Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biomedical Informatics 36 (2003) 45–60
Journal of Biomedical Informatics
www.elsevier.com/locate/yjbin

# Usability in the real world: assessing medical information technologies in patients' homes

David R. Kaufman,[a,*] Vimla L. Patel,[a] Charlyn Hilliman,[a] Philip C. Morin,[b] Jenia Pevzner,[a] Ruth S. Weinstock,[b,c] Robin Goland,[d] Steven Shea,[a,e] and Justin Starren[a]

[a] *Department of Biomedical Informatics, Columbia University, New York, NY, USA*
[b] *Joslin Diabetes Center and Division of Endocrinology, Diabetes and Metabolism, SUNY Upstate Medical University, Syracuse, NY, USA*
[c] *VA Medical Center, Syracuse, NY, USA*
[d] *Naomi Berrie Diabetes Center, Columbia University, New York, NY, USA*
[e] *Mailman School of Public Health, Columbia University, New York, NY, USA*

Received 17 July 2003

**Abstract**

*Objective:* This paper presents an approach to usability evaluation of computer-based health care systems designed for patient use in their homes. Although such devices are becoming more prevalent, there is very little known about their usability.

*Design:* The theoretical foundations for the methods are discussed. The approach incorporates a cognitive walkthrough usability evaluation and new methods for usability testing that can be conducted in patient's homes. The method was applied to the IDEATel intervention, a multi-institution randomized controlled trial of the feasibility, acceptability, and clinical utility of a home-based telemedicine system for diabetic Medicare population. The usability study was designed to assess barriers to optimal use of the system. The focus was both on dimensions of the interface and on dimensions of patient skills and competency. The usability field research involved testing 25 patients in their homes using the system. The analysis included a range of video-analytic methods of varying levels of granularity.

*Results:* The usability evaluation revealed aspects of the interface that were sub-optimal and impeded the performance of certain tasks. It also found a range of patient-related factors such as numeracy and psychomotor skills that constituted barriers to productive use.

*Conclusions:* A multifaceted usability approach provided important insight regarding use of technology by an elderly chronic-care patient population and more generally, for understanding how home health initiatives can more effectively use such technology.


*Keywords:* Usability evaluation; Field usability testing; Cognitive methods; Telemedicine; Diabetes; Chronic illness

## 1. Introduction

Chronic illness affects over 100 million individuals in the United States [1]. Many of these individuals suffer from multiple afflictions and over 40 million of them are limited in their daily activities by their condition. The societal and financial costs of chronic illness are increasing as the population ages and effective treatments forestall mortality. In addition, people with lower incomes, especially those who have less access to quality health care, tend to be burdened with more serious conditions. Bodenheimer and colleagues [2] argue that the current model of primary-care medicine, a system designed for acute rather than chronic care, is ill-suited to the task of taking care of chronic-care patients. Wagner and Groves [3] advocate a new model of chronic care, one in which patient self-management plays a more central role.

Self-management initiatives increasingly rely on the use of technologies to facilitate the process of care in the home. These technologies range from medical devicessuch as glucose monitors to comprehensive

* Corresponding author. Present address: Laboratory of Decision Making and Cognition, Department of Biomedical Informatics, Columbia University, 622 W. 168th Street, VC-5, New York, NY 10032-3720, USA. Fax: +1-212-305-3302.
*E-mail address:* davek@dbmi.columbia.edu (D.R. Kaufman).

1532-0464/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S1532-0464(03)00056-X

computer-mediated telemedicine systems that provide interactive support as well as World Wide Web access. Although such devices are required to meet certain standards, very little is known about their usability [4]. Problematic user interfaces can induce errors and thereby compromise patient safety. The use of such systems can present some difficulties for health care professionals. Therefore, it is reasonable to assert that these same systems may present formidable challenges to chronic care patients who are typically older, less educated, and often have minimal experience with computers.

The World Wide Web is emerging as a vital knowledge resource for patients and consumers of health information. In *providing* unprecedented access to high quality information, it is increasingly serving as a mediator of health education, decision making, and management. Home telemedicine is a medium with potential to transcend social, economic, and geographic barriers [5]. With the rapid growth of the Internet and related technologies, telemedicine may serve to bridge significant gulfs of accessibility in the delivery of quality health care. In inner cities, populated largely by minorities, the obstacles include language, low educational attainment, disempowerment, and lack of social support for health-related behaviors and activities. The delivery of traditional health care services to rural settings is similarly challenged by factors such as geography, climate, social, and economic barriers [5]. Telemedicine affords the possibility of breaking down these barriers to improve access and thereby contributes to reduction in disparities among socio-demographic groups in access to care, quality of care and health outcomes [6].

There is a paucity of evaluation research on patient populations using home health care technologies. The greatest threat to the effective and safe use of complex technological systems is events that are unfamiliar to users and that have not been anticipated by designers [7]. There is a need to understand the usability of these devices and also the set of core competencies and knowledge that are required to productively operate this technology.

This paper presents a methodological approach to the study of usability of medical information systems in patients' homes. We are particularly interested in understanding how to facilitate the design and re-design process of home telecare systems for patients with chronic health conditions. In addition, a focal point of this research is characterizing the barriers to productive, efficient, and safe use of these systems towards the goal of sustainable autonomous self-management by patients. The theoretical and methodological framework for usability evaluation is detailed in Section 2. The framework is illustrated in the context of elderly patients' use of a home telemedicine system. The final Section 4 provides a summary of the lessons learned from field usability research. The authors argue that such research is needed to address a critical gap in knowledge regarding the use of technology by elderly chronic-care patient populations and more generally, for understanding how home health initiatives can more effectively use such technology.

## 2. Methodological framework

The research presented in this paper is informed by a cognitive engineering approach to the study of human–computer interaction. This is an interdisciplinary approach to the development of principles, methods, and tools, to guide the analysis and design of computer-based systems [8]. This work is influenced by Norman's theory of action [9] which posits a cyclic model of human computer interaction. This is a continuous process of iterative interaction with a system, beginning with a users' goal (e.g., open application), leading to an action (click on icon) resulting in a change in the state of the system (application opens new document). Ideally, the user recognizes that the intended change in the state of the system has occurred, thereby satisfying the goal and leading to a subsequent cycle of goal, action, and system response. The research is also informed by a distributed cognition framework in which routine human cognition is seen as distributed across individuals (e.g., a team of medical professionals and a patient) and technology. A more complete exposition of this perspective is reported by Horsky et al. [10].

The research is predicated on a two-pronged methodological approach to the study of human–computer interaction [11]. The first component consists of a *cognitive task analysis* of the system carried out by the team of investigators. The second part of this work involves *field usability testing* of patients performing a series of tasks using the system. This includes a video-analytic approach to the study of human computer interaction. In general, there are only a few resources on the subject of field usability [12]. The research is informed by methods from the ethnography of work and education, and in particular, interaction analysis [13]. Interaction analysis is an interdisciplinary method for the empirical investigation of humans interacting with others and with objects in their environment. This method investigates human activities such as discourse, nonverbal interaction, and the use of artifacts and technologies, identifying routine practices and problems and the resources for their solution [13]. Most importantly for the present purposes, it offers a series of guidelines for conducting video-analytic field research, conventions for the transcription and annotation of verbal protocols, and strategies for identifying analytic foci (e.g., the structure of events). Jordan and Henderson [13] also offer suggestions on how to draw reasonable inferences that are sanctioned on the basis of evidence from a corpus of video data.

### 2.1. Cognitive walkthrough

The cognitive walkthrough (CW) is a usability inspection[1] method in the form of a cognitive task analysis [14] which has been applied to the study of usability and learnability of several distinct medical information technologies [11,15]. The purpose of a CW is to evaluate the cognitive processes of users performing a task. The method involves identifying sequences of actions and goals needed to accomplish a given task. The specific aims of the CW procedure are to determine whether the user's background knowledge and the cues generated by the interface are likely to be sufficient to produce the correct goal–action sequence required to perform a task. The method is intended to identify potential usability problems that may impede the successful completion of a task. To perform a CW analysis, a researcher/analyst or group of analysts performs a task simulation, "walking through" the sequence of actions necessary to achieve a goal. Both behavioral or physical actions such as mouse clicks and cognitive actions (e.g., inference needed to carry out a physical action) are coded. The principal assumption underlying this method is that a given task has a specifiable goal–action structure (i.e., the ways in which a user's objectives can be translated into specific actions).

