6 * \mathrm{ToT}_7}$$

The average ToT a user takes is $M = 10$ seconds ($SD = 3.71$s). The minimal ToT is 4.4 seconds and the maximal ToT is 20.8 seconds.

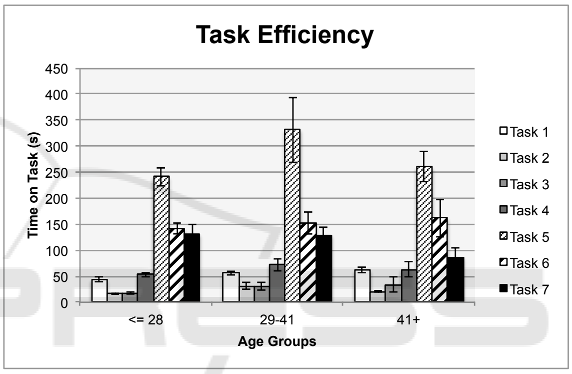

Figure 6: Efficiency measure marginal means for all tasks over age. Error bars denote standard error.

It seems necessary to incorporate the success rate into the efficiency measurement, because users could be faster at a task, when they simply skipped certain elements of a task. Therefore a corrected time on task (CTOT) is calculated using the effectiveness measure (SR=success rate) of the last three tasks.

$$CToT = \sqrt[3]{\mathrm{ToT}_5/SR_5 * \mathrm{ToT}_6/SR_6 * \mathrm{ToT}_7/SR_7}$$

This measure shows a mean of 14.4 seconds ($SD = 14$) with a minimum of 4.4 seconds and a maximum of 62.4 seconds.

Univariate ANOVA analysis indicates that age groups do not differ in regard to efficiency ($F(2,25) = 2.245$, $n.s.$). Correlation analysis confirms this ($r = .306$, $p = .14$).

9.2 Effects on Performance

When trying to understand how performance is determined multiple linear regression analyses are performed.

First a model for success rate is generated using age, diabetes duration, technical expertise and computer self-efficacy as possible predictors. Interestingly only a model with a single predictor remained significant. The model used expertise in household technology as a predictor and was able to explain 27% (adj. $r^2 = .27$) more variance than the scale mean ($F(1,17) = 7.714$, $p < .05$). The constant term showed a coefficient of $B = 2.185$ ($SEB = 0.473$) and the predictor EHT showed a coefficient of $B = -0.265$ ($SEB = 0.09$) and a standardized slope of $\beta = -.559$. This means that household technology actually influences effective negatively.

Secondly a model of corrected time on task was generated using the same predictors as in the last regression. A model using EHT and CSE as predictors was shown to explain 47% (adj. $r^2 = .47$) more variance than the scale mean ($F(2,15) = 8.52$, $p < .01$). Expertise with household technology again showed negative influence on performance while computer self-efficacy showed positive effects (see Table 1). Variance inflation was negligible ($VIF = 1.036$). The only downside of this model is the relatively high standard error of the constant term. Removing the weakest predictor from the model, did

Table 1: Linear regression table for corrected time on task. All predictors increased the explained variance significantly.

Predictor	Unstand. coeff. B	SEB	Standardized slope β
(Constant)	-32.476	37.140	
EHT	16.952	5.732	.531
CSE	-8.508	3.711	-.412

The only downside of this model is the relatively high standard error of the constant term. Removing the weakest predictor (i.e. CSE) from the model, did not decrease the standard error to a drastically more acceptable size ($SEB = 32.7$), while at the same time the explained variance decreases to 33% (adj. $r^2 = .33$, $F(1,16) = 9.358$, $p < .01$).

9.3 Pen Evaluation

After the experiment was finished it was of high interest to find out, how much participants liked the digital pen-based diabetes diary. We based the assessment on a TAM model using both perceived ease of use and perceived usefulness as predictors for the behavioral intention. Furthermore we wanted to investigate how technical expertise influenced behavioral intention.

We measured perceived ease of use and usefulness for nine features and combined them into two scales. Reliability was analyzed using Cronbach's α. Perceived ease of use showed excellent reliability with a Cronbach's $\alpha = .914$, similarly to perceived usefulness, which showed also an excellent reliability of $\alpha = .909$. The scales are calculated as the mean of the items.

The three behavioral intention items showed an acceptable reliability of $\alpha = .782$. The scale BI was calculated as the mean of these items.

When looking at how the age group influences the three scales PEU, PU, and BI (see Figure 7), differences are hard to make out. Only perceived ease of use seems to show a difference between young and medium aged users. Though Univariate ANOVA rejects this hypothesis ($F(2,19) = 2.373$, *n.s.*). In general perceived ease of use is high ($M = 5.01$, $SD = 0.75$), perceived usefulness is very high ($M = 5.44$, $SD = 0.58$) and behavioral intention is also very high ($M = 5.38$, $SD = 0.58$).

