**Judith A. Deutsch**
Research in Virtual Environment And Rehabilitation Sciences (RiVERS) Lab
Department of Developmental and Rehabilitative Sciences
University of Medicine and Dentistry of New Jersey
School of Health Related Professions
Newark, New Jersey
deutsch@umdnj.edu

**Jeffrey A. Lewis**
Research in Virtual Environment and Rehabilitation Sciences (RiVERS) Lab
Department of Developmental and Rehabilitative Sciences
University of Medicine and Dentistry of New Jersey
School of Health Related Professions
Newark, New Jersey
and
VR LAB
Center for Advanced Information Processing
Rutgers University
Piscataway, New Jersey

**Elizabeth Whitworth**
Department of Information Systems
New Jersey Institute of Technology
Newark, New Jersey

**Rares Boian**
VR LAB
Center for Advanced Information Processing
Rutgers University
Piscataway, New Jersey

**Grigore Burdea**
VR LAB
Center for Advanced Information Processing
Rutgers University
Piscataway, New Jersey

**Marilyn Tremaine**
Department of Information Systems
New Jersey Institute of Technology
Newark, New Jersey

Presence, *Vol. 14, No. 2, April 2005,* 198–213


# Formative Evaluation and Preliminary Findings of a Virtual Reality Telerehabilitation System for the Lower Extremity

## Abstract

Usability studies are an essential and iterative component of technology development and ease its transfer from the laboratory to the clinic. Although such studies are standard methodology in today's graphical user-interface applications, it is not clear that current methods apply to new technologies such as virtual reality. Thus experimentation is needed to examine what existing methods can be viably transferred to the new user-interaction situations. In this paper, 5 integrated interfaces with 3 simultaneous users are evaluated via a set of usability studies, which adapt traditional methods for assessing the ease of use of the interface design. A single expert domain user was run in an intensive study that examined the therapist manual and interfaces of the Rutgers Ankle Rehabilitation System (RARS). The interface and manual were extensively modified based on this evaluation. A second study involving 5 therapists was then conducted to evaluate the telerehabilitation component of the RARS system. In both studies, the tester and developer's observations, along with the session videotapes and therapist-user questionnaires, were triangulated to identify user problems and suggest design changes expected to increase the usability of the system. Changes that resulted from the analysis with the domain expert are described and recommendations for how to conduct usability studies in such multiuser remote virtual reality situations are proposed. Results from the pilot usability telemonitoring studies are also presented. The validity of usability studies in the development and refinement of rehabilitation technology is highlighted.

## 1 Introduction

A key usability engineering methodology used to evaluate how easily the user understands and employs the application is called a usability study. The iterative application of usability studies to the design process serves as an integral component to the generation of easier-to-use interfaces (Mayhew, 1992; Nielsen, 1993).

Standard methods for conducting usability studies are well documented and described in basic human-computer interaction textbooks (Preece, Rogers, & Sharp, 2002). Typical usability studies involve giving a user a real task to perform while recording the process in some fashion (usually videotaping) (Mantei & Teorey, 1988). Literature on usability-study operation has focused on the evaluation of web-page usage (Cato, 2001) and the evaluation of single-user interfaces (Mayhew, 2002). Assessing the usability of rehabilitation interfaces involving multiple users, virtual reality exercises, remote connections, and integrated processes still lacks standard approaches. A usability methodology involving virtual reality navigation tasks was recently proposed (Hix & Gabbard, 2002), which involved the traditional four stages: User task analysis, expert-guided-based evaluation, formative usability evaluation, and, finally, summative comparative evaluation. Elements of Hix's methodology are applied in this paper.

The design of any rehabilitation technology requires an assessment and understanding of how the technology will be used and deployed. This is especially critical when transferring technology from a research laboratory to clinical use, where the technology must be robust and easy to learn and use, since the focus is on the care of the patient. To achieve this goal in the development of the Rutgers Ankle Rehabilitation System (RARS), a set of specific methodologies called usability engineering was applied. This paper describes how usability testing and software design iteration were performed collaboratively by a group of engineers and clinician scientists. Key issues in this description are the adaptation of standard usability methodology to assess the ease of use of a virtual reality and related telerehabilitation system. Findings from a formative evaluation consisting of an early-stage usability evaluation with an expert domain user followed by pilot usability studies run on 5 physical therapists using the remote-monitoring system and the resulting modifications to the system are reported here.

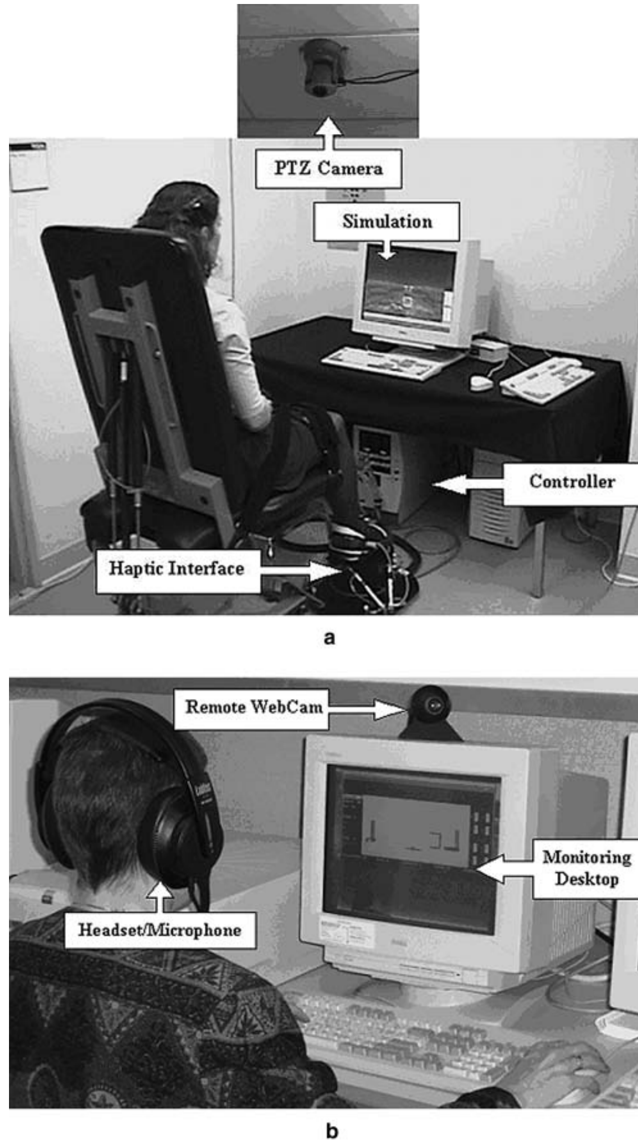



**Figure 1.** *RARS system overview: (a) patient's station; (b) remote-monitoring station.* © UMDNJ, Rutgers 2003.

## 2 Brief Description and Application of the Technology

The Rutgers Ankle Rehabilitation System (RARS) is a VR-based rehabilitation environment for patients having lower-extremity dysfunction. As shown in Figure 1a, the patient sits in the chair and places the affected lower extremity in a boot mounted on a platform that has six degrees of freedom. By moving their feet, patients manipulate a virtual object presented on the computer screen. In the particular VR rehabilitation application described here, the platform serves as a "joystick" used for flying an airplane through 3D targets. The therapist can set a variety of parameters that match the patient's therapy needs. In addition, the therapist can view these parameters remotely and interact with the patient through a telemonitoring system that includes a graphics display, a camera for observing the patient, and an audio channel for communication.

The RARS consists of local hardware and software components, as well as a remote-monitoring subsystem.