The CW method assumes a cyclical pattern of interaction as described previously. The codes for analysis include *goals* which can be decomposed into a series of *subgoals* and *actions.* For example, opening an Excel spreadsheet (goal) may involve locating an icon or shortcut on one's desktop (subgoal) and double clicking on the application (action). We also characterize the system response (e.g., change in screen, update of values) and attempt to discern potential problems.

The CW analysis also provides us with substantial insight into the cognitive demands of a task. For example, tasks that require the user to execute lengthy sequences of actions or require movement between different screens make heavier demands on a user's working memory. Similarly, a graphical representation or display that is littered with objects and text will necessitate extensive perceptual processing. We can anticipate that such systems will place a strain on a user's limited attentional resources and may be an ineffective tool. An important consideration in carrying out a walkthrough is an understanding of the target population. In this context, elderly users of a system are likely to have a lower tolerance for excessive memory or attentional demands. One of the most desirable properties of the walkthrough is that it yields a theory of competent performance [15], which can also be used as a basis for coding user data. The competence theory specifies the set of skills and knowledge needed to perform a task. Usability is therefore a function of both the system interface and user characteristics.

[1] Usability inspection methods are a software evaluation process that involves trained evaluators rather than end-users.

### 2.2. Overview of field usability testing

Usability testing refers to a process that employs participants who are representative of a particular target population to evaluate the degree to which a product or a system satisfies basic usability criteria [12]. It is regarded as perhaps the most informative test of the adequacy of a particular system. Usability testing names a range of designs and methods, ranging from controlled experiments (e.g., comparison tests) to informal studies with a single participant. Although usability testing is more commonly conducted in laboratory settings, field testing at clinical sites has become increasingly possible with the advent of portable usability laboratories [16]. At present, we do not know of any published field usability research that has been conducted in patient's homes. Relatively little is known about seniors as a population of computer users even though basic research on aging provides us with some insights into this group. Many of the study design decisions informing this research were predicated on the relative novelty of studying this population, especially in their own homes. In addition, a primary objective of this research was to investigate dimensions of competency and barriers to productive use of systems. Laboratory testing affords a degree of control that is not possible in a naturalistic setting. On the other hand, field research provides a glimpse of human–computer interaction under more realistic conditions that approximate their actual context of use.

In our research, we draw on theories and methods derived from qualitative research, in particular, ethnography, participant observation, and interaction analysis. However, field usability research can differ in important respects from more naturalistic or observational studies. Investigators may endeavor to exert a degree of control over test subjects, for example asking them to perform a series of tasks. To reiterate, our hypothesis is that field usability testing will confirm many of the findings from the cognitive walkthrough analysis, but more importantly, will provide us with additional insights into the barriers of productive participation in a telemedicine intervention. The methodological approach is discussed in detail in Section 2.3.

Ideally, CW analysis and usability testing form a tightly coupled process. The CW analysis provides a descriptive and procedural characterization of a system task and this enables us to formalize an analysis for usability testing. On the other hand, a CW is predicated on an understanding of a target population and usability testing provides a more in-depth characterization of the population. This facilitates more targeted walkthroughs in the context of iterative design. Field usability testing is a relatively expensive and time-intensive undertaking whereas the CW is less expensive and can be performed in a fraction of the time. A CW can be repeated multiple times over the course of iterative design.

In the following sections, we describe an approach we employ for usability testing of health care systems. Kushniruk et al. [17] outline a stepwise approach for usability testing. Similarly, Cimino et al. [16] detail an approach to field usability testing in a clinical setting that provides an interesting study in contrast. The goals of this section are to describe the approach employed in this research and to provide some general guidelines and practical advice for carrying out field usability research.

### 2.3. Development of a test plan and selection of tasks

In this critical first step, the evaluators outline the task and procedure for user testing. The plan is informed by an objective, which in this case is to understand barriers to productive use of a telemedicine system. The plan may involve an exploratory test, typically conducted early in the development cycle with a few users or a more structured test such as the comparison of two interfaces [12]. In the latter case, the test may utilize an experimental design. Field usability is more likely to be employed later in the development cycle in view to characterize how a fully developed system works in a practical context.

### 2.4. Selection of representative users

Usability testing, like other areas of qualitative research, tends to employ participants based on specific criteria rather than through a random sampling process. Users may differ on a range of dimensions including age, education, gender, computer experience, domain expertise, and areas of specialization (if applicable). It is not possible to employ a fully representative sample. However, relying exclusively on a convenience sample such as power users or early adapters is not likely to provide a sufficiently robust or realistic test of the system. It is important to identify relevant criteria to distinguish user types and select a reasonably diverse sample.

### 2.5. Setting up the testing environment

Video provides a remarkably rich and vivid reproduction of an encounter. It also provides a permanent record of an event and supports multiple viewings and re-analyses of data [13]. In our field studies, all sessions are audio taped and videotaped. In addition, screens are captured to a digital camera via a VGA to TV scan converter (MicroJack, Ontario, Canada). The computer display is transformed into a video signal (S-video) and sent to a small digital camcorder (Canon ZR25). The subject is also videotaped using a video 8 camcorder (Sony CCD-TRV11). The video 8 camcorder is mounted on a tripod and placed in reasonably close proximity to the subject. The audio signals from the two cameras are synchronized and a standing omnidirectional microphone is placed on the table to record the subject's verbalizations. A cassette recorder is used to provide a redundant source of audio. It is also easier to transcribe an audiotape than it is a videotape. Maintaining an inventory of all equipment including cables and tapes is advisable to prevent loss and to ensure that the team is prepared with all essentials when you arrive at a site (cf. Fig. 1).

Fig. 1. Subject using the system. The scan converter is located on the corner of the desk and is attached to a digital camcorder.

The setup typically takes 10–15 min if no problems should arise. The camera should ideally be placed in such a way as to capture the patient's profile (it is more difficult to protect their confidentiality if they are recorded from the front) and their fingers. The primary focus of the analysis is on their interaction with the computer and capturing hand and finger movements are important. Problematic spatial layouts, the availability of quality electrical outlets, noise in surrounding areas are just a few of the intangibles that can complicate the data collection process. It is vitally important to test one's setup locally before embarking on a usability road trip.

Although video provides a durable record of an event and supports repeated viewing by multiple participants, it is not necessarily a neutral tool or one that perfectly reproduces direct observation. The researcher must choose a location and particular field of view for the camera [18]. Invariably certain phenomena must be selected for inclusion and others excluded. This issue is discussed further in the data analysis section.

### 2.6. Role of researcher

Unlike most usability testing where the focus is exclusively on the user, the researchers may play a more interactive role in field testing. As a consequence, it is not possible to delete them from the analysis or treat them as a completely neutral agent. As mentioned previously, the researchers instruct the subject and may guide them in their performance of the task. The extent of the researcher's involvement is dependent on the facility of the user. A skilled user will need minimal guidance, whereas a novice user may need step-by-step instructions or even hands-on demonstrations in which the researchers will execute some of the actions. Therefore, the researcher is viewed as a participant and his or her contributions are coded accordingly.

### 2.7. Data analysis

#### 2.7.1. Transcriptions

The audio segments of the session are first transcribed verbatim from the audiotape. The transcriber attempts to faithfully reproduce the discourse and exchanges between researchers and patient. The participants are marked accordingly and the time is noted at 30 s intervals (information can be obtained from the tape counter). This *time stamping* process enables us to identify the beginning and end of events and sync the audio with the video. The transcribing of the audiotape represents only the first pass. The transcript emerges as an iteratively modified document that increasingly reflects the categories that the analysts views as central [13]. The next step is to add field notes and observations to the transcript, and this represents a working document for video analysis.

#### 2.7.2. Video analysis

There are two sources of video data: the video of the subject and the video of the screen display captures. The video cameras are synched and digital camera is hooked up to a computer through a Firewire port (IEEE 1394) and is controlled by software such as IMovie on Macintosh computers and Ulead Video Studio on PCs. Both are relatively inexpensive and enable the analyst to control the camcorder (e.g., rewind, fast-forward) and capture video to a hard disk (or other storage medium). In addition, both enable the editing of video for various purposes such as presentation or archiving. There are more elaborate (and expensive) video-analytic software/hardware solutions that provide extensive resources for coding, indexing, quantifying, and storing video sequences in a database. These are valuable tools, but we recommend that one begin with relatively simple and inexpensive options. It is worth pointing out that even compressed video can require substantial space on a hard disk. The most common data rate for uncompressed digital video is roughly 4 Mbyte/s or one gigabyte every 4 min.