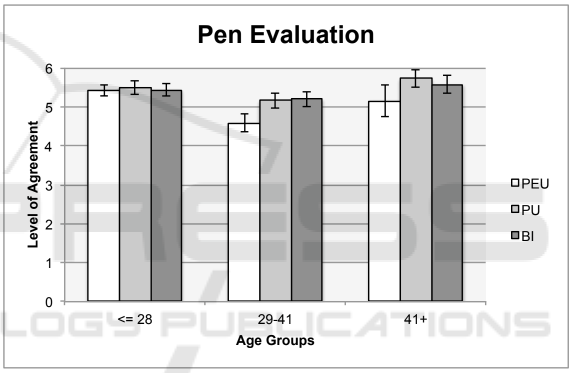

Figure 7: Marginal means of PEU, PU, and BI over age groups. Error bars denote standard error.

In order to determine what factors predict the behavioral intention a TAM based model using both PEU and PU as predictors was used. Additionally diversity factors like age, gender (dummy coded), diabetes duration, technology-expertise, and computer self-efficacy were used as predictors. Only the pure TAM based model using a two predictor model with PEU and PU as predictors was able to significantly explain more variance than the scale mean. This model was able to explain 86% (adj. $r^2 = .859$) more variance than the scale mean ($F(2,15) = 52.747$, $p < .01$). This is a very large effect. The perceived usefulness was about three times stronger in predicting the behavioral intention than the perceived ease of use (see Table 2). Both predictors had a positive influence on the behavioral intention.

Table 2: Linear regression results for behavioral intention.

Predictor	Unstand. coeff. *B*	*SEB*	Stand. slope β
(Constant)	-0.04	0.54	
PEU	0.210	0.123	.244
PU	0.798	0.155	.734

10 QUALITATIVE INSIGHTS

During the experiments positive feedback about the pen in general was extensive. Participants were surprised that a pen was able to "understand" what they were writing. One participant (female, 55 years) was very reluctant to take part in the experiment because, she "hate[s] all technology." She said:

"Stay away from me with this nonsense. All you engineers should try to learn that normal people like me cannot use these complicated devices."[1]

After persuading her to take part in the experiment and using the pen she stated that she loved the friendly interaction with the pen, and the pen actually felt good in her hand. She would like to have pens be as heavy as this one, as normal pens were to tiny for her to use.

Most participants showed similar reactions. Even the non-diabetic participants were delighted with the pen usage and in particular with the feeling that technology could "understand" them. Some even asked whether they could buy a pen for non-diabetic related tasks.

11 DISCUSSION

The results (see Figure 8) from the Diabetto user studies are interesting because some of the results are rather unexpected, while others are very much expected. We could confirm that age does have a negative effect on computer self-efficacy and expertise with household technology (supporting H_1 and H_2), while at the same time it has no effect on computerized technology (disagreeing with H_2). This could be a sample effect. The size of the sample is rather small and the sample is very tech savvy.

Much more interesting though is that the expertise with household technology exerts a negative effect on the efficiency with the pen (disagreeing with H_3), while computer self-efficacy exerts a positive effect (supporting H_3). A possible explanation might be that a pen that is used by hand with no buttons or other interaction features that are typical for computerized technology is more like a non-technology object. People that tend to type more and write less might be slower at writing with a digital pen as well. The only troublesome finding is that computer self-efficacy in turn does influence efficiency positively. This could be explained in a way that people that are more confident in using technology feel less intimidated by the computerized interaction with the pen.

Influences of diabetes expertise through duration or domain knowledge were not found in regard to effectiveness (rejecting H_4). This could mean that the diary was as easy to use for diabetics as it was for non-diabetics. This effect could be explained by the extensive help-function that explained many of the features anytime during the experiment.

Any effect of efficiency or effectiveness on the pen evaluation was not found (rejecting H_5). These were expected as similar findings existed in the experiments described in (Calero Valdez et al., 2009).

The influence of perceived ease of use and perceived usefulness on behavioral intention was similarly strong as predicted by theory (supporting H_6). Both showed a strong influence. The weakness of prediction of perceived ease of use could be explained with the general high rating of perceived ease of use for the whole sample. Looking at perceived usefulness though reveals a similar distribution. Usefulness still seems to be more important than ease of use, when ease of use is high. Ease of use could therefore be seen as a necessary condition for acceptance, while usefulness poses as a sufficient condition.

12 LIMITATIONS AND FUTURE WORK

This study tried to use both, qualitative and quantitative methods in order to develop a deeper understanding of the requirements of diabetes patients and to implement the user requirements into the development of the prototype. However, future studies will have to validate the device performance data with a larger sample. On the base of the present data, extrapolation is not yet feasible.

