

Chapter 5: Usability Evaluation

Usability testing provides a systematic approach to the evaluation of human-machine interfaces. It comprises a set of techniques which can be implemented to support iterative design decisions. Results of usability testing include diagnosis of design problems, identification of avenues for problem solutions, and the determination of adequate specification compliance for the product being evaluated.

Usability methodologies have developed over the past three decades in the field of evaluation of human-computer interaction (Treu, 1994). Developers in the field have sought to produce computer systems and software that people find usable. While the approaches have been tested and validated through usability experimentation with human-computer interfaces, the concepts can be applied to any evaluation of a human-machine interface. The following evaluation was conducted with wheelchair operators to determine usability characteristics of the new wheelchair PAU.

Objective

The objective of the usability testing procedures was to evaluate the user interaction characteristics of the PAU. The specific objectives of the testing were to:

- diagnose problems with the PAU design which may interfere with user satisfaction and performance effectiveness.
- identify avenues to resolve these problems in future design iterations.
- rank the design problems identified in terms of importance and frequency in order to provide a priority list for design alterations.
- demonstrate that wheelchair operators are capable of using the PAU.
- provide a user satisfaction evaluation for the new PAU.

Usability Evaluation Tools

Several developed tools and methods for collecting data on human-machine interactions were utilized in the usability evaluation. User performance testing was the method employed to evaluate the new PAU. This evaluation method collected data by two general means: evaluating wheelchair users in the performance of defined tasks, and soliciting data through the use of a user satisfaction questionnaire (Hix and Hartson, 1993). There are numerous ways to procure information about the user-machine interaction while performing tasks. The following discussion explains each of the measurement tools implemented in the usability evaluation.

Direct Quantitative Measures

The attribute evaluated by the task performance measures was the "initial use" of the product. Other attributes which can be tested through this type of usability evaluation may include the learning rate, a preference over similar products, error recovery, infrequent use, etc. (Whiteside, Bennett, and Holtzblatt, 1988). Methods used to evaluate the initial use characteristics of the PAU are both quantitative and qualitative. Time to complete the performance of a defined task was the direct quantitative measure implemented in the evaluation. In addition, each subject provided pre- and post-task estimates of the time he or she would be willing to spend performing the task.

Qualitative data were collected through critical incidents and critical observations tools by means of verbal reports and verbal responses to impact analysis questions. These methods and tools are explained here.

Critical Incidents

Identification of critical incidents which occur during user task performance can be a useful qualitative data collection method. Incidents can identify elements within the human-machine interaction which may lead to negative results. Hix and Hartson (1993), define a critical incident as "something that happens while a participant is working that has a significant effect, either positive or negative, on task performance or user satisfaction, and thus on usability of the interface." Analysis of critical incidents involves identification of common features within incidents that classify design deficiencies (Del Galdo, Williges, Williges, and Wixon, 1986). In this way, the critical incident analysis was utilized to seek out aspects of the PAU design which should be improved with future design iterations.

Historically, the critical incident technique was initially used by Fitts and Jones (1947) in an effort to identify contributing factors to "pilot-error" experiences with aircraft controls. Others have modified and implemented the technique to study a variety of situations such as management concepts (Flanagan, 1954), anesthesiology procedural errors (Cooper, Long, Newbower, and Philip, 1982), and software documentation design (Del Galdo et al., 1986). The successful application of the critical incident technique to such diverse studies implies that it is a versatile tool which may be adapted successfully to varied product usability evaluations.

While critical incidents may be classified as either positive or negative, this study only collected data on negative critical incidents identified by subjects. Hix and Hartson (1993) point out that negative critical incidents help to indicate problems in the interface design. This problem identification is one of the main objectives of the usability evaluation.

Positive critical incidents, on the other hand, generally identify positive interaction design aspects. This type of data may be more useful in human-computer interface design than in mechanical prototype testing. For example, a positive design feature of a human-computer interface may be identified at one point in testing, and then implemented in various other appropriate places on the interface because it works so well. This type of data does not contribute significantly to the PAU design effort. Therefore, only negative critical incident data were collected during usability testing.