The key hardware element is a small parallel-kinematics robot in which patients place their feet. The main local software component is a library of virtual reality simulations written in WorldToolKit (Sense8 Co, 1998). The remote-monitoring station is a standard PC with a video, audio, and network connection with the local PC driving the simulation seen by the patient. The RARS includes a pan-tilt-zoom (PTZ) camera and software used for remote monitoring. The remote-monitoring screen runs a Java 3D (Sun Microsystems 2002) graphics applet, which follows, in real time, the local VR exercise.

A therapist in a remote location can manipulate the camera to see the patient or the patient's lower extremity (Figure 1b). The current exercise configurations and performances are detailed on the monitoring applet and mimic the information that the local therapist can set and examine both numerically and graphically. The center of the Java 3D applet shows several views of the ankle's relative motion to the leg as well as a 3D mock-up of the actual exercise simulation. The RARS and telerehabilitation subsystem are described in detail elsewhere (Girone, Burdea, Bouzit, Popescu, & Deutsch, 2000; Girone, Burdea, Bouzit, Popescu, & Deutsch, 2001; Boian, Lee, Deutsch, Burdea, & Lewis, 2002; Boian, Deutsch, Lee, Burdea, & Lewis, 2003; Lewis, Boian, Burdea, & Deutsch, 2003).

The system was developed collaboratively by a clinician scientist (JED) and several engineers (GB, MG, RB, JAL). It was designed for physical therapists to allow them to apply principles of exercise in designing a treatment session. Thus, with basic knowledge of the system and the parameters that can be modified to create exercises of varying difficulty levels, any physical therapist can use the system with a variety of patients. The system is currently designed for individuals who have lower-extremity dysfunctions that interfere with their functional mobility for use in a sitting position. It has been used to rehabilitate individuals who have had ankle sprains or fractures (Girone et al., 2000; Deutsch, Latonio, Burdea, Boian, 2001a), and poststroke individuals to address neuromuscular deficits that have interfered with ambulation (Boian et al., 2002; Deutsch, Latonio, Burdea, Boian, 2001b).

The software evaluated in this study consists of four interfaces used sequentially; to set a patient baseline (Figure 2a), configure an exercise session (Figure 2b), supervise and modify the exercise task in the virtual environment (Figure 2c), and monitor the patient remotely (Figure 2d). The user interfaces are listed below as they appeared in the therapist's training manual, which was the fifth interface evaluated in this study.

## 3 Testing Usability

Usability was tested in two phases, first in a formative evaluation by a domain expert and subsequently in a pilot study of the telemonitoring system using 5 physical therapists. The methods used in both phases of the study were the same. In this section of the paper we present (a) an overview of the study design, (b) challenges faced in testing usability, (c) detailed description of the study protocol and (d) how this study differs from other usability studies.

### 3.1 Overview of Study

Testing during the formative evaluation and remote telemonitoring pilot study was conducted over two sessions. The setting and terminology describing the study design are presented in Figure 3.

In the first session, therapist-users learned how to use the RARS at the rehabilitation site. They alternated between reading the manual (one of the interfaces), using the hardware, and practicing with the other four interfaces (shown in Figure 2). Training included: (a) first positioning a patient in the chair and establishing a baseline (shown in Figure 2a); (b) configuring an exercise (using the screen shown in Figure 2b) requiring setting parameters, which include: platform range of motion, platform resistance, exercise time, airplane speed, air turbulence, visibility, and presence or absence of haptic effects; and (c) monitoring exercise performance in real time (see Figure 2c). Therapists practiced on themselves first and then had an opportunity to try the RARS with an experienced user. The session was overseen at the rehabilitation site (also called the local

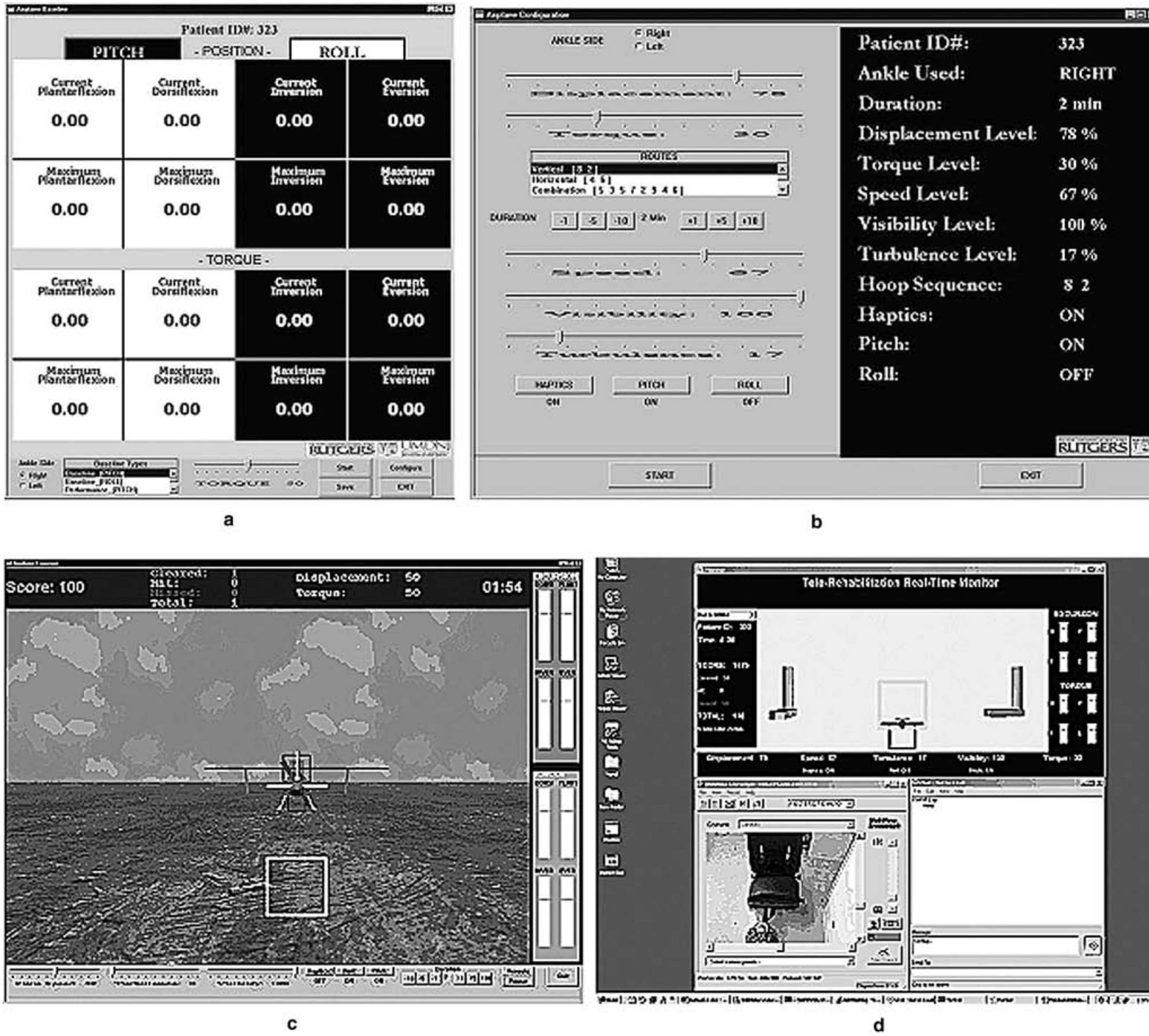
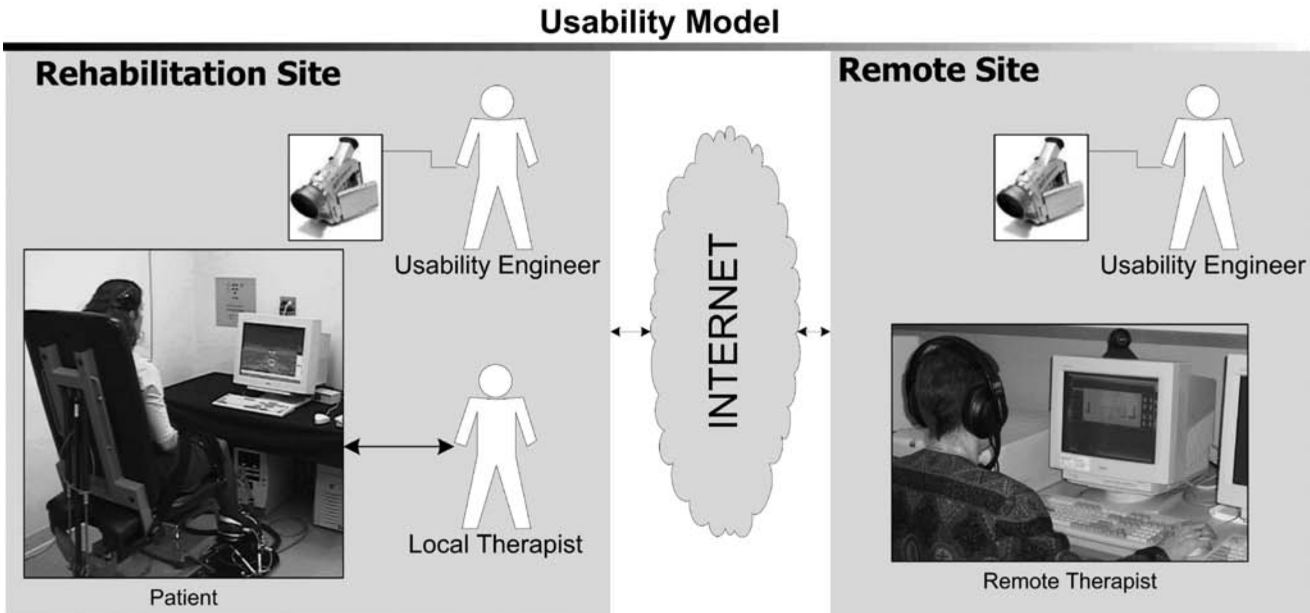