#### 2.7.3. Macroanalysis

The video editing session may include a single analyst or a team who view and comment on the video session. The videos are coded in multiple ways at varying levels of granularity. The first pass is a relatively coarse or macro-level analysis which involves segmenting the session into events or episodes and noting their duration on the transcript. The analysts also note particular difficulties that the subjects encountered and anything else that is significant. At this point the analysts' subjective impressions are included and will be scrutinized more carefully in subsequent video analysis sessions. The first phase of evaluation will require 5–10 h of analysis for each hour of transcript. Finer granularity of coding may require upwards of 20 h for each 1 h session. Macrocoding of videotapes can reveal a host of problems pertaining to usability and contribute substantively to the iterative design process. In fact, many analyses can stop at this point. However, since we were also interested in understanding dimensions of user competency, a finer level of granularity was required.

#### 2.7.4. Microanalysis

The microcoding of video is very time-consuming; it is, therefore, necessary to code selectively. It is possible to begin microanalysis by selecting a subsection of subjects and/or a subsection of tasks. The macro-level coding should provide guidance to the aspects that warrant further analysis. For example, one may choose to focus on the subjects who are genuine beginners in order to understand how to change the interface to meet their needs or to develop a training protocol. Alternatively, tasks that cause users more difficulty than others may be prioritized. In our case, web access and interpreting blood pressure values were tasks that presented difficulties for several subjects.

#### 2.7.5. Segmenting event structure

There are different ways to partition a video into manageable units of analysis. Chronological time provides a convenient way to characterize the activities observed on tape. One may divide a 1-h video into arbitrary 5–10 min segments. However, it is often more meaningful to divide the video into events. For the purposes of analysis, events can be defined as "stretches of interaction that cohere in some manner that is meaningful to the participants" [13]. In this context, events correspond to a task such as measuring blood pressure or viewing patient data.

Tasks often begin with the researcher instructing the subject to perform a task. The end of the event can be indicated either by task completion or the point at which the researchers decide that the subject is having too much difficulty and that it is best to move on to the next task. The cognitive walkthrough (CW) provides a basis for coding the task into constituent units of analysis. The basic codes include *goals*, *subgoals*, *actions*, and *system response.* The goal is typically expressed by the researcher when he or she instructs the user and the subgoals flow from a user's understanding of the task. When we conduct usability testing with participants who are reasonably skilled at using computers, we can code actions at a coarser level such as clicking on link or entering text. However, novice subjects encounter difficulties executing actions and it may be necessary to code at a finer level for mouse movements and keyboard presses.

#### 2.7.6. Verbal and nonverbal analysis

Speakers are noted as researcher and subject. In addition, requests and comments are coded accordingly. The propositional content of the patient's speech provides evidence on what he or she may understand and the sorts of difficulties they experience with the system. In addition to the verbal exchanges and behavioral actions, we also code for nonverbal behaviors including: (a) change in body position, (b) gesture, and (c) gaze. The analysis of nonverbal behavioral is predicated on the belief that cognition and knowledge is embodied, meaning that it is literally in the eyes and hands of the knower as well as the head [19].

Many novice computer users lack the expressive vocabulary to talk about objects on the screen such as scroll bars and buttons. Therefore gesture is an important expressive tool for these participants and needs to be considered for the present analysis. Gesture is emerging as an active area of research in the social science [20]. Goldin-Meadow [21] suggests that gestures serves as a tool for communication for listeners, and a tool for thinking for speakers. Gestures alternatively serve to complement speech or as a substitute for speech. Gesture is a reliable phenomenon that is found across cultures, ages, and tasks [21]. There are extensive taxonomies and sources for interpreting gestures [22]. However, we focus on a restricted class of gestures, namely deictic or pointing gestures that indicate entities in the conversational space (e.g., objects on the screen). In addition, we code for expressions of understanding or agreement (e.g., a head nod) and alternatively misunderstanding and/or disagreement.

In this usability testing situation, the participant is focused on either interacting with the system or with a researcher. Gaze plays an important role in coordinating both conversational interaction and in carrying out physical tasks [13]. It can be revealing to track people's eyes, when and how gaze moves between objects, from persons to objects, sustaining or shifting the focus of attention as events unfold (e.g., the display changes). There is currently very little research on nonverbal behavior in human–computer interaction and the coding of gesture and gaze is best viewed as exploratory at this point.

## 3. Usability evaluation of a home telemedicine system

The application of these methods is illustrated in the context of a usability evaluation of a home telemedicine system. The system and intervention are first described. The subsequent section focuses on understanding a target population prior to undertaking the usability evaluation research, drawing on both prior research on seniors and the knowledge we have gained about the specific population involved in the home telemedicine initiative. This is followed by an illustration of the cognitive walkthrough with two representative tasks and an analysis of task complexity. The last part of this section focuses on selected examples from the field usability testing demonstrating different facets of the methods and analyses.

### 3.1. The informatics for diabetes education and telemedicine project

The IDEATel project is a large-scale multi-institution randomized controlled study, designed to assess the efficacy of a home-based telemedicine system [6,23]. The target population is Medicare beneficiaries living in medically underserved areas including individuals in rural regions of Upstate New York and in urban areas, including Northern Manhattan and the Bronx. The Upstate population consists mostly of English speakers and tends to be somewhat more computer literate. The Downstate population is predominantly Hispanic and has less computer experience.

The focal point of the intervention is the home telemedicine unit (HTU) which provides the following functions: (1) synchronous videoconferencing, (2) electronic transmission of fingerstick glucose and blood pressure readings, (3) secure email to a physician and nurse case manager, (4) web-based review of one's clinical data, and (5) access to web-based educational materials. The system is designed to be accessible to elderly novice computer users. Simplicity is a guiding design principle. All components of the system and related services are available in both English and Spanish.

### 3.2. Understanding the target population

An effective usability study is predicated on a careful consideration of the intended population of users: in this case, seniors. Although seniors are using computers and the Internet with greater frequency, the gulf remains rather wide in comparison with other adults. According to a recent Department of Commerce Report [24], adults over the age of 65 are less likely to have ever used computers than any other demographic age group. The elderly are also more likely to be less affluent and have less education than younger adults, factors which are also associated with the so-called digital divide.

At present, there is a paucity of cognitive and/or human–computer interaction research that addresses the challenges seniors confront in learning to negotiate the Internet [25]. There is, however, a growing body of cognitive aging research that can inform design and health-care interventions for older adults [4]. This literature documents changes in psychomotor skill, memory, and learning retention rate and also provides some broad guidelines for design.

At the time of the study, the IDEATel project was in its second year of operation. Log file analysis indicated that patients performed certain tasks such as monitoring and uploading their blood pressure and glucose better than others such as web access. In addition, we interviewed several patients who had just been enrolled in the study. The objectives were to get a better sense of the study population and to attempt to find out whether we can explain patterns of access of various facets of the IDEATel system. The interviews were semi-structured and covered a range of topics, including the users' experience with computers and the Internet, health-related matters, satisfaction with current level of health care, willingness to learn about health matters on the web, and their experience and understanding of the diabetes.

### *3.3. Cognitive walkthrough evaluation*

The IDEATel intervention enables several distinct superordinate tasks. We can define a task functionally as that which orients a user towards achieving an objective related to diabetes health care. Several of the tasks support multiple constituent tasks. For example, IDEATel includes a web-based diabetes diary called the Diabetes Manager that allows patients to track their own progress. The diabetes diary enables several goals for five variables: BP, glucose, medications, exercise, and viewing Hgba1c. For the first four of those variables, a patient can view his or her record over a period of time, edit an event, delete an event, and add a record. Tracking glucose progress also allows a user to graph glucose levels.

The glucose and blood pressure measurement tasks employed tightly coupled goal action sequences and were reasonably easy to execute. A tightly coupled sequence is one in which an action transparently flows from a goal and the user can readily perceive that the system has responded thereby signaling the next subgoal and action sequence. A partial walkthrough of the glucose task is illustrated below (cf. Fig. 2).