Critical Observations

Critical observations are distinguished from critical incidents by their source. While critical incidents are identified by the participant, a critical observation is an incident identified by an experimenter other than the subject performing the task. Both negative critical incidents and negative critical observations were utilized as means to collect qualitative data throughout the usability evaluation. The data were analyzed to help identify deficiencies in the design and introduce solution options for the problems.

Verbal Reports

The usability evaluation collected critical incident data by means of verbal reports. Subjects were asked to "think aloud" as they reviewed a videotape of their task performance. Types of information retrieved from these data included why the subject was having a problem, descriptions of design features contributing to the problem, what they expected to happen, etc. This verbal report technique was retrospective with respect to the physical actions evaluated, and concurrent with respect to a videotape record of the event. The videotape was reviewed immediately after the task was completed so that the experience was fresh in the subject's mind.

Verbal report data were collected after the actual task performance (retrospectively) for two reasons. The first reason was to eliminate any impact on the main quantitative task performance measure, time to complete the task. Karat (1988) explains that while thinking aloud generally does not change behaviors, it does usually slow down task performance. The second reason for collecting task retrospective verbal reports was to minimize investigator interference with the task performance. It is suggested by Karat (1988) that prompts are often required by the investigator to spur the subject to think aloud (this was found to be the case in this usability study). Such prompts may have interrupted subject performance of the defined tasks. Hix and Hartson (1993) specifically suggest that investigators should not prompt subjects for verbal reports when timing data are being taken.

Questionnaire

Further data on user satisfaction were collected by means of a questionnaire presented to the subject at the completion of the usability evaluation (shown in Appendix K: User Evaluation Questionnaire and Subject Responses).

Impact Analysis

Impact analysis is a tool which can be used to estimate the relative importance of design problems discovered in usability testing and present them in a rank order list (Del Galdo et al., 1986). Whiteside et al., (1988) explain that an impact analysis on data collected from tests with users can estimate which design choices will most likely advance achievement of usability goals. Problems can be prioritized by consideration of several variables including:

- the amount of time which the subject required to deal with the problem.
- a rating of importance or criticality of the problem (provided by either the subject or the design team).
- the frequency with which the problem occurred.

Two separate impact analyses were conducted with the wheelchair PAU usability tests. The first analysis ranked problems derived from incidents which have been identified by subjects through the critical incident data collection tool. This analysis considered both the frequency with which the incidents (from which the problem was derived) were identified across all subjects, as well as the criticality rating of the incidents. Ratings for an incident were only solicited from those subjects that identified the critical incident in the verbal report. The rating reflected the opinions of the subjects from the perspective of possible consumers of the product.

A second impact analysis was conducted on the problems derived from incidents identified through critical observations. Problems were ranked based on the same two factors in the first analysis: frequency with which the critical observation (from which the problem is derived) occurred across all subjects, and the criticality rating of the observation. In this case, the rating was provided from the perspective of the design team. One member of the team defined the critical observations and then three of the team members rated the criticality of the observations after the evaluation was completed. Data were obtained by reviewing recorded notes, the videotape, and the audio tape of the tests.

Due to the unobservable cognitive aspects of the problem-solving process, it was anticipated that accurate time estimates of time wasted on a particular problem were not possible through observation. Therefore, the amount of time which the subject spent dealing with a problem was not incorporated into the impact analyses.

Methods

Subjects

Subjects who represented possible consumers for the new device were selected. In this case, PAU consumers are also possible end users (manual wheelchair operators). PAU consumers as discussed throughout this dissertation may also include any purchaser of the product. See Appendix D: Assistive Device Consumers, for a more complete explanation of possible consumers.

In order to test the usability characteristics of the PAU attachment, detachment, and operation, subjects with the anticipated capabilities to perform these tasks were needed. The ability to operate the unit while attached required that the subjects have several upper-body capabilities. These included:

- manual dexterity sufficient to manipulate a finger trigger control.
- arm strength and flexibility sufficient to steer a wheelchair with a tiller steering handle attachment.
- postural stability sufficient to maintain stature while seated in a sling-type wheelchair seat without assistive restraints (e.g. seatbelts, molded posture supports, etc.).