**Figure 2.** *The RARS software components (prepilot): (a) Baseline screen; (b) Configuration screen; (c) VR exercise simulation seen by the patient and local therapist; (d) screen view of the remote-monitoring station.* © UMDNJ, Rutgers, 2003.

site) by a usability engineer and at the remote site by a system developer (see table in Figure 3).

In the second session, therapists returned and were tested on their retention of knowledge about the RARS. Testing was done using a naïve patient at the rehabilitation (or local) site. Therapists were then asked to move to the remote site. An experienced patient and local therapist remained at the rehabilitation site, while the therapist-user and the usability engineer moved to the remote site. At the remote site, therapists were instructed in the use of the telemonitoring system and had an opportunity to remotely monitor the patient-user and local therapist at the rehabilitation site (see Figure 1b). Features of the system related to storing




| Session | Part | Patient | Local Therapist | Remote Therapist | Usability Engineer |
|---|---|---|---|---|---|
| Session I | Part a | Therapist-User | N/A | Therapist-Developer | Rehabilitation Site |
| | Part b | Experienced-User | Therapist-User | Therapist-Developer | Rehabilitation Site |
| Session II | Part a | Naïve User | Therapist-User | Therapist-Developer | Rehabilitation Site |
| | Part b | Experienced-User | Therapist-Developer | Therapist-User | Remote Site |

**Figure 3.** *The usability model for the rehabilitation and remotes sites: in Session I, instruction and practice with RARS at the rehabilitation site, and in Session Two, retention testing on RARS at the rehabilitation site, and instruction and practice with remote monitoring at the remote site.* © UMDNJ, 2003.

patient-exercise data in online databases and interpreting these data were not included in this study.

During both sessions a usability engineer interacted with the patient and therapist-user. Both sessions were videotaped. Questionnaires were administered at the end of each session to capture the therapist's impressions about ease of use of the system.

### 3.2 Specific Challenges Faced in Testing Usability

Typically a usability study involves users who are asked to interact with software and to perform a set of tasks that they would normally carry out using the system (Nielsen, 1993). A record is made of the users' interaction either via tester coding, audio recording, transaction logging, or videotaping. Often, more than one such data-collection method is used. In the present study the same procedures were used except that testing was performed simultaneously, or in sequence, on five integrated user interfaces (a training manual, baseline testing, exercise configuration, exercise testing, and remote monitoring) and with multiple users (a local therapist, a remote therapist, and a patient). Normally, each of the interfaces (except for the training manual) is tested in an unrelated usability study. However, given the manner in which rehabilitation is practiced, it was important to maintain the tight coupling between each of the interfaces and the patient's therapy. Therefore a set of integrated tasks had to be designed. These tasks related the user interfaces to each other. This created a problem of separating the effects of one interface study on the next interface evaluation. Because of this, methods were needed to sort out the interactions between the integrated user interfaces.

To run a realistic usability study, one needs to use real patients. Normally, usability engineers conduct usability studies, but such individuals would not have experience in patient therapy. Thus they would not be able to make corrective suggestions if the therapist-user in the study generated an error due to a system-interface misunderstanding. Having the therapist-designer of the application present solved this problem, but this approach is not normally accepted in usability evaluations because of the possible unintended effect on the user in the study.

A challenge, also unique to this study, was how to evaluate a remote-monitoring interface. Specifically, it was important to ascertain whether the remote therapist had enough information about the patient to accurately direct the remote therapy session. Part of the difficulty in evaluating the effectiveness and usefulness of the patient information provided through remote monitoring was that the therapist was new to a complex system. Here, the effect of being a novice RARS user is confounded with the usefulness of the remote information provided. Thus, the data collected on this user may not reflect the experiences of more experienced users.

A final challenge was the number of users involved in the interaction. The local sessions have two users, a therapist-user and a patient-user. The remote monitoring involves three users, the therapist-user at a remote location, a local therapist (or therapist assistant), and the patient-user. Although multiuser evaluation sessions encourage users to talk to each other about the problems they are having with the interface, some data are lost as the number of users increases, as one user may solve a problem before it can be observed and solved by the other two users (Sullivan, 1991).

### 3.3 Methods Used in the Present Study

**3.3.1 Sessions.** The two usability studies were conducted on the interface over two days and in several stages. The protocol was the same for both studies. In the first usability study, there was an expert domain user, in the second usability study, there were 5 physical therapists. A training manual was written by two system codesigners (JAL and JED), in which the RARS system and its telerehabilitation component are described. The manual was designed as a teaching tool to be used over two sessions. It was 30 pages long, with figures on the left side and text in bullet form on the right side. It contained questions and exercises to stimulate learning.