**Task/Goal: Measure Blood Glucose Level**
**1. Subgoal:** Begin Measurement
**Action:** Press Blue Power Button
**System response:** Meter Displays Last Blood Glucose Result
**2. Subgoal:** Obtain a Blood Sample
**Subgoal:** Use Sterile/Sharp Lancet
**Action:** Replace Lancet (if necessary)
**Subgoal:** Draw Blood using Instrument
**Action:** Pierce Finger with Lancet
**3. Subgoal:** Apply blood to test strip
**Subgoal:** Locate Pink Test Area
**Action:** Dab Finger/Touch Strip
**Potential problem:** Missing Test Area, Applying Excessive Blood, Applying too Little Blood
. . .. . .
**4. Subgoal:** Take Measurement with Device
**Subgoal:** Determine readiness of the device
**Action:** Look for flashing test strip on meter
**System response:** Code 4
**System response:** Flashing test strip
**5. Subgoal:** Insert Pink Test Strip
**Action:** Push test Strip in Firmly (Pink Side Up)
**System response:** Flashing Clock Signals Waiting

The walkthrough of the glucose-monitoring device revealed five subtasks, nine subgoal action pairings, 12 actions, and five device/screen transitions. Familiarity with the device components and related objects (e.g., meters, lancets, and test strips) facilitates the relative ease with which patients execute the task. This is in contrast to some of the web-based tasks. Accessing the web necessitates nine actions and seven screen transitions. The transitions include a series of displays with connecting messages and security-related screens. For the most part, the screens (including several security dialogue boxes) were not meaningful to the participants and might encourage passive responding. The transitions are likely to be a source of considerable confusion to beginner users and, in fact, usability testing appears to have confirmed this conjecture (cf. Fig. 3).

The web-based Diabetes Manager (Siemens Medical Solutions, Malvern, PA) screen represents the command center of the IDEATel system. Most of the tasks are initiated from this page. The screen is somewhat cluttered and tasks are not well segregated. In addition, there are labeled links such as glucose and blood pressure that appear twice but reference somewhat different functions (i.e., today's readings versus readings for some extended period of time). There is also an abundance of text and some of it appears in small lettering which proved to be difficult for some of the participants to read. The goal–action sequences and affordances across

Fig. 2. Glucose monitoring device.

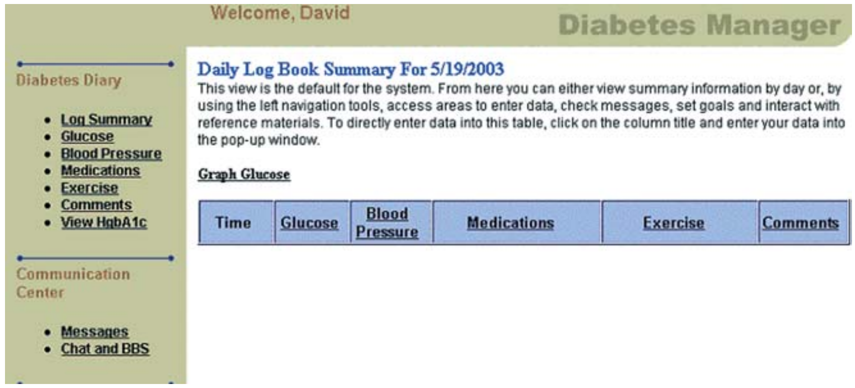


Fig. 3. Top half of the Diabetes Manager screen.

several of the subtasks are relatively consistent (e.g., using dialogue boxes for retrieving values) which supports and reinforces the learning process.

The Diabetes Manager was developed as a general purpose tool for the broader diabetic population and was not optimized for the elderly. For many elderly patients, vision and dexterity are significant issues. The analysis would suggest that subjects could experience difficulties reading the screens. In addition, many of the widgets present unique problems for this population of users. The problem appears to be both the lack of familiarity and the necessity of fine eye–hand coordination. The following task analysis involving tracking blood pressure illustrates some of these issues. Fig. 4

The tracking BP application enables an individual to perform a range of tasks in view to monitor one's glucose over a certain period of time. The following CW analysis pertains to changing the dates to a certain period of time in which to view one's glucose values.

**Goal:** View Progress over a specified period of time (other than the default value)
**Subgoal:** Determine Period of Time
**Subgoal:** Change Dates
**Subgoal:** Select Date
**Subgoal:** Change Month in "From" Slot
  **Action:** Bring Cursor to Month Field
  **Action:** Click on diamond on Pull Down Menu
  **System response:** Pull Down Menu Unfolds
  **Action:** Scroll Down to Correct Month
  **System response:** Selected Month is Highlighted
**Subgoal:** Change Date in "From" Slot
  **Action:** Bring Cursor to Date Field
  **Action:** Click on diamond on Pull Down Menu
  **System Response:** Pull Down Menu Unfolds
  **Action:** Scroll Down to Correct Date
  **System response:** Selected Date is Highlighted
**Subgoal:** Change month in "To" Slot
  **Action:** Click on diamond on Pull Down Menu
  **System response:** Pull Down Menu Unfolds
  **Action:** Scroll Down to Correct Month
  **System response:** Selected Month is Highlighted
**Subgoal:** Change Date in "To" Slot
  **Action:** Click on diamond on Pull Down Menu
  **System response:** Pull Down Menu Unfolds
  **Action:** Scroll Down to Correct Date
  **System response:** Selected Date is Highlighted

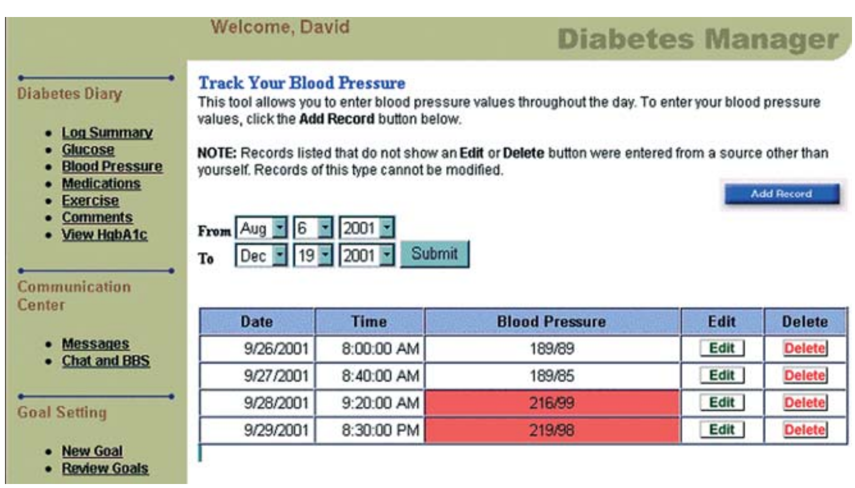


Fig. 4. Tracking blood pressure interface.

The calendar widget is common to a range of applications and it is relatively straightforward from a cognitive vantage point necessitating the repetition of the aforementioned goal–action sequence applied to each of the date fields. However, elderly subjects could experience significant difficulty with the narrow scroll bars and pulldown menus. The use of these widgets necessitates fine eye–hand coordination. So we may anticipate that new users could be challenged to master this task.

Table 1 displays five common tasks and two basic measures of complexity. Sending an email message (to a nurse or physician) necessitates 13 actions, whereas accessing the web site requires only nine actions but involves seven screen transitions. As illustrated previously, changing the calendar is basic to viewing ones recorded values (e.g., glucose and BP) over some period of time. Certain tasks are likely to introduce additional complexities. For example, sending an email message involves keyboarding skills which novice users are not likely to have. This may prove to be more difficult than negotiating screen transitions. The cognitive walkthrough cannot precisely or quantitatively predict errors or problems, but rather highlights aspect of the tasks that are likely to be problematic or that make excessive demands for a particular population of users. The CW can reveal certain dimensions of user problems, but can be used most effectively in concert with user testing.

### 3.4. Field usability evaluation

#### 3.4.1. Test plan and task selection

In this critical first step, the evaluators outline the task and procedure for user testing. The plan is informed by the objectives, which in this case is to understand barriers to productive use of a telemedicine system. The objective of the test plan was to employ a representative set of tasks that were likely to be among the most commonly used by patients. Subjects were asked to perform the following series of tasks: (1) measure blood pressure, (2) upload results, (3) access the Diabetes Manager (Siemens Medical Solutions, Malvern, PA) web page, (4) review patient data, (5) generate and interpret a table of blood glucose results, and (6) visit the American Diabetes Association educational website specially designed for the IDEATel project. Subjects were asked to think aloud throughout the task and offer comments on each screen (and screen transition). In actuality, the discourse more typically resembled a conversation between the researchers and the patient.