The ability to attach and detach the new PAU required that the subjects have additional physical capabilities including:

- upper body flexibility and agility sufficient to lift objects off of the ground while seated.
- upper body strength sufficient to lift a nine pound battery several inches off of the ground.
- coordination and dexterity sufficient to align an electrical plug with a compatible socket, push it into place, and rotate slightly clockwise. It was assumed that this level of dexterity was also sufficient to align the device with the securing blocks on the wheelchair.

Subject Screening

Potential subjects were screened through a phone interview to determine if they matched the consumer profile for the PAU. Each potential subject was given a description of the device and manual wheelchair to be used in the tests. They were then asked several questions to determine if they anticipated that they were capable of performing the tasks associated with operating the new PAU. The questions were posed as follows:

- Do you think that you have the manual dexterity to sufficiently manipulate a finger trigger control such as the trigger on a power drill?
- Do you think that you have the arm strength to steer a wheelchair with a tiller steering handle similar to a rudder on a boat?
- Do you have postural stability sufficient to sit in a sling-type wheelchair seat without postural restraints such as seatbelts or supports?

The tasks were then described with examples in the following manner:

- Do you think that you are capable of lifting a wheelchair battery weighing approximately nine pounds several inches off of the ground?
- The power unit attachment process involves plugging an electrical plug into a socket and slightly rotating it clockwise. Do you think that you are capable of doing this?

It was further explained that the evaluation required the subject to transfer to and from a test wheelchair and to view a television screen. Each person was asked if there would be a problem performing these activities.

A potential subject who responded negatively to any of the screening questions was politely released from the subject pool. This occurred in only one case due to the careful explanation of participant requirements in advertisement literature. Potential subjects who responded positively to all screening questions were scheduled to participate.

Apparatus and Facilities

The usability testing was conducted at the Industrial Ergonomics Research Laboratory in Whittemore Hall on the Virginia Tech campus in Blacksburg, Virginia. The room provided an open area for testing which minimized the opportunity to collide with walls and other obstacles.

The following apparatus was used to collect data during the usability evaluations.

- One General Electric VHS video camera recorder model number 9-9806.
- One video camera tripod.
- One General Electric color television model number 8-2668.
- One four foot tall television cart.
- One cassette audio tape recorder.
- One clipboard.
- Data sheets and pen.

In addition, the following apparatus was utilized in the usability testing:

- One Everest and Jennings 1987 manual wheelchair model Ultra Lite Premier with the new power assist attachment.
- Two circular plastic obstacles approximately 10 inches in diameter.
- One tape measure.

Data Recording Apparatus Configuration

The video camera recorded the tasks from an optimum location for recording task performance. Efforts were also made to avoid interference with the tasks from the camera, tripod, or associated electrical attachments. The experimenter was located sufficiently out of the area of task performance. The camera and experimenter were oriented at different positions in order to obtain maximum exposure to the task events.

The experimenter maintained constant access to a clipboard which contained the data collection sheets and pens for recording data throughout the evaluations.

Specification of Procedures

Subject Greeting and Instructions

Each subject was tested individually at the defined testing facility. Upon arrival at the testing facility, the subject was greeted by the usability experimenter. First, the experimenter explained the outline of the procedures for the usability testing. The experimenter explained that an informed consent form must be read, agreed to, and signed before any of the testing could be conducted. At this time the experimenter explained that the subject had the right to stop the testing at any point without offering a reason. Next, the experimenter explained that the PAU to be tested would be demonstrated for them prior to any testing.

It was then briefly explained to the subject that the testing consisted of performing a series of predefined tasks while information was collected about the events. Also, the information collected would not be used to evaluate his/her skills, but could be used to improve the design of the PAU. The subject was then told that the procedures would be videotaped and that afterward, the videotape would be reviewed by the subject and the experimenter together to gather more information through discussion. Additionally, the videotape review sessions would be recorded on audio tape for future reference. Finally, the subject was told that he/she would be requested to fill out a questionnaire in order to receive his/her evaluation of the new PAU.