In the first session, the therapist-users read the manual and practiced on the system. Therapists completed the following sections: System Overview, Haptic Interface, Patient Set Up, Session Configuration, and Exercise Training. First, therapists were given an overview of the system and an explanation of the main exercise, which consisted of using the foot as a joystick to "fly" an airplane through successive targets (such as the light square shown in Figure 2c). Therapists were given the opportunity to "fly" the airplane to experience the exercise being administered to the patient. Second, therapists were instructed on how to set the system parameters. A pseudopatient who knew the VR system was used in this stage of the study. Therapists learned to use the system by first positioning the experienced patient in the chair and then establishing the baseline parameters (the screen for setting these parameters is shown in Figure 2a). The baseline parameters were captured by asking the patients to move their feet into plantar/dorsiflexion and inversion/eversion as far and as fast as possible in order to establish the patients' torque and range-of-motion capabilities. This is needed so that the subsequent exercises configured for the patient neither exceed the patient's capabilities nor fail to challenge the patient. Third, therapists then learned how to configure an exercise for the patient (the screen for this is shown in Figure 2b). This task included setting the platform range of motion, platform resistance, placement of targets alignment in space (horizontal, vertical, or combination), exercise duration, airplane speed, amount of air turbulence, amount of visibility, and presence or absence of haptic effects, for example, the sensation of a jolt when a target frame is inadvertently hit by the airplane. Varying these parameters allows the therapist to configure different types of exercises, such as strengthening or endurance, and varying levels of difficulty. Therapists were then given several clinical scenarios and asked to oversee the experienced patient using the simulation. At the end of the session, the therapist-users completed two questionnaires, a standard usability questionnaire (IBM Ease of Use Survey, 2003), and an instrument-specific questionnaire that tested knowledge about the RARS system.

In a second session, therapists were tested on their retention of knowledge by using a naïve patient. They were required to set the patient up, perform a baseline, and then configure exercises based on specific criteria (e.g., increase strength or endurance). Therapist-users reviewed the training manual for 15 minutes and were allowed to ask questions of the usability engineer about the manual content. Therapist-users then proceeded to position and train a naïve patient. This required setting a baseline and configuring several exercises. For part two of the second session, therapist-users worked at the remote site. Guided by the manual, therapist-users then followed the tutorial on the remote-monitoring system and had an opportunity to direct a session remotely. In the remote-monitoring session, therapist-users were instructed on how to use the PTZ camera to view different aspects of the local scene. They learned as well to communicate with the local therapist and the patient by using the audio system and the web-chat window. The local therapist established the patient's baseline and set up an exercise while the remote therapist monitored the session. Deliberate errors were made by the local therapist to probe the remote therapist's attention. A 10-minute break while the camera was set and communication was checked was provided before the remote therapists began a treatment session. The local therapist configured the baseline and the remote therapist instructed the patient on how to use the system and what to do during a strengthening and coordination exercise. The remote therapist was required to adjust the exercise parameters (by communicating with the local therapist) during the simulation to maximize the patient's performance. At the conclusion of the session, a usability questionnaire was administered for the remote-monitoring component of the system.

During both sessions, a usability tester (EW) interacted with the therapist and the patient while collecting video data from the screens and making notes of her observations. One of the system developers (JED) remotely monitored the sessions, also taking notes and serving as a resource to the tester if questions about the system and its use arose during testing. Throughout both sessions, users were required to talk aloud about what they were experiencing within the system. Talking aloud is a well-established method in usability testing that gives experimenters a better understanding of the user's thought processes. This method allows the capture of cognitive difficulties that occur in real time that may be forgotten, or even pass unnoticed by the user.

**3.3.2 Formative Evaluation.** Only a single expert domain therapist was used in this first usability study, designated as the formative evaluation. This person represented someone who was familiar with the types of therapy that RARS was designed to provide, and who could therefore give us feedback on not just how usable the interface was for her but also on how usable the interface was for the patient and for the integral task of the therapist working with the patient. In addition, this person, because of her knowledge of the field, could give us viable feedback on how well the user's manual would be received and understood by other therapists.

**3.3.3 Remote-Telemonitoring Pilot Study.** In the next set of usability studies that were run, 5 additional physical therapists were trained in the use of RARS following the protocol described above. The results of pilot usability studies conducted on the telerehabilitation part of the study are reported here. In this final stage that was run at a separate time from the initial training, the remote-monitoring interface was described to the therapist and used in a second exercise session with the pseudopatient. The remote-monitoring screen is shown in Figure 2d. Features and interfaces of RARS related to storing and interpreting the data from the exercises were not included in this study. The videotape of the expert user was reviewed for a gestalt impression of the therapist-users' problems, and data from the tester's notes, developer's notes, and user questionnaires were triangulated to identify the strengths and limitations of the five integrated user interfaces for the formative evaluation.

### 3.4 How This Study's Methods Differ From Existing Usability Studies

Although common usability methods were used in this study, as indicated earlier, the study involved several difficulties because of its multiuser, multi-interface, and remote-testing aspects. The common methods were the use of videotaping, experimenter coding during the study, and the completion of user-satisfaction questionnaires following the study. In addition to this approach, the following methods were applied: First the entire set of user studies was embedded into a single system manual, which was used as a guide for the therapist-user. Thus, the therapist-user could review other parts of the system during the process of learning the new features. The entire manual was used because this best simulated the real-life situation in a clinic where therapists have a single complete manual.

Second, the studies involving the local RARS components were run with one of the interface designers (JED) remotely monitoring the subject. Although remote monitoring has been conducted in other usability work, this is usually done when no experimenter is in the room. A key problem with application designers being present during usability studies is the possibility for their nonverbal and subverbal behavior to influence the user being tested. Nevertheless, in a study where therapy is being applied to another human through the interface, it is essential to have the advice of the therapist-designer, who has a deeper understanding of the therapy than a usability engineer. Having the remote-observation capability of RARS permitted the therapist-designer to be "present" without being in the same room and thus unduly influencing the subject in the study. An added advantage was the ability to have a second, remote, note taker.

Third, a set of true-false questions, to assess how much the therapist in the study understood about the operations of the system, was administered to the user. A verbal report of the problems with the system was recorded on video and through experimenter coding. However, this report did not ascertain what problems were due to a misunderstanding of the interface by the user, or due to the cognitive complexity of the task, that is, handling a patient, and, in the remote session, communicating to an assistant therapist. Thus, although cognitive-comprehension questionnaires are used in usability work, this set was used specifically to evaluate the depth of problems found in the video and coded data capture.

Finally, the therapist-user was required to explain the usage of the system to both the patient and to the local therapist in the study. This is the most unique of the new approaches that were taken in this study. Having a user explain what this person has just learned can give a usability engineer much more detail about what is misunderstood, omitted, or misrepresented by the user. The data captured with this technique was especially useful for the evaluation of the remote session. Because the remote therapist-user not only explained to the local therapist what steps to take in adjusting therapy, but also gave reasons for the system adjustments, it could be determined how well the therapist-user understood the patient-user problems via the remote monitoring.

## 4 Results

### 4.1 User Questionnaires

**4.1.1 Formative Evaluation.** Table 1 summarizes the expert domain's ratings of the system's ease of use during the formative evaluation. The therapist-user generally rated the RARS system positively on "comfort of use," "ease of learning," "ability to effectively complete work," "organization of information on the system," "enjoyment and pleasantness of interfaces," and "overall satisfaction." These items received scores between 6 and 7, with 7 being the highest rating on the scale. Items that received lower ratings were "ability to complete work efficiently and quickly," "the system's ability to signal the user when they have made an error," and the "information provided for the system being easy to understand." These features received scores between 3 and 4.