Table 1
Task complexity for five common tasks

| Task | Actions | Screen transitions |
|---|---|---|
| Measuring glucose | 12 | 5 |
| Taking blood pressure | 9 | 5 |
| Access web site | 9 | 7 |
| Sending email message | 13 | 3 |
| Changing calendar | 10 | 1 |

#### 3.4.2. Representative sample

IDEATel intervention includes two distinct geographic populations and patients who differ in other important respects including language and level of education. Our study sample was selected to include participants from both urban and rural regions. In addition, it was important to include subjects who had been using the system with some regularity and those who did not. We conducted field usability testing in 25 subjects' homes, including 14 subjects in the New York City (NYC) area and 11 in Upstate New York Table 2 summarizes certain patient characteristics. A notable difference is in the years of education. The mean number of years of education was 12.1 for the Upstate subjects and 8.5 for the NYC area participants. In addition, 12 out of 14 NYC subjects were Spanish speaking, whereas all of the Upstate patients were English speaking. We selected both subjects who had been using the system to access the web with some regularity and those who had not.

Since participants in the usability testing were volunteers, it is inevitable that they cannot be fully representative. Our sample size was unusually large for usability testing because of the heterogeneity. Usability testing can often produce informative results with a small sample size. Our objectives were not merely to contribute to iterative design, but to gain understanding of the challenges confronted by seniors in a computer-mediated home health initiative.

#### 3.4.3. Procedure

Subjects were visited in their homes by two researchers. One of the researchers is a fluent Spanish speaker and also serves as a translator. The researchers explain the procedure to the patient and emphasize that the goal is to improve the system design with the intent of rendering all aspects of the intervention to be more accessible for all participants. They are assured that they will not be judged in any way and their participation in the intervention study will not be affected. The subjects

Table 2
Subject characteristics (means ± SD or %)

| | New York City $N = 14$ | Upstate $N = 11$ |
|---|---|---|
| Age (years) | 69.6 (6) | 73.7 (8) |
| Education level (years) | 8.5 (5) Range: 0–14 | 12.1 (3) Range: 7–16 |
| Living with diabetes (years) | 10.4 (10) | 12.1 (9) |
| Language | Spanish 86% | English 100% |
| Web site use[a] | 36% | 45% |

[a] Use of IDEATel web site prior to usability study.

are informed that they may stop at any time. They are then given the consent form and asked to read it carefully while the researchers set up the equipment. If they are unable to read the form, a family member or a researcher will assist them. They are free to ask any questions before the study commences. In a brief interview, prior to the user testing tasks, subjects were asked a series of demographic questions about their age, level of education and health status. They were also asked about their prior use of computers and their experience with diabetes. The interview segment, which typically lasted no more than 5–10 min, was audiotaped. Following the brief interview, the subjects were then instructed to sit down by the computer (if they were not already doing so). The cameras would then be turned on. The first task asks the subjects to take their blood pressure and then to upload (send) their results. Subsequent to that, they are asked to login to the system and access the web.

The participants varied considerably in terms of their levels of computer experience. Fifteen out of twenty five subjects had never used a computer prior to this session and the researcher assisted them accordingly. The subjects were provided with as much assistance as necessary, but no more than necessary. The goal was for the patient to assume as much autonomy as possible in using the system, but at the same time not allow them to become too frustrated. The researchers play a very different role in this kind of field usability research. He or she is not a neutral or objective observer, but a central participant in the encounter. However, it is important that the researchers maintain a level of consistency between each encounter and adhere to basic principles (e.g., encourages patient autonomy).

Using the mouse proved to be a formidable obstacle for some of the participants, and these subjects were not able to complete all of the tasks. Training was an integral aspect of this procedure. The entire session lasted between 45 and 70 min. The participants were asked to think aloud during the entire session. Subjects were also offered an opportunity to take a break if necessary. The session concluded after all of the tasks were complete or the participant proceeded as far as possible. The subjects were then provided the opportunity to ask any questions or request assistance if necessary.

During the session, one of the researchers interacted closely with the patient while the other researcher (when not translating) monitored the audio signal and took field notes. The video is able to provide a rich record of the encounter, but invariably there is much that is not captured (for example, a description of the setting). Field notes are an informative source of ancillary information. Each tape is logged and the field notes are written up as soon as possible after each session. This minimizes errors of recall that naturally transpire after an elapsed period of time.

### 3.5. Video analysis

#### 3.5.1. Macroanalysis

As mentioned previously, the videos are coded in multiple ways at varying levels of granularity. For the purposes of analysis, the Spanish dialogue was translated into English by one of the researchers who was also present at the encounter. This was necessary because most of the analysts were not fluent speakers of Spanish. Macroanalysis involves segmenting the session into events or episodes, noting their duration on the transcript, and the general difficulties that the subjects encountered. As illustrated in Fig. 5, the initial coded transcript contains the subject's verbalization and exchanges between the researchers and subject in the left-hand column. The analysis and comments are presented in the right-hand column. The comments include observations from the field notes, summarization of the interview, and descriptions of problems the user encountered. The analysts' impressions about the comfort level of the user were also noted.

The interview revealed that the patient had been diagnosed with diabetes 4–5 years ago and that her self-management skills were not well developed. The subject, a 69-year-old woman, had a relatively low level of literacy and no prior computer experience. The table presents the first part of the session up to the point of the blood pressure measurement task. This analysis revealed her anxiety and lack of comfort in using a computer, and the level of difficulty she had in negotiating a mouse. In addition, we observed that she positioned herself a considerable distance from the computer. For the purposes of this paper, both the dialogue and the coding are presented in an abbreviated form. This first level of video analysis or macrocoding was applied to all transcripts before any further analysis was undertaken. This provided us with some indicators of how different subjects performed and the overall usability of different aspects of the system. It also enabled us to develop a profile of each subject. Perhaps most importantly, it suggested areas that warrant further evaluation.

The macroanalysis indicated that there were a range of barriers that impeded participants' abilities to use the system more effectively. Barriers reflect a combination of individual competencies, system design limitations, and environmental variables such as social support. System limitations included problematic widgets (e.g., narrow scroll bars), small font size, links that are inadequately spaced, unnecessarily complex tasks (e.g., too many steps), nontransparent screen transitions, and system stability. Cognitive and skill-based barriers include psychomotor skills as reflected in mouse and keyboard use, mental models of the system (essentially understanding how the system works on a very basic level), and literacy and numeracy. We also observed that patients differed in terms of their level of anxiety in using

**Dialogue**

**Feb. 26 2002.**

Investigator: How old are you?

Patient: 69, almost 70

I: Your highest level of education, high school?

P: Second grade. After the first year.

I: So two years?

P: Yes, two years.

……

I: How long have you had diabetes?

P: Four years. Five or four…

I: Does it affect your life, for example your daily routine?

P: I cannot walk a lot. I have blood pressure and sugar.

I: Before you had a computer did you monitor your blood sugar daily?

P: Yes, I asked my doctor. The doctor gave me a machine.

I: So you measured your sugar every day?

P: Yes, a little. But with the computer I do it more.

I: Do you use the computer besides checking you blood pressure and sugar? Do you use it to see your results?

P: No. I see the results there. {On the blood pressure machine} And the nurse sees them there, in the hospital.

I: Could you please take a seat at the computer and complete some tasks. Can I ask you first to take your blood pressure?

P: Yes. How do I put it on? Like this?

I: How do you normally do it?

P: Usually he {the companion} helps me.

I: Could you please tell us what the three values are?

P: 93 and 170.

I: What does it mean?

P: These small ones?

I: Yes, the smallest ones?

P: I do not see them well. Ah, 69.