The experimenter then presented the informed consent form to the subject and asked the subject to carefully read the form and sign it if the subject agreed to proceed. A copy of the informed consent form can be found in Appendix J: Usability Evaluation Informed Consent Form. During this time, the experimenter remained close to the subject to answer any questions. However, the experimenter made an effort not to behave in any manner which may have been interpreted as impatient. Once the informed consent form was signed, the experimenter gave a copy of the form to the subject to retain for personal record.

Demonstration

Prior to the usability evaluation, the experimenter introduced the subject to the PAU with a demonstration. The following steps in the demonstration are listed in order of presentation:

- 1) Explain that the premise of the design is to add drive wheels which are electrically powered to the wheelchair.
- 2) Show the battery with the battery box and explain how it is wired to the motor.

- 3) Explain that the battery sling stays on the wheelchair when the unit is removed.
- 4) Demonstrate how the tiller column with the steering handle at the top, rotates.
- 5) Describe the connection between the outer column tube, the horizontal crossbars, and the wheelchair frame.
- 6) Explain the different column positions: power-drive, manual-drive, and transfer.
- 7) Transfer into the wheelchair and readjust to the power-drive position.
- 8) Discuss the on/off, forward/reverse switch and finger actuated speed control.
- 9) Demonstrate maneuvering including moving forward and reverse and turning.
- 10) Readjust the column to the transfer position.
- 11) Transfer out of the wheelchair.
- 12) Remove the tiller column and crossbars to demonstrate how it slides off of the wheelchair.
- 13) Explain that the securing blocks for the horizontal crossbars remain on the chair when the device is removed (the wheelchair can still be folded without obstruction).

User Performance Task Sequence and Grouping

A total of thirteen tasks was evaluated with the wheelchair operators. The tasks are listed here in the order performed.

- Battery Attachment Task.
- Battery Detachment Task.
- Column Attachment Task.
- Column Detachment Task.
- Transfer Into Wheelchair with the PAU in Place Task.
- Transfer Out of Wheelchair with the PAU in Place Task.
- Switching Column Mode: Transfer to Manual Task.
- Switching Column Mode: Manual to Power Task.
- Switching Column Mode: Power to Transfer Task.
- Straight and Clockwise Driving Task.
- Straight and Counterclockwise Driving Task.
- Forward Figure-Eight Maneuvering Task.
- Reverse Figure-Eight Maneuvering Task.

This order was selected to minimize time requirements for the testing.

Tasks were grouped as follows during the evaluation to avoid unnecessary repetition of procedures.

- Group 1: Battery Attachment/Detachment Tasks.
- Group 2: Column Attachment/Detachment Tasks.
- Group 3: Transfer Into/Out of Wheelchair with the PAU in Place Tasks.
- Group 4: Switching Column Modes Tasks.
- Group 5: Driving Tasks.
- Group 6: Figure-Eight Maneuvering Tasks.

General Procedure for all Tasks

The following description of the testing procedures applies to all tasks evaluated. The set of procedures was completed for each task group before the procedure for the subsequent task group was initiated. Specific procedures and apparatus configurations for the individual task groups are detailed following the general procedure description.

All equipment including wheelchairs, PAU attachments, etc., were placed in positions which provided optimum access prior to task performance. It was anticipated that convenient apparatus placement would minimize any opportunity for injury. Once the equipment was in place, the experimenter demonstrated the task procedure while verbally explaining the steps as they were demonstrated.

Following the demonstration, it was verbally explained a second time what steps were to be accomplished and in what order. The experimenter then asked the subject if he or she understood all of the steps to be completed and the order in which they were to be completed. The experimenter then placed the equipment back in the original positions and had the subject place his/her wheelchair or the test wheelchair (as appropriate) into the most convenient position for performing the task group. Prior to the task attempts, the experimenter asked the subject "As a potential consumer, how much time would you be willing to spend to perform this task?," for each of the tasks within the group. The responses were recorded on the data collection sheet.

The experimenter then instructed the subject that the first attempt at the task group was just practice and that he or she should feel free to ask questions and take his or her time. The subject was then asked to begin the sequence of events described in the procedure. During the practice tasks, the experimenter served to assist the subject by answering questions. Physical assistance was only provided when specifically asked for by the subject. In addition, the experimenter took notes with a paper and pen on any informative statements or actions. When it was determined by the experimenter that the subject would not be capable of carrying out the task, the experimenter terminated the procedure and proceeded to the videotape evaluation session.