The therapist-user's response to open-ended ques-

**Table 1.** *Formative Evaluation Questionnaire Scores by Category*

| Categories | RARS | Monitor |
|---|---|---|
| Comfort of use | 6 | 6 |
| Ease of learning | 6 | 6 |
| Ability to effectively complete work | 7 | 6 |
| Organization of information on the system | 7 | 5 |
| Enjoyment and pleasantness of interfaces | 7 | 5 |
| Ability to complete work efficiently and quickly | 4 | 5 |
| The system's ability to signal the user when they have made an error | 3 | N/A |
| Information provided for the system being easy to understand | 4 | 5 |
| Overall satisfaction | 7 | 6 |

tions about the positive aspects of the system included "enjoyment of the exercises," "flexibility of the system that allowed customizing the exercises for the patient," and "the ease with which [he/she] learned the system." Accuracy for responses to the RARS true/ false questions was 87%. The results for the telemonitoring component were slightly lower than those for RARS as a whole, with 6 being the highest rating. The therapist-user rated the organization, enjoyment, and functionality of the remote-monitoring screen as a 5. Positive comments were made about the "opportunity to direct treatment from a remote location" and the "ability to monitor the patient from various viewpoints," including focusing on the area of treatment, in this case the ankle. In addition, the therapist stated, "With some clarification, the system is easy to figure out, especially with a prior training session." Specific comments on the negative aspects of the system were regarding the time lag of the voice feedback (microphones), the inconsistency of buttons between the local and the remote interface, and the difficulty of multitasking, especially during remote monitoring, where the therapist-user would be typing in the chat box, engaging in verbal interactions with the local therapist, and monitoring patient performance at the same time.

#### 4.1.2 Remote-Telemonitoring Pilot Study.

Table 2 summarizes the 5 therapists' responses to the usability of the telerehabilitation system. Responses were recorded using a 7-point Likert scale with "strongly disagree" rated as a 1, to "strongly agree" rated as 7. The lowest ratings were for the system's ability to provide information on errors. The highest ratings were for liking the use of the interface. There was variability in the responses, with ratings ranging from 2 to 7 for comfort and ease of use of the system.

Feedback was also elicited from the therapists using open-ended questions such as: "What were the most positive and negative aspects of the system?" Therapists commented on issues related to communication. They were distracted by the lag between the remote and the local site and complained about the quality of the audio communication. They stated as well that the web-chat feature was not responsive to real-time needs. One therapist commented that the multiple-screen layout was too complicated to manipulate, and another, similarly, said that a larger screen would be beneficial. Requests were made to have information available from the baseline configuration, which could not be monitored remotely, as well as to be able to remotely control the settings for the session.

### 4.2 Tester and Developer Observations Resulting in Interface Changes

This section describes the observations of the tester and developer coupled with support from the therapist-user comments made during the video sessions. These observations were used to make changes to the therapist-training manual and the interfaces. The manual lacked clarity and its construction produced cognitive overload for the user. To address these issues, the manual format was changed, reducing the information on each page and labeling and numbering figures to make it easier for the reader to follow sequential information (see Figures 4a, 5, 6b).

**Table 2.** *Remote-Monitoring Questionnaire Results (each initial corresponds to one of the 5 therapists participating in the study)*

| Usability of remote-monitoring items and ratings | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Overall, I am satisfied with how easy it is to use this system | | | | C | A E | D | B |
| It was simple to use this system | | | | C | A E | D | B |
| I can effectively complete my work using this system | | | A | C | B E | D | |
| I am able to complete my work quickly using this system | | | A | | B E | C D | |
| I am able to efficiently complete my work using this system | | | A | | B C E | D | |
| I feel comfortable using this system | | C | A | | E | D | B |
| It was easy to learn to use this system | | C | | A | E | D | B |
| I believe I became productive quickly using this system | | | A C | | E | B D | |
| The system gives error messages that clearly tell me how to fix problems | A | D | | E | | | |
| Whenever I make a mistake using the system, I recover easily and quickly | | A | C | B E | D | | |
| The information (such as on-screen messages and other documentation) provided with this system is clear | | | | A | E | B C D | |
| It is easy to find the information I needed | | | | E | A | B C D | |
| The information provided for the system is easy to understand | | | | E | A | B C D | |
| The information is effective in helping me complete the tasks and scenarios | | | | A | B C E | D | |
| The organization of information on the system screen is clear | | | | | A C E | B D | |
| The interface of this system is pleasant | | | E | | B C | | A D |
| I like using the interface of this system | | | | | C E | B D | A |
| This system has all the functions and capabilities I expect it to have | | A | | B | E | C D | |
| Overall, I am satisfied with this system | | | | B | E | C D | A |

Terminology was inconsistent across four of the interfaces and was not the preferred terminology of a physical therapist. For example "Displacement" was used when the term "Range of Motion" was preferable as the common term used in physical therapy (Figures 4a, 5, 6a,b). Inconsistent placement of the same terms in different parts of the interface also caused confusion. For example, the term "pitch" preceded the term "roll" on two of the screens, and then the order was reversed to "roll" and "pitch." This proved to be distracting and was corrected (see Figure 6c).

The "mapping" between the buttons and the tasks they performed (see Figure 4b) for the buttons labeled "Start," "Save," "Configure," and "Exit" was unclear to the users. It was unclear partially because the buttons were designed with words that involved large conceptual meanings, for example, "Start" seemed to have the same function as "Configure," and "Save" seemed to

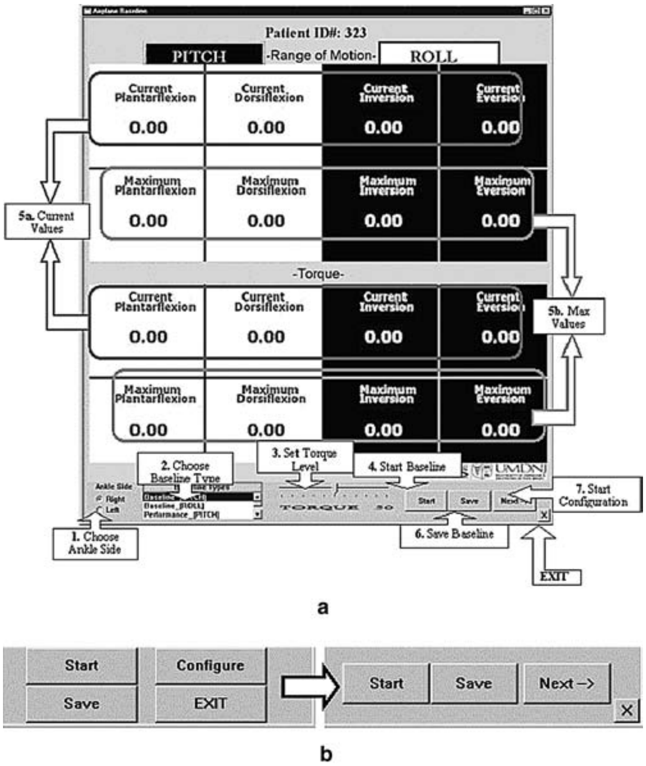



**Figure 4.** *Interface modifications: (a) baseline diagram in manual clarifies the screen's elements; (b) command modification of baseline screen (the "Configure" button was changed to "Next" and the "EXIT" button was minimized).* © UMDNJ, Rutgers, 2003.

have similar functionality to "Exit." These buttons represented a sequence of tasks that were to be performed by the user to establish the baseline parameters for the patient. Some of the buttons had no meaning until a portion of the task was accomplished. Thus, it would have been more appropriate to design the buttons so that they matched the sequence of procedures the user was to take when setting the baseline, and even to change their size and accessibility to make the process visible to the user.

Making procedures more apparent within the interface by reorganizing the presentation of the functions solved this command-structure problem. The new design involved three large buttons, "Start," "Save," and "Next," in a line, and a small "X" button to quit the screen (see Figure 4b). The order and direction of the buttons in the revised interface correspond to the actions that the therapist-user is supposed to take to complete a patient baseline. The "Configure" button, which would take the user to the next screen to configure an exercise, was changed to "Next," being an easily understood term, and a more obvious indicator that the user would be taken to a new screen. The button that would exit the user from the screen was made smaller and separated from the other buttons, making it less likely that users would click on the exit button by mistake.