**Codes**

**Profile:**

**INT** 69 year old (almost 70) female

**INT** Native Spanish speaker

**INT** Second grade education

**INT** Resides in US 36 years

**INT** Originally from Dominican Republic

**OFN** Lives with male companion

**OFN** Rents one room in large apartment in Harlem.

**INT** Never previously used computer

**Diabetes History:**

**INT** Has diabetes 4-5 years

**INT** Had monitoring device before the computer, but used it less often

**INT** Had not used computer to measure glucose for months

**OFN** Has some few notes of values on one page of a notebook

**Taking Blood Pressure:**

**VU** Cannot take blood pressure alone without help

**VU** Does not see or know what pulse value stands for

**VU** Positions herself too far from the computer

**VU** Cautious about touching computer

**VU** Points at buttons but does not touch them when discussing what she has to do next

**VU** Visibly uncomfortable with computer

**Legend: I-Interview, P-Patient. Sources of Evidence: INT-Interview, OFN-Observation/Field Note, VU-Video of User, VSC-Video Screen Capture.** The video analysis transcript include 5 sources of evidence: 1) the verbal exchanges or dialogue during the videotaping, 2) the interview data prior to the video session, 3) observation or field notes (not derived from the video), 4) video of the patient, and 5) video screen capture (from the computer display) which is not in use in this brief excerpt. The dialogue is indicated in the left hand column and all codes are presented on the right hand side. The profile is listed at the beginning of every interview.

Fig. 5. Macro video analysis transcript.

the system, self-efficacy, and motivation. Physical health was also a limiting factor in some instances. In most respects, a macro-level analysis of video constitutes a relatively complete usability evaluation. However, an in-depth understanding of these issues warrants a finer level of analysis. This is discussed in the next section.

### 3.5.2. Microanalysis

As discussed in the methodological framework, microcoding of video is a very time consuming process and one has to be rather selective. The microanalysis of the video recording focuses on certain subjects and particular tasks. Both the subjects and tasks have been singled out because they can reveal important facets of the interaction with the system. In this section, we present a contrast of two subjects. The first subject is the same one discussed in the macroanalysis section. The subject's encounter with the system is revealing of the challenges that a novice computer user with relatively low literacy confronts when trying to learn the system. The second subject, despite having no prior experience with computers prior to receiving an IDEATel system, developed substantial mastery.

The video transcript of the microanalysis is represented in Fig. 6. The first segment of the encounter, from 39:10 to 40:48 (time taken from the digital video counter), represents an attempt by the researchers to orient

| Time | Dialogue and Conversational Code | Action Code | System Response | Comments |
|---|---|---|---|---|
| 39:10 | **R:** Could you explain what goes on now?<br>**Request Exp1:** to explain page | **Gesture:**Exp.1 motions up and down the screen (blood pressure values)<br>**Gesture:** P. turns away to retrieve her eyeglasses from her companion | Diabetes Manager Page | The task is to interpret daily values of blood pressure and glucose |
| 39:36 | **P:** It says... the glasses do not help me. I have to read it? I cannot read this, I can read only few things, but I know what is says. | **Action:** P. starts reading, moves her lips but does not utter clear sounds or words, adjusts glasses, and brings her hand to chin. | | |
| 40:02 | **R:** Do you know what these values represent? | **Action:** Exp.1 takes mouse in hand and uses cursor as pointer to illustrate his words.<br>**BPC:** P. turns to Exp.2, then turns back to screen. | Cursor moves to table | Experimenter asks if the patient can understand glucose values |
| | **P:** They represent health. (Intonation goes up, suggesting a question) | | | Patient is confused by tabular representation |
| 40:36 | **R:** Could you explain a little?<br>**P:** It explains how the patient…<br>I2: Can you see?<br>**P:** What happens that these glasses do not help. I know a general idea of what it says. | **Gesture:** P. motions weakly and vaguely with hand, then pushes glasses up nose bridge | | |
| 40:48 | **R:** OK. This is blood pressure | **Action:** Exp.1 takes mouse, and points at BP with cursor on screen | Cursor moves | Patient is still experiencing difficulties with mouse and experimenter is executing many of the actions. |

| Time | Dialogue and Conversational Code | Action Code | System Response | Comments |
|---|---|---|---|---|
| 48:16 | **R:** What about blood pressure?<br>**P:** OK, we can see the pressure | **Action:** Exp.1 takes mouse, clicks BP link.<br>**Gesture:** In response to question, P. nods, 'yes' and adjusts eyeglasses. | System goes to BP page | Patient is still somewhat disoriented by tabular representation and experimenter tries to orient her. |
| 48:52 | **P:** This, this… (doesn't complete sentence). | **Gesture:** Exp.2 points with hand in general direction of screen. Moves hand up and down in front of the screen.<br>**Gesture:** P. strokes chin<br>**Gesture:** Patient points at table. | | Experimenter tries to orient patient to focus on BP values in table. |
| 49:16 | **R:** Can you tell us what you see here?<br>**P:** Here? | **Gesture:** Exp.2 points specifically at table at lower part of the screen.<br>**Gesture:** Patient points at BP values in one of columns. | | Patient signals that she has established a point of reference |
| 49:21 | **P:** Here, the numbers?<br>**R:** The numbers.<br>**P:** They are very small. | | | |
| 49:32 | **P:** Two, one, three.<br>**R:** February 13, that was the date.<br>**P:** So this is … | **Gesture:** P. points to values on screen with finger.<br>**Gesture:** Exp.2 waves hand up and down to the side of the monitor. | | Patient reads the date as a set of discrete numerals, can't parse cell |
| 49:50 | **R:** Can you tell what these numbers represent?<br>**P:** Nine and four.<br>**R:** Blood Pressure | **Action:** Exp.1 takes mouse and uses cursor as pointer to indicate values highlighted in red. | Cursor moves to values | Patient experiences similar difficulty with time feld |
| | **R:** What do the numbers represent?<br>**P:** It is red. | **Gesture:** P. points at values with index finger | | Does not recognize that red denotes an abnormal value |

Fig. 6. Micro analysis of video transcript Subject 1.

the patient to the table of values on the Diabetes Manager page. The researcher is trying to provide guidance and orient the patient to interpret the blood pressure and glucose values in the table. At this point, the subject is still struggling to use the mouse and the experimenter is assisting her in carrying out various actions. The patient is clearly confused by the display, but the experimenters do not yet fully appreciate the nature of the difficulty. At time 40:02, the researcher probes as to the meaning of the blood pressure values and uses the mouse as a pointing device. The subject responds that "they represent health," signaling that she lacks a point of reference to understand the expression of systolic over diastolic (e.g., 212/89). The problem may reflect a lack of familiarity with the representational formalism as BP is indicated somewhat differently on the

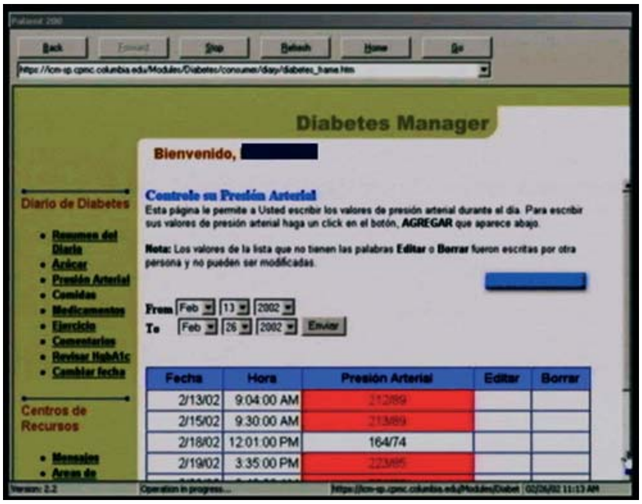


Fig. 7. The screen capture from Subject 1 examining her blood pressure values.

Fig. 8. The video of Subject 1 examining her blood pressure values (as indicated in Fig. 7).

monitoring device (i.e., systolic and diastolic presented as discrete values in a vertical orientation).

The second segment lasts for about 94 s and takes place about 7.5 min later after the researchers explained to the patient the various facets of the Diabetes Manager page. The screen (illustrated in Fig. 7), indicates the blood pressure tracking task. In this task, a subject can examine BP readings over a specified interval (e.g., a month). At 48:52, the researcher directs the gaze of the patient toward specific cells in the table representing blood pressure. The patient's pointing gesture indicates that she understands the common reference point (Fig. 8). The researcher probes as to the meaning of the values. At 49:21, the subject reads the date cell with value February 13, 2002 as "two, one, three" (a set of discrete values rather than a date). She has similar difficulty with the screen representation of the time. The subject was also not familiar with the convention that a value marked as red signaled an abnormal reading.