When reviewing how effectiveness is measured and what other effectiveness measures exist, we can see that only effectiveness of introductory tasks correlated with age. Further research could therefore look into other tasks and learnability of these tasks as well.

Strangely trust in doctors competence played a large role in perceived ease of use and perceived usefulness. This exploratory finding might be a fluke or could indicate high agreeableness in a participant.

[1] Translated from german.

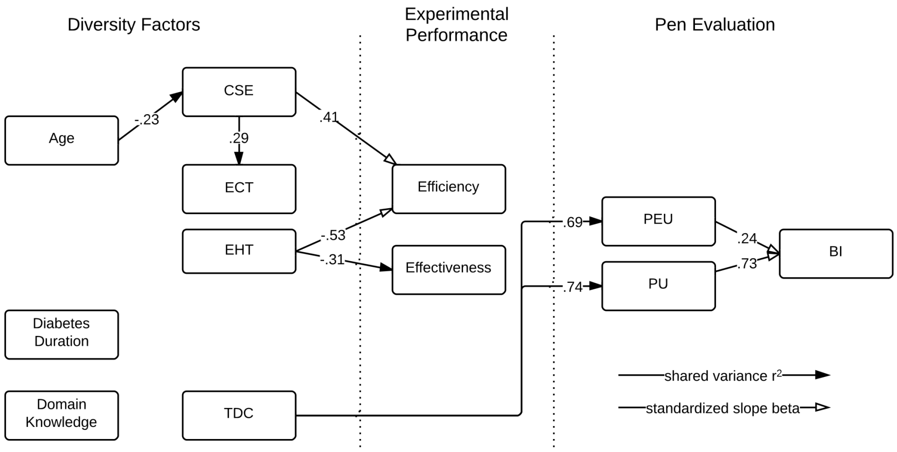

Figure 8: Visual representation of results, showing both correlational and linear regression data. The minus sign in front of a value indicates the direction of the correlation, and is not result of an complex number squared. It was added after squaring.

Examining this and further personality types might be of interest in further research.

ACKNOWLEDGEMENTS

We would like to thank Hennadiy Verkh for the development of the software prototype. Furthermore we would like to thank the participants for taking part in the study and sharing their insights. We would also like to thank the anonymous reviewers for their input on a previous version of this article.

REFERENCES

Alagöz, F., Calero Valdez, A., Wilkowska, W., Ziefle, M., Dorner, S., & Holzinger, A. (2010). From cloud computing to mobile internet, from user focus to culture and hedonism: The crucible of mobile health care and wellness applications. In *2010 5th International Conference on Pervasive Computing and Applications (ICPCA)*, pages 38–45. IEEE.

Alagöz, F., Ziefle, M., Wilkowska, W., & Calero Valdez, A. (2011). Openness to accept medical technology – a cultural view. In *Information Quality in e-Health*, pages 151–170. Springer.

Arning, K. and Ziefle, M. (2007). Understanding age differences in PDA acceptance and performance. *Computers in Human Behavior*, 23(6):2904–2927.

Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer Networks*, 54(15), 2787–2805.

Beier, Guido. (2004). Kontrollüberzeugungen im Umgang mit Technik: ein Persönlichkeitsmerkmal mit Relevanz für die Gestaltung technischer Systeme. (Doctoral Thesis) dissertation.de.

Brands, A. M., Biessels, G. J., Haan, E. H. d., Kappelle, L. J., & Kessels, R. P. (2005). The effects of type 1 diabetes on cognitive performance a meta-analysis. *Diabetes Care*, 28(3):726–735.

Brands, A., Van Den Berg, E., Manschot, S. M., Biessels, G. J., Kapelle, L., De Haan, E. H., & Kessels, R. P. (2007). A detailed profile of cognitive dysfunction and its relation to psychological distress in patients with type 2 diabetes mellitus. Journal of the International *Neuropsychological Society*, 13(02):288–297.

Broderick, J. E., & Stone, A. A. (2006). Paper and electronic diaries: Too early for conclusions on compliance rates and their effects – Comment on Green, Rafaeli, Bolger, Shrout, & Reis (2006).

Burke, L. E., Warziski, M., Starrett, T., Choo, J., Music, E., Sereika, S., Sevick, M. A., et al., (2005). Self-monitoring dietary intake: current and future practices. *Journal of Renal Nutrition*, 15(3), 281–290.

Burke, L. E., Conroy, M. B., Sereika, S. M., Elci, O. U., Styn, M. A., Acharya, S. D., ... and Glanz, K. (2011). The Effect of Electronic Self-Monitoring on Weight Loss and Dietary Intake: A Randomized Behavioral Weight Loss Trial. *Obesity*, 19(2), 338–344.