Another problem involved the toggle switch in the exercise screen, where the system started off paused so that the patients could orient themselves with the airplane they were supposed to pilot with their ankles, and be ready for the exercise. In the initial design, the user would have to press the "Pause" button again in order to start the system, which is not very intuitive. In order to make the process more intuitive, the previously large "Quit" button was reduced to a small "X" at the bottom right of the screen, and the space created used to include a large "Start" button (see Figure 6d).

Observations of the remote-monitoring component of RARS highlighted several of its limitations. There was a reliability issue with the audio communication and audio feedback. The order of the elements on the remote screen did not match that on the local interfaces, making it more difficult for the user to learn. The therapist-user did not consistently read the messages presented on the chat window, which is attributed to having to attend to three elements: the VR exercise simulation, chat window, and camera view of the patient. As a result of these difficulties, the terms used on the remote-monitoring station were altered and reordered before the 5 therapists were introduced to the system.

Modifications made after the formative evaluation were successful in improving use of the remote-monitoring system; however, other areas required change. To reduce confusion, a foot side view was added to each remote screen. The parameters were color-coded to identify those that could be changed (gray background) and those that could not (blue background) (Figures 7a and 7b). To increase the capabilities of the remote therapists, screens will be developed to allow for baseline and exercise configurations to be viewed and controlled re-

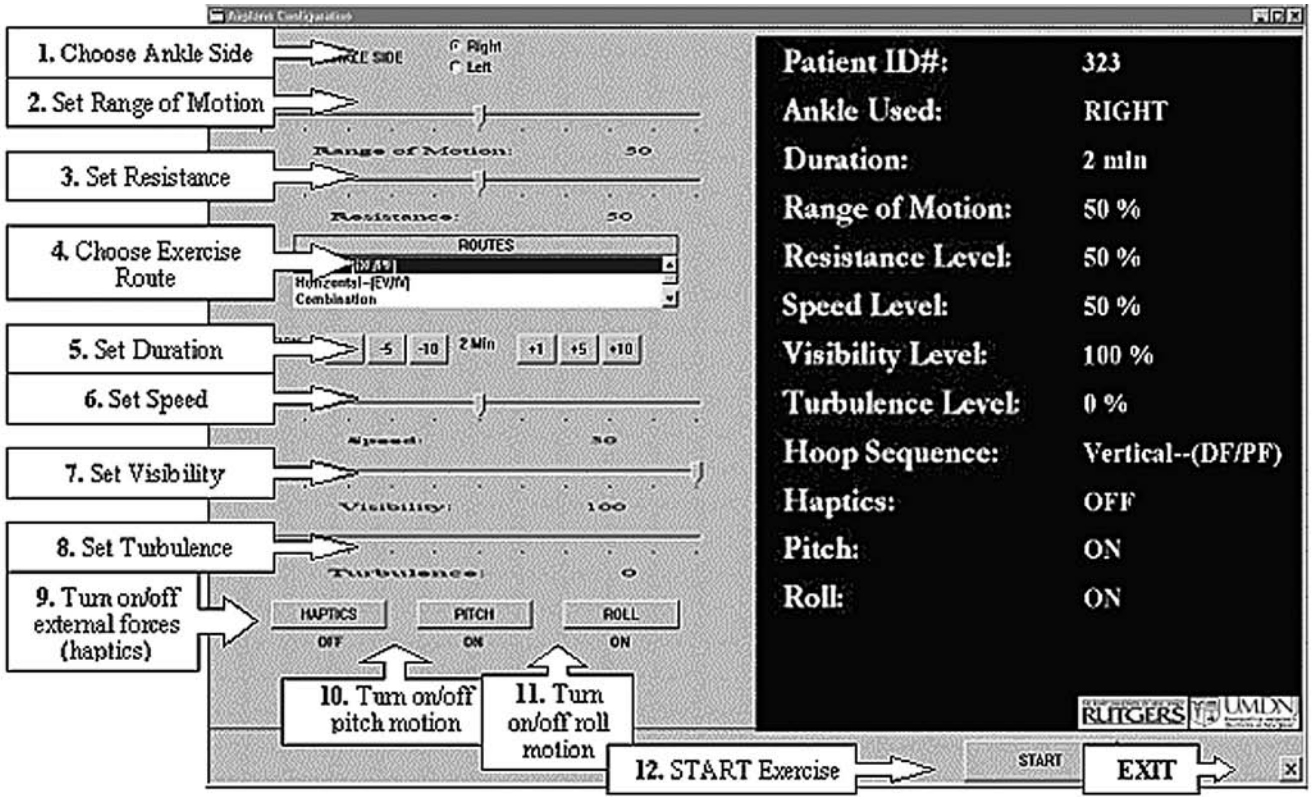



**Figure 5.** *Configuration diagram in manual clarifies the screen's elements.* © UMDNJ, 2003.

motely. The red letters on the black background will be revised for visual clarity. This change, as well as the new remote capabilities of the system, will be evaluated in future usability studies.

## *5* Discussion

A formative evaluation of the RARS was conducted adapting existing usability-study methodology. The findings from the tester's notes, videos, user questionnaires, and developer's notes found key problems, which required modifications. Changes were made to the user interfaces to be consistent with the therapists' language and use (Preece et al., 2002). Small order changes on the screen were made to reduce the difficulty the user-therapists had in finding information, therefore consistency in layout was implemented (Teitelbaum & Granda, 1983). The command structure was altered to better reflect the clinical decision-making process. The toggle switch, which started the machine in a state that the user did not specify and then required a toggle switch to alternate between states, was also modified (Norman, 1988). The most substantial changes were made to the therapist-training manual to reduce the confusion experienced by the reader.

While all these changes served to refine and ease the use of the system, it is noteworthy that the expert domain user rated her experience with the "comfort of use," "ease of learning," "enjoyment of the interfaces," and "overall satisfaction" as highly satisfying. The aspects of the system that received lower ratings were related to the training manual and its use. Importantly, the therapist-user was able to correctly apply principles of exercise to create effective exercise sessions for an experienced patient-user, and knowl-

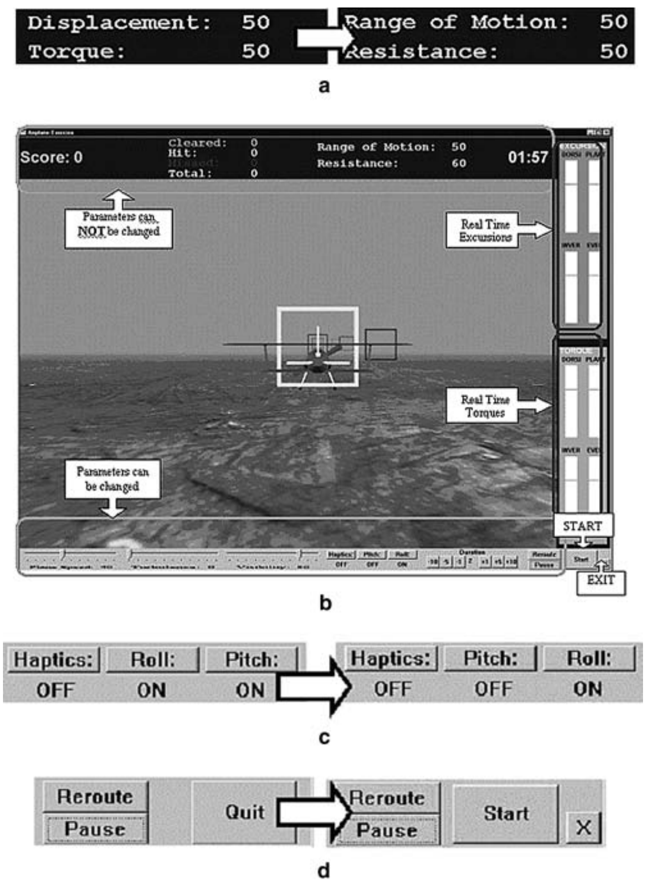



**Figure 6.** *Interface changes to the exercise screen following the study: (a) terminology modification; (b) simulation diagram in manual clarifies the screen's elements; (c) ordering modification for consistency; (d) command modification ("Quit" button was minimized and replaced with a "Start" button).* © UMDNJ, 2003.

edge retention of the system was 87% after using the system only once.