Over the course of this segment, it was evident that the subject was not able to understand the values in a table and could not draw any inferences about her health status from these representations. Two other subjects had similar numeracy difficulties. In addition, several other subjects who exhibited higher levels of literacy also experienced problems with the tabular representations. These problems included establishing a correspondence between data on a monitoring device (e.g., BP) and their presentation in tabular format (systolic/diastolic) on the computer screen and recognizing cues that values were outside the normal range. Some subjects also had difficulty tracking values over bounded periods of times, for example, to compare patterns of results over different weeks. These are all core competencies in using a telemedicine system and more generally, reflect basic health literacy.

Many novices, including those who negotiate many of the tasks without difficulty, lack an expressive vocabulary for referring to objects on the screens (e.g., scroll bars). Analysis of the nonverbal aspects of behavior such as gesture, gaze, and body position change, as exemplified in this microanalysis, is essential for making informed judgments about the interaction. This analysis is a particularly important tool in understanding how aspects of systems, as instantiated in various representations (e.g., tables, charts, and graphs) and interfaces, can be used to mediate behavioral change or health-related decisions (cf. Fig. 9).

An interesting study in contrast is subject 2, a 74-year-old woman. Like subject 1, she had been in the IDEATel intervention for about a year and had no prior computer experience. However, she developed a remarkable mastery in a short while. Two brief excerpts from the coded video transcript are presented in Table 6. In the first segment, the patient is accessing the web (7:28) and in the second, the task is to explain the Diabetes Manager Page including her glucose and blood pressure results. The session is markedly different in several respects. The patient is in full control and needs no assistance or prompting. She has a remarkable facility with the mouse and a relatively robust mental model of the system (basic understanding of how it works), which enables her to effortlessly negotiate screen transitions and complete tasks. Interestingly, for most of the session, the patient had her hand on the mouse. The mouse was used as pointer to orient her to the details on the screen (such as values in her patient record) much in the way that one would use a finger to guide the reading of a book. The subject also used the mouse/pointer to focus the gaze of the researchers and establish common points of reference for purposes of discussion. The patient was a

| Time | Dialogue and Conversational Code | Action Code | System Response | Comments |
|---|---|---|---|---|
| 7:28 | **R:** Can you explain what you see now?<br>**P:** It is calling the Internet. There it turns off, and leaves me with the system. | **BPC**: Patient moves head, looks from side to side<br>**Gesture**: Points at screen with index finger<br>**Action**: Takes mouse and clicks security screen away<br>**BPC**: keeps hand on mouse for a couple of seconds, then off<br>**Action**: Takes mouse and clicks identification screen away. | Security screen/dialogue box appears<br><br>Sec. Screen goes away<br><br>Password screen appears | The task is to access the web and patient immediately assumes the control |
| 7:54 | **R:** So far, so good. | **Action**: Takes hand off mouse, hunches over keyboard, and puts in password.<br>**Action**: Takes mouse and clicks OK button and hand off mouse.<br>**Action**: Takes mouse in hand immediately as DM page comes up. | Password screen transitions to Diabetes Manager Page | Patient performs all actions without any prompting. |
| 12:03 | **R:** Can you explain the screen to me, not all the details?<br><br>**P:** OK, it says daily manager. This is the sugar. This says blood pressure. These are food, medications, exercises, comments; to revise and change dates. | **BPC**: Sits holding mouse and looking at screen.<br><br>**Action**: Clicks on downward arrow on scroll bar<br>**Action:** Brings cursor over to Messages link | Cursor moves over to Messaging | Experimenter asks patient to explain the different facets of the Diabetes Manager Page. |
| 12:35 | **P:** This shows what happens today. At 9.35 AM, the sugar this morning, and this is the sugar now. And this is blood pressure from morning, and this is for now. | **Action**: Brings cursor over to side of screen<br>**Action**: Takes cursor to first value<br>**Action**: Takes cursor to second value<br>**Action**: Clicks on scroll bar arrow | Cursor moves<br>Cursor moves to point at value<br>Scrolls up page | Patient employs the mouse both to orient herself and to focus the gaze of the experimenters on common points of reference. |

Fig. 9. Micro analysis of video transcript Subject 2.

high school graduate and literacy and numeracy were not issues.

The two subjects represent opposite ends of the continuum with respect to system mastery and facility. On the surface, there are many similarities between these individuals. They are both Hispanic women of similar ages living in Northern Manhattan. They do differ considerably in their level of education. In addition, the second subject had access to social support (a son who would sometimes provide assistance), whereas the first subject had none. The contrast highlights the remarkable heterogeneity evidenced by subjects in this study.

## 4. Conclusions

Telemedicine is an emerging technology initiative that promises to transform patient care. Thus far, the predominant focus of telemedicine evaluation has been on technical feasibility, cost effectiveness, and measures of health care outcomes. A few programs have also evaluated patient satisfaction with the system and changes in behavior as a function of participating in a telemedicine intervention [5]. However, systematic and comprehensive research on other aspects of accessibility, including its social, cultural, and psychological dimensions is lacking [5]. In this paper, we present a multifaceted approach to usability evaluation that incorporates a cognitive walkthrough and field usability testing in patients' home. The methods are illustrated in the context of an evaluation of a comprehensive diabetes telemedicine program. The approach is predicated on an in-depth understanding of both the tasks involved in the process and the intended target population. We argue that such research is needed to improve our understanding of the obstacles to effective use of technology by an elderly chronic-care patient population and more generally, for understanding how home health initiatives can more effectively use such technology.

Although we know relatively little about seniors' use of technologies, research on cognitive aging provides some insight into the challenges seniors confront in using these systems. There are age-related declines in psychomotor skills, especially in dexterity and hand–eye coordination. There is some evidence that these physical limitations can impair individuals from learning to use a keyboard and mouse [26]. Age reduces processing capacity as measured by working memory [4]. Older adults are more affected by distracting context (e.g., clutter on a screen) and this limits their ability to selectively attend to relevant screen features and perform concurrent tasks (e.g., work on a computer and hold a conversation). Research by Rogers and Fisk [27] indicates that seniors are limited in their ability to develop automated responses. As a consequence, actions that become highly automated for younger adults may continue to exert a substantial cognitive load for seniors. Despite these age-related declines, older adults are remarkably adaptive and can continue to perform at a high level [4].

The cognitive walkthrough provides a meaningful measure of task complexity and a means to anticipate potential user problems. The method is predicated on a sound understanding of the target population. Field usability research provides a window into the process of human–computer interaction in a natural setting (i.e., a patient's home) under realistic conditions that approximate the actual context of use. The CW predicted certain patterns of difficulty with the system, but field usability testing revealed a host of other problems, many of which were not anticipated by the developers or evaluators. Usability evaluation sheds light on dimensions of a system that erects barriers to fruitful, efficient, and safe use of products.

An analysis of a system is only half of the battle. Users vary considerably in terms of their knowledge, competencies and other personal attributes (e.g., self-efficacy). There are also extraneous variables such as social support and patient health that impact the success of a given program. Field research can provide a remarkably vivid portrayal of how these variables shape uptake of the intervention by different individuals.

Field usability testing is perhaps the most expensive and time-consuming of the available usability methods. In a mature field of information technology application (for example, where the target population and domain of use is well understood), field research may not be cost effective. However, in the exploding area of Internet-mediated health care, there are many issues that warrant the kind of close scrutiny only afforded by this kind of research. As e-health initiatives continue to proliferate, usability inspection and user testing methods should play an increasingly important role in characterizing the obstacles to safe and productive use of home-health technologies.

A recent Institute of Medicine report [28] suggests that a "profound cultural change" is needed to enable patients to play a more active role in the management of their chronic conditions. In order to achieve such a change, the health-care system needs to foster a supportive environment that offers ready access to reliable and understandable sources of clinical knowledge and "actively encourages health literacy" by providing patients with relevant information. For chronic conditions, patients themselves become the caregivers and assume substantial responsibility for their own health care. Patients can be taught proper management of diet, exercise regimen, self-monitoring of blood pressure and glucose, and adherence to medication regimens. Telemedicine is a medium that can serve as an engine for the envisioned profound culture change in chronic-care management.

In the IDEATel project, the results of this research led to software changes, development and subsequent revision of a patient tutorial and the creation of a field training program. The ultimate objective of this work is to develop a comprehensive design and evaluation framework for enabling seniors to more effectively participate in Internet-mediated health care initiatives. However, there exist significant cognitive, perceptual-motoric, scientific literacy and innumeracy barriers that preclude some patients from fully exploiting the benefits of web-based telemedicine. An in-depth understanding of these barriers is a prerequisite for tapping into the vast potential of such innovative interventions.