Calero Valdez, A., Ziefle, M., Horstmann, A., Herding, D., & Schroeder, U. (2009). Effects of aging and domain knowledge on usability in small screen devices for diabetes patients. In *HCI and Usability for e-Inclusion*, pages 366–386. Springer.

Calero Valdez, A., Ziefle, M., Alagöz, F., & Holzinger, A. (2010a). Mental models of menu structures in diabetes assistants. In *Computers Helping People with Special Needs*, pages 584–591. Springer.

Calero Valdez, A., Ziefle, M., Horstmann, A., Herding, D., & Schroeder, U. (2010b). Task performance in mobile and ambient interfaces. does size matter for usability of electronic diabetes assistants? In *Information Society (i-Society), 2010 International Conference on the Information Society*, pages 514–521. IEEE.

Calero Valdez, A., Ziefle, M., Horstmann, A., Herding, D., & Schroeder, U. (2011). Mobile devices used for medical applications: Insights won from a usability study with diabetes patients. In *The International Journal of Digital Society (IJDS)*, 2(1):337–346.

Chen, Y. (2010). Take it personally: accounting for individual difference in designing diabetes management systems. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, DIS '10, pages 252–261, New York, NY, USA. ACM.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, & user acceptance of information technology. *MIS Quarterly*, 13(3):319–340.

Davis, G. B. (2002). Anytime/anyplace computing and the future of knowledge work. *Communications of the ACM*, 45(12), 67–73.

Duke, D. L., Thorpe, C., Mahmoud, M., & Zirie, M. (2008, March). Intelligent Diabetes Assistant: Using machine learning to help manage diabetes. In *Computer Systems and Applications, 2008. AICCSA 2008*. IEEE/ACS International Conference on (pp. 913–914). IEEE.

Forjuoh, S. N., Reis, M. D., Couchman, G. R., & Ory, M. G. (2008). Improving diabetes self-care with a PDA in ambulatory care. *Telemedicine and e-Health*, 14(3), 273–279.

Gregory, C. L. (2008). But I Want a Real Book. *Reference & User Services Quarterly*, 47(3), 266–273.

Hader, C., Beischer, W., Braun, A., Dreyer, M., Friedl, A., Füsgen, I., Gastes, U., Grüneklee, D., Hauner, H., Köbberling, J., et al., (2004). Diagnostik, Therapie und Verlaufskontrolle des Diabetes mellitus im Alter. *Diabetes und Stoffwechsel*, 13:31–56.

Hien, P., & Böhm, B. (1997). *Diabetes-Handbuch: eine Anleitung für Praxis und Klinik*. Springer.

Kerkenbush, N. L. (2003). A comparison of self-documentation in diabetics: electronic versus paper diaries. In *AMIA Annual Symposium Proceedings* (Vol. 2003, p. 887). American Medical Informatics Association.

Kerkenbush, N. L., & Lasome, C. E. (2003). The emerging role of electronic diaries in the management of diabetes mellitus. *AACN Advanced Critical Care*, 14(3), 371-378.

Klack, L., Schmitz-Rode, T., Wilkowska, W., Kasugai, K., Heidrich, F., & Ziefle, M. (2011). Integrated Home Monitoring and Compliance Optimization for Patients with Mechanical Circulatory Support Devices. *Annals of biomedical engineering*, 39(12), 2911–2921.

Krishna, S., Boren, S. A., & Balas, E. A. (2009). Healthcare via cell phones: a systematic review. *Telemedicine and e-Health*, 15(3), 231–240.

Scheermesser, M., Kosow, H., Rashid, A., & Holtmann, C. (2008, January). User acceptance of pervasive computing in healthcare: Main findings of two case studies. In *Pervasive Computing Technologies for Healthcare*, 2008. PervasiveHealth 2008. Second International Conference on (pp. 205–213). IEEE.

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2002). Patient non-compliance with paper diaries. *BMJ*, 324(7347):1193–1194.

Tani, S., Marukami, T., Matsuda, A., Shindo, A., Takemoto, K., & Inada, H. (2010). Development of a health management support system for patients with diabetes mellitus at home. *Journal of medical systems*, 34(3), 223–228.

Venkatesh, V. and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, pages 186–204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3):425–478.

Venkatesh, V. and Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2):273–315.

Venkatesh, V., Zhang, X., & Sykes, T. A. (2011). Doctors do too little technology: A longitudinal field study of an electronic healthcare system implementation. *Information Systems Research*, 22(3):523–546.

Yeung, S. E., Fischer, A. L., & Dixon, R. A. (2009). Exploring effects of type 2 diabetes on cognitive functioning in older adults. *Neuropsychology*, 23(1):1–9.

Ziefle, M. (2002). The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. *Behaviour & Information Technology*, 21(5), 303-311.

Ziefle, M., & Wilkowska, W. (2010, March). Technology acceptability for medical assistance. In Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on (pp. 1–9). IEEE.