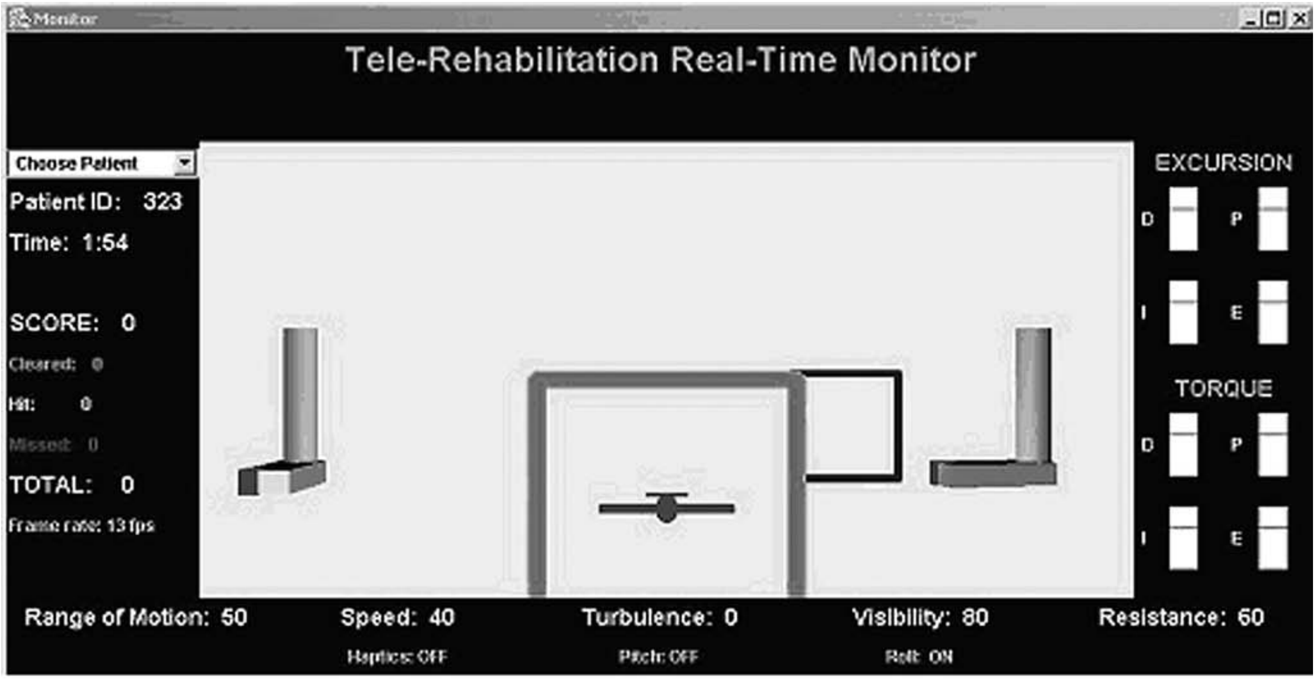
Substantial changes to the manual and user-integrated interfaces were made after the formative usability evaluation. Modifications to allow for language consistency made to the telerehabilitation component after the formative evaluation were successful. Subsequently, the 5 therapists in the pilot study of the remote-monitoring subsystem did not have complaints about language. This reinforces the value of an iterative process in developing interfaces and the usefulness of a formative evaluation.

Visualization of the multiple screens used to monitor a remote session will require improvements. The current system is not integrated, as the remote user has to open and monitor three independent screens. Possible solutions are synchronizing these screens as a single application that could be opened and resized simultaneously. In addition, more time allocated for therapist-user training may be required than that provided in this study.

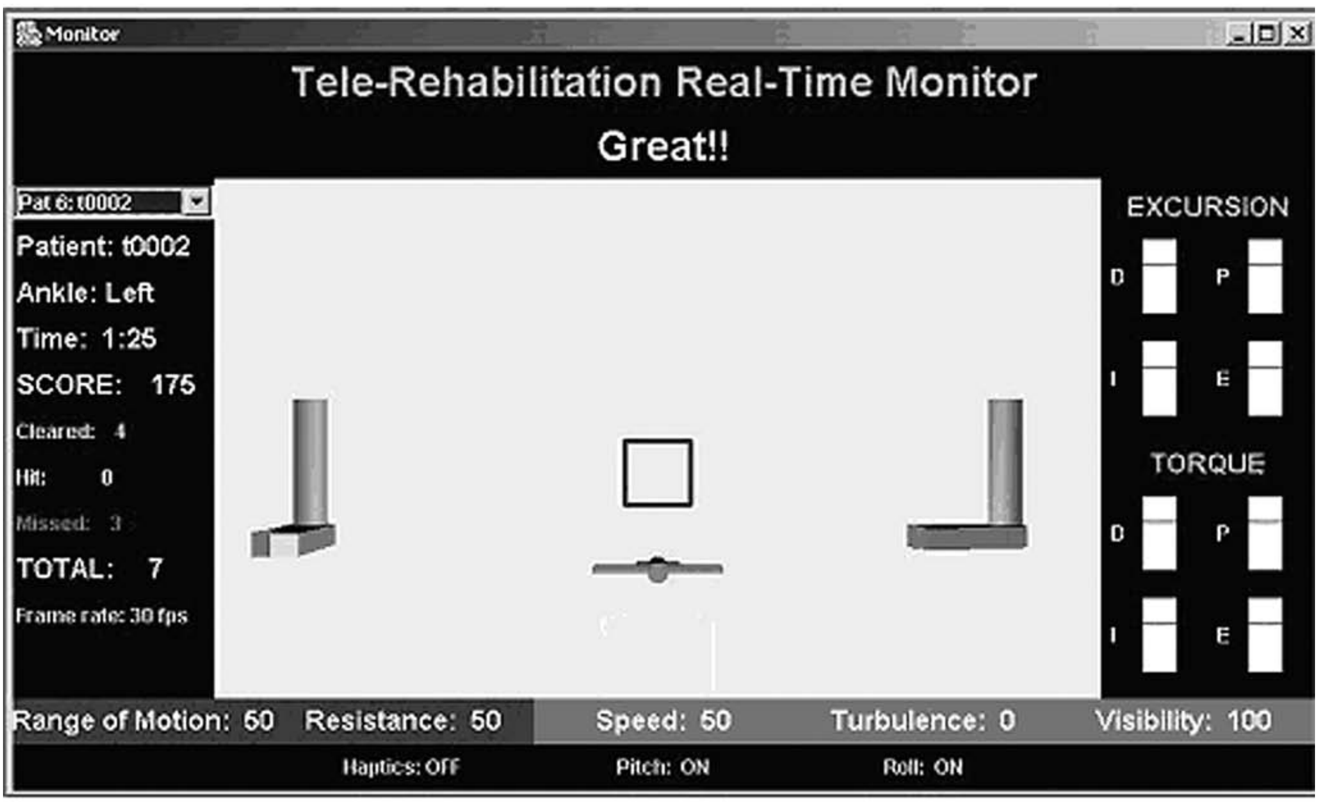
Communication with the local therapist and patient also remains an area for improvement. The system that was developed for these studies also relied on a chat mechanism that did not elicit the local therapist's attention and thus caused delays in communication. Although improvements to the audio communication were made following the formative evaluation (use of headset by remote therapist), the pilot study demonstrated the need for further refinement. Delays in the transmission of the audio communication could be attributed in part to the Local Area Network (LAN), which mediated the communication. The lack of reliable and fast communication infrastructure plagues telework systems in general, and it is expected that delays will diminish with the (future) introduction of faster and more reliable networks as part of medical informatics infrastructure.