After the practice sequence, the experimenter placed the equipment back in the original positions and asked the subject to repeat the task group procedures. During this completion of the task group, the experimenter measured the times to complete the individual tasks with a stopwatch. Time was measured from the point the subject first touched the equipment, to the point when the equipment was secured for the task (or as indicated in the task procedure).

If the experimenter detected frustration or fatigue from the subject, the subject was either asked if he or she would like assistance, or informed that he or she no longer needed to complete the task. In such cases, the experimenter reminded the subject that the inability to complete a task was the result of design problems with the power attachment, not a user inadequacy. Preset time cutoffs for performance were not used based on performance variation throughout pretest and early evaluations. It was found that the time required to reach a point of frustration or fatigue was too variable among the subjects to set specific limits. This was supported by findings throughout the usability evaluation.

Generally, a "Worst Acceptable Level" is calculated prior to evaluation to establish a consistent cutoff point at which the experimenter requests that the subject discontinue performing the task. Hix and Hartson (1993), define the "Worst Acceptable Level" as the lowest acceptable level of user performance. If there were time constraints on a specific task, this time would identify a specific boundary which should be measured against. However, no such time constraints exist with the PAU evaluation tasks and

therefore only fatigue and frustration criteria were used to establish the point for task discontinuation.

Once the subject had performed the task for a measured time, the subject was told how much time was required to complete the task. With this knowledge, he or she was again asked the question "Knowing that it required ___ seconds to perform the task, how much time would you be willing to spend performing the task as a consumer of the product?" This question was only posed to subjects who completed the task.

Task Performance Data Collection Sheets

The quantitative measures for the user task performance, time to complete the task, and pre- and post-task time estimates, were recorded on task performance data collection sheets. Table 7, titled Time Data Collection Sheet, shows how the information was recorded.

Measuring Concept. The first column in the sheet describes the measuring concept which was the task performed by the subject. The value measured is listed in the second column. This is the time to complete the task as listed in the tables. The next two columns of the tables are usability specifications titled "Current Level," and "Planned Target Level."

Current Level. Hix and Hartson (1993) define the "Current Level" usability specification as the "present level of the value to be measured for the usability attribute"(initial use). A specification of the "Current Level" offers a starting point for reference and ensures that the attribute is measurable. The method used for establishing the "Current Level" specification with the wheelchair PAU was design team performance with the prototype. Three design team members who were familiar with the operation of the PAU performed the specified tasks outlined for the usability evaluation. The three times measured to perform a specific task were then averaged to establish a "Current Level" reference time. The values calculated for each of the task performance times in the "Current Level" were then manipulated to set the usability specification at the "Planned Target Level."

Planned Target Level. Defined by Hix and Hartson as "the target level indicating attainment of unquestioned usability success," the "Planned Target Level" provided a goal for the measured times collected in the PAU evaluation. These are the levels of times which represent successful usability in terms of time required to perform tasks for initial use of the wheelchair PAU.

The "Planned Target Level" specifications were calculated with a technique adapted from a heuristic proposed by Hix and Hartson. The "Planned Target Level" was specified as twice the amount of time required for the design team to perform the task. Therefore, the "Planned Target Level" is set as the "Current Level" multiplied by two. This level served as a starting point to measure against.

Pre-Task Time Estimate, Observed Results, and Post-Task Time Estimate. Columns five, six, and seven of Table 7 were used to record times collected during the evaluation. The pre- and post-task time estimates represent the subject's verbal response to the question "As a potential consumer, how much time would you be

willing to spend performing this task?" The post-task estimate also included a statement telling the subject how long he or she required to perform the task. This performance time was recorded in the "Observed Results" column of the table.

Critical Incident Procedure

Following completion of the first group of tasks (battery attachment/ detachment), the experimenter explained and demonstrated the critical incident data collection procedure to the subject. The experimenter explained that the design team needs to understand the problems which users of the PAU are experiencing. It was carefully stated again that the purpose of the testing is to identify problems with the design and not to evaluate his or her capabilities as a consumer. Any difficulties which they may experience may also be experienced by future consumers and the design team needs to identify these problems.

The experimenter then explained that one tool which is used to identify problems is the verbal report technique. It was explained that:

- The technique involves "thinking aloud" while a procedure is conducted or remembered so that the experimenter can understand the way that the subject is thinking through the procedure.
- The subject will be asked to talk through the procedure which they just completed as it is displayed on the videotape.
- The design team is interested in everything that they are thinking, especially thoughts which may lead to incorrect actions or incidents which they may feel are problems with the power attachment.
- The words spoken will be recorded by a tape recorder for careful analysis at a later time.

At this point the experimenter demonstrated verbal protocol by "thinking aloud" through the procedure of placing a cassette tape in the audio tape recorder and rewinding the tape. This sequence was prerecorded on videotape so that the demonstration matched the verbal report task the subject was required to perform (thinking aloud while watching the video of the task performance). The subject was then asked to practice verbal protocol by "thinking aloud" through the practice procedure of attaching and detaching the battery. At this point the experimenter offered appropriate feedback to the subject to encourage optimal information collection. It was explained that information such as why an action was taken, what reaction they anticipated, aspects of the design he or she did not like, etc., should be included in the report.

Next, the experimenter asked the subject to "think aloud" as he or she reviewed the videotape of the subject performing the timed task. The subject was then told that he or she may request the videotape to be stopped or rewound at any point. This was to insure that there was an opportunity for the subject to provide all possible information they wished to contribute.

While the videotape was reviewed, the experimenter did not interfere as long as the subject was attempting to think aloud. Karat (1988) warns against intervention with prompting questions but suggests that an experimenter should be present to remind the subject of the task to "think out loud." Therefore, when the subject stopped talking

while the tape was playing, the experimenter prompted the subject with the following statements as appropriate:

"Tell me what you are thinking."

"What were you thinking at this point in the task?"

These statements were used when the experimenter felt that the subject was withholding information that may be useful.

At the conclusion of the videotape and the subject's verbal report, the experimenter asked the subject if he or she has more information to provide about the task. The subject was told that the videotape can be rewound to watch again to help the subject's memory if he or she wishes.

Impact Analysis

Once the data from the verbal report were collected, the experimenter identified which of the statements were classified as critical incidents. This was accomplished by stepping through the notes taken during the verbal report. Critical incidents of interest for the evaluation are defined as any statement indicating an element of the design or procedure which may lead to a negative effect on task performance or user satisfaction. The following list of examples demonstrates the type of statements which were interpreted as critical incidents.

- "tends to jerk when turning."
- "when I turn it that way it goes that way [the opposite direction]."
- "can't see what I'm doing."
- "forgot to turn it."
- "hard to reach both of them [clasps] with one hand."
- "hit my knee [on the column section]."

These statements represent both observable and nonobservable (cognitive) events which may negatively impact user performance or satisfaction.

Subjects were offered a chance to rest or tend to personal needs during the time required to identify the critical incidents. The experimenter then described the first identified incident to the subject and asked the subject to describe features of the design which led to the incident. The experimenter then explained to the subject that the design team would like them to evaluate the incident according to the following question:

"From your perspective as a potential consumer, how critical would you rate this incident in determining your decision to buy this product?"

The subject was shown the graphic rating scale in Figure 13 with the polar endpoints identified as "Critical" and "Non-Critical" (adapted from Del Galdo, 1986). This scale was visually available to the subject at all times during the critical incident review process. The subject was asked to identify the point along the scale which he or she felt each incident fit in terms of criticality of the incident with respect to the decision to buy the product. The experimenter recorded the responses on the critical incident data collection sheet (Figure 13). Subjects often chose a rating between two numbers, e.g. "2 or 3." This rating was then quantified as the midpoint between the numbers, the example rating is then 2.5.

<p>Incident:</p> <p>Task:</p> <p>1) Can you describe aspects of the design which may have led to the incident?</p> <p>2) From your perspective as a potential consumer, how critical would you rate this incident in determining your decision to buy this product?</p> <p>Non-Critical ----- -----Critical</p> <p>1 2 3 4 5 6 7</p>
--

Figure 13. Critical incident tool (adapted from Del Galdo, 1986).

After the subject provided a criticality rating for each incident identified in the review session, the subject then proceeded to the next task group. The videotape review, including critical incident identification and rating evaluation, was repeated following each task group performance in which the subject participated. After the evaluation was completed, all critical incident data for the subject were consolidated onto a sheet(s) for analysis purposes. The summarized data sheet is shown in Table 8: Critical Incident Data (a similar sheet was used for the critical observation data).

The thirteen tasks performed and evaluated are detailed in the following section. For each task group, the experimenter demonstrated the tasks and then instructed the subject to practice and perform the task group. The subject and experimenter then reviewed the subject's performance of the task group on videotape to identify and evaluate the critical incidents.

Attachment and Detachment of the Battery

Apparatus Configuration

The subject was seated in his or her personal wheelchair to perform this task group. The test wheelchair was placed adjacent to the subject's wheelchair in an opened configuration with the back of the wheelchair facing the subject. The covering flap (which covers the battery box when in place) was lifted out of the way of the battery

box. Brakes on the test wheelchair were locked to avoid unwanted movement of the wheelchair. The column section of the PAU was in the manual-drive position attached to the front section of the test wheelchair. The distance between the two wheelchairs was determined by the placement of the battery box as follows.

The battery box which contained the battery was placed on the ground between the two wheelchairs. The plastic battery box has a handle on top of the box for handling the battery, and a plug at the end of the box for wiring to the PAU. The battery box was oriented in the same manner that it assumes once it is placed into the battery sling on the wheelchair. The location of the box on the floor was approximately four inches from the end of the battery box to the end of the battery sling. The distance from the subject's wheelchair was arranged such that the handle on the box was at the most convenient distance for the subject to reach.

Procedure

- 1) The handle of the battery box was grasped and lifted into the battery sling on the wheelchair.
- 2) The battery box covering flap was pulled over the top of the battery box and secured to the lower battery sling flap with the clasp.
- 3) The clasp which pulls the sides of the battery sling together was secured.
- 4) The socket and wire located on the test wheelchair were plugged into the battery box plug located on the upper right-hand corner of the box.
- 5) At this point, the experimenter checked to see if the plug was sufficiently secured by actuating the finger trigger on the PAU.
- 6) The socket was pulled out of the plug on the battery box.
- 7) The clasp which holds the sides of the battery sling together was released.
- 8) The clasp which secures the covering flap over the battery box was released.
- 9) The covering flap was lifted off of the battery box.
- 10) The handle on the battery box was grasped and the battery box was lifted onto the ground.

Attachment and Detachment of the Column

Apparatus Configuration

The test wheelchair was opened and positioned with the brakes locked to avoid unwanted motion. The subject remained seated in his or her personal wheelchair and aligned the personal wheelchair to the position which provided optimal access to the front of the test wheelchair. The column section of the PAU was placed on the floor midway between the front vertical supports of the test wheelchair. It was oriented such that the column was parallel to the front-to-back dimension of the test wheelchair. Also, the column unit was facing the ground such that the crossbars could be lifted directly into position on the test wheelchair without the need to rotate the device into a different position.

Procedure

- 1) The column attachment was lifted into the lower securing blocks on the wheelchair frame. This was accomplished by first lifting the lower crossbar to a position above the lower set of securing blocks on the wheelchair. Next, the release mechanism

- (finger grips) on the lower crossbar was squeezed and the lower crossbar was dropped into the grooves in the lower set of securing blocks.
- 2) The plug which extends from the column section of the PAU was grasped and plugged into the socket located on the front section of the wheelchair. This motion required aligning the plug tongs with the socket holes, pushing the plug into the holes, and then rotating the plug slightly clockwise to secure it in place.
 - 3) At this point, the experimenter checked to see if the plug was sufficiently secured by actuating the finger trigger on the PAU.
 - 4) The plug was grasped and rotated counterclockwise to unlock, and then pulled from the socket.
 - 5) The finger grips on the lower crossbar were squeezed and the column section of the PAU was lifted up