## Acknowledgments

This research was supported by Cooperative Agreement Centers for Medicare and Medicaid Services (95-C-90998). We are grateful to all the participants who graciously gave of their time to participate in the study. In addition to the authors, the IDEATel Research Consortium includes the following: Columbia University and New York Presbyterian Hospital (NYC): Walter Palmas, MD, Lesley Field, RN, MSN, Rafael Lantigua, MD, George Hripcsak, MD, Soumitra Sengupta, PhD, Edward Shortliffe, MD, PhD, Mohammed Ashraf, Sean Mota, Catherine Tuck, MD*, Patricia Kringas, RN, Jessica Rivera, RN, Rene Bachmann, RN, CDN, CDE; Harlem Hospital Center (NYC): Linnea Capps, MD; St. Luke's-Roosevelt Hospital Center: Alejandro Prigollini, MD; Hebrew Home for the Aged at Riverdale (Bronx, NY): Douglas Holmes, PhD, Jeanne Teresi, EdD, PhD, Stephanie Silver, MPH; Jian Kong, MS, Eva Petkova, PhD, Lucja Orzechowska, MA, EdM, Joseph P. Eimicke, BS; SUNY Upstate Medical University at Syracuse (Syracuse, NY): Roberto Izquierdo, MD, Paul E. Knudson, MD, L. Thomas Wolff, MD, Suzanne Meyer, RN, CDE, Suzan M. Carusone, RN, CCRC, Fraser G. Tudiver, MD**, Teresa J. Wagner, MS, Carina Lagua, MA, RD, CDE, Susan Fox, RN, Susan West, RN, MSN, CDE, JoAnn Shupe, AOS, Donald R. Roller, BS, Joseph Ziemba, MBA, David P. Thomson, MD, Michele Malone, AOS, Kelly Duncan BSW, CCRC, Wendy VanRiper, Sarah Doolittle, BS, CCRC, Joanne Kearns, MA, RD, CDE, Deborah Bowerman, RN, CDE, Phyllis T. Bazen, MSN, ACNP-C, Pamela Suddaby, RN, BS, Kristine M. Brindak, RN, Fronda Benson, LPN; A.L. Lee Memorial Hospital, Inc. (Fulton, NY): Shirley Seabury, MHA; Arnot-Ogden Medical Center (Elmira, NY): John F. Carroll, MD, MBA, Frederic P. Skinner, MD, PhD, Terri L. Ennis, RN; Bassett Health care (Cooperstown, NY): Kathleen Sellers, RN, PhD, Barbara Leo, BA, BS, RN, Elsie Haus, MEd, FNP, Catherine Cooney, BS, MA, RN; Delaware Valley Hospital

(Walton, NY): Michael J. Freeman, DO, MPH; Family Health Network of Central New York, Inc. (Cortland, NY): Susan Ulrich, MS, RN, CPNP; Guthrie Health care System (Sayre, PA): Brian D. Cassetta, MD; Hudson Headwaters Health Network (Glens Falls, NY): Patricia Auer, MSN, FNP, W. Emily Hammond, AAS, RN, Cairenn L. Young, MSW, Mary M. Beadnell, LPN; Lourdes Hospital (Binghamton, NY): Timothy C. Howland, MD; Olean General Hospital (Olean, NY): Robert Catalano, MD, Galvin Anderson, MD, Terri N. Kolasinski, RN; Samaritan Medical Center (Watertown, NY): Anil Shah, MD, Kate R. Fraczek, RN, Terry Flynt, BSN, RN; State University of New York at Buffalo (Buffalo, NY): John Taylor, MBA, Lurene R. Tumiel, PhD, Marc Shilling, MA, MPA, Chester Fox, MD, United Health Services (Binghamton, NY): David Kwiatkowski, MD; American Telecare, Inc. (Eden Prairie, MN): C.R. Abbruscato, MSEE, Karen Boril. *Deceased. **Current address: E. Tennessee State University (Johnson City, TN).

## References

[1] National Academy on an Aging Society Report, 1999. Chronic conditions: a challenge for the 21st century.

[2] Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002;288:2469–75.

[3] Wagner E, Groves T. Care for chronic diseases. BMJ 2002;325:913–4.

[4] Fisk AD, Rogers WA. Health care of older adults: the promise of human factors research. In: Fisk AD, Rogers WA, editors. Human factors interventions for the health care of older adults. Mahwah, NJ: Lawrence Erlbaum; 2002.

[5] Bashshur R, Reardon T, Shannon G. Telemedicine: a new health care delivery system. Annu Rev Public Health 2000;21:613–37.

[6] Shea S, Starren J, Weinstock R, et al. Columbia University's informatics for diabetes education and telemedicine (IDEATel) project: rationale and design. JAMIA 2002;9(1):49–60.

[7] Vicente K. Cognitive work analysis: toward safe, productive, and healthy computer-based work. Mahwah, NJ: Lawrence Erlbaum; 1999.

[8] Roth EM, Patterson ES, Mumaw RJ. Cognitive engineering: issues in user-centered system design. In: Marciniak JJ, editor. Encyclopedia of software engineering. New York: Wiley; 2002.

[9] Norman DA. Cognitive engineering. In: Norman DA, Draper SW, editors. User centered system design. Hillsdale, NJ: Lawrence Erlbaum; 1986.

[10] Horsky J, Kaufman DR, Oppenheim MI, Patel VL. A framework for analyzing the cognitive complexity of computer-assisted clinical ordering. J Biomed Inform 2003;36(1–2):4–22.

[11] Kushniruk AW, Kaufman DR, Patel VL, Levesque Y, Lottin P. Assessment of a computerized patient record system: a cognitive approach to evaluating an emerging medical technology. MD Comput 1996;13:406–15.

[12] Rubin J. Handbook of usability testing: how to plan, design, and conduct effective tests. New York: Wiley; 1994.

[13] Jordan B, Henderson A. Interaction analysis: foundations and practice. J Learning Sci 1995;4(1):39–103.

[14] Polson P, Lewis C, Rieman J, Wharton C. Cognitive walk-throughs: a method for theory-based evaluation of user interfaces. Int J Man-Machine Stud 1992;36:741–73.

[15] Patel VL, Kaufman DR. Medical informatics and the science of cognition. JAMIA 1998;5:493–502.

[16] Cimino JJ, Patel VL, Kushniruk AW. Studying the human–computer-terminology interface. J Am Med Inform Assoc 2001;8:163–73.

[17] Kushniruk AW, Patel VL, Cimino JJ. Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces. Proc AMIA Annu Fall Symp 2001:163–73.

[18] Mackay W. Ethics, Lies and videotape. In: CHI'95 Proceedings: Conference on Human Factors in Computing Systems: Mosaic of Creativity, Denver, CO, USA, 1995. p. 138–145.

[19] Stevens R, Hall R. Disciplined perception: learning to see in technoscience. In: Lampert M, Blunk ML, editors. Talking mathematics in school: studies of teaching and learning. Cambridge, MA: Cambridge University Press; 1998. p. 107–49.

[20] Kendon A. Do gestures communicate? Res Language Soc Interact 1994;27(3):3–28.

[21] Goldin-Meadow S. The role of gesture in communication and thinking. Trends Cogn Sci 1999;13(11):419–29.

[22] McNeill D. Hand and mind: what gestures reveal about thought? Chicago: University of Chicago Press; 1992.

[23] Starren J, Hripcsak G, Sengupta S, et al. Columbia University's informatics for diabetes education and telemedicine (IDEATel) project: technical implementation. JAMIA 2002;9(1):25–36.

[24] U.S. Department of Commerce. A nation online: how Americans are expanding their use of the Internet. Washington, DC. 2002.

[25] Hawthorn D. Possible implications of aging for interface designers. Interact Comput 2002;12(5):507–28.

[26] Bosman EA, Charness N. Age-related differences in skilled performance and skill aquisition. In: Blanchard-Fields S, Hess T, editors. Perspectives on cognition in adulthood and aging. New York: McGraw-Hill; 1996.

[27] Rogers WA, Fisk AD. Age-related differences in the maintenance and modification of automatic processes: arithmetic stroop interference. Hum Factors 1991;33:45–56.

[28] Corrigan JM, Greiner A, Erickson SM. Fostering rapid advances in health care: learning from system demonstrations. Committee on rapid advance demonstration projects: health care finance and delivery systems, Institute of Medicine, 2002.