### 5.1 Limitations of the Current Study

This study was a single formative-evaluation study in which usability of the RARS was tested using the advice of a single expert domain user and usability of the telerehabilitation component was tested using 5 physical therapists. Only the key features of the system were tested because the focus of the study was in understanding how the interface would work in the multiattentive task of patient therapy and monitoring. During the telerehabilitation tasks, video information was gathered only at the remote-monitoring site because of the shortage of qualified experiment personnel and equipment. Collecting data from the local therapist would involve collecting video data and having a second camera person present on the local therapist's side of the interaction. Collecting data from the patient-user within the same interaction would prove even more difficult. As




a




b

**Figure 7.** *Remote-monitoring screen modifications; (a) original monitoring window; (b) modified monitoring window; the font size of the display on the left-hand side was made larger. The parameters that could not be changed were highlighted in blue and those that could be changed were highlighted in gray. The order of the parameters was made consistent with the exercise screen (see Figure 6).* © UMDNJ, Rutgers 2003.

described earlier, it is important to hear users talk through their ideas. However, with both the patient-user and the local therapist in the same room, having both talk aloud would be cumbersome, if not entirely unnatural. Lack of sophisticated video capture equipment and trained users (e.g., the assistant therapist) to serve in secondary roles in the study restricted some of the control of the study and the information that could be captured.

An additional limitation is the design of the telerehabilitation study, which was conducted in a single session. The 5 therapists may not have had adequate time to learn the new system completely, and there was no opportunity to measure retention of knowledge for each individual stage of the system learning.

## 6 Conclusions and Future Work

The five integrated interfaces of the RARS were evaluated first by an expert domain user. Based on the expert therapist's report, the RARS and its telerehabilitation component were easy to use and consistent with her style of therapy practice. Many modifications were prompted by this formative evaluation, validating usability studies as an integral component of design and implementation of rehabilitation technology. In addition, the modifications made to traditional usability methods aided in capturing data necessary for the interface evaluation, and helped to explain the meaning of data that were confounded by the complexity of the study.

The remote-monitoring system was subsequently tested on 5 therapists from different practice areas and with varying levels of experience. In anticipation of this study, the training manual was modified for clarity. Modifications from the initial formative evaluation eased use during the subsequent part of the study. Areas that still need to be improved in the telerehabilitation component related to communication with the local therapist and patient as well as visualization of the remote-monitoring scene. Controlling all aspects of the session remotely will reduce the need for a local therapist and perhaps increase the capabilities of the system.

There remain other aspects of the RARS system and its remote-monitoring component that warrant further study. Patient data is stored by the system transparently in an Oracle database, which is then accessed by the treating physician or therapist using a web interface portal (Boian et al., 2002). The database portal and its ease of use to generate graphs and to follow patient therapeutic history remains to be studied. Another aspect that requires more attention is the patient's subjective evaluation of a telerehabilitation intervention in general and one using the RARS in particular. Finally, as the remote-monitoring station is placed farther from the local site, the influence of Internet-based communication will become an important focus area for further research.

## Acknowledgments

Research reported here was supported by grants from the National Science Foundation (BES-0201687) and the New Jersey Commission on Science and Technology (R&D Excellence Grant).

## References

Boian, R. F., Deutsch, J. E., Lee, C. S., Burdea, G., & Lewis, J. A. (2003). Haptic effects for virtual reality-based post-stroke rehabilitation. *Proceedings of the Eleventh Symposium on Haptic Interfaces For Virtual Environment and Teleoperator Systems,* 247–253.

Boian, R. F., Lee, C. S., Deutsch, J. E., Burdea, G., & Lewis, J. A. (2002). Virtual reality-based system for ankle rehabilitation post stroke. *Proceedings of the First International Workshop on Virtual Reality Rehabilitation (Mental Health, Neurological, Physical, Vocational),* 77–86.

Cato, J. (2001). *User-centered web design*. London: Addison-Wesley.

Deutsch, J. E., Latonio, J., Burdea, G., & Boian, R. (2001a). Rehabilitation of musculoskeletal injuries using the Rutgers ankle haptic interface: Three case reports. *Proceedings of Eurohaptics,* 11–16.

Deutsch, J. E., Latonio, J., Burdea, G., & Boian, R. (2001b). Post-stroke rehabilitation with the Rutgers ankle system—A case study. *Presence: Teleoperators and Virtual Environments, 10*(4), 416–430.

Girone, M., Burdea, G., Bouzit, M., Popescu, V. G., & Deutsch, J. (2000). Orthopedic rehabilitation using the 'Rutgers ankle' interface. *Proceedings of Virtual Reality Meets Medicine 2000,* 89–95.

Girone, M., Burdea, G., Bouzit, M., Popescu, V. G., & Deutsch, J. (2001). A stewart platform-based system for ankle telerehabilitation. *Autonomous Robots, 10,* 203–212.

Hix, D., & Gabbard, J. (2002). Usability engineering of virtual environments. In K. Stanney (Ed.), *Handbook of virtual environments: Design, implementation and applications.* Mahwah, (pp. 681–699). NJ: Erlbaum.

IBM. (2003). *IBM ease of use survey*. Available from http://www-3.ibm.com/ibm/easy/eou_ext.nsf/ Publish/2003.

Lewis, J. A., Boian, R. F., Burdea, G. C., & Deutsch, J. E. (2003). Real-time web-based telerehabilitation monitoring. *Proceedings of Medicine Meets Virtual Reality 11,* 190–192.

Mantei, M., & Teorey, T. J. (1988). Cost/benefit analysis for incorporating human factors in the software lifecycle. *Communications of the ACM, 31,* 4, 428–439. Paper republished in B. Boeinm (Ed.), *Software Engineering*. Los Angeles: Computer Society Press, 1989.

Mayhew, D. J. (1992). *Principles and guidelines in software user interface design*. Englewood Cliffs, NJ: Prentice-Hall.

Mayhew, D. J. (2002). *The usability engineering lifecycle: A practitioner's handbook for user interface design*. San Francisco: Morgan Kaufman.

Nielsen, J. (1993). *Usability engineering*. New York: Academic.

Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic Books.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: Beyond human-computer interaction*. New York: Wiley.

Sense8 Co. (1998). *WorldToolKit release 8 technical overview*.

Sullivan, P. (1991). Multiple methods and the usability of interface prototypes: The complementarity of laboratory and focus groups. *Proceedings of the ACM Conference on Systems Documentation,* 106–118.

Sun Microsystems (2002). The Java 3D API Specification, Java 3D vl. 3, June 2002.

Teitelbaum, R. C., & Granda, R. E. (1983). The effects of positional constancy on searching menus for information. *Proceedings of the ACM CHI'83 Conference on Human Factors in Computing Systems,* 150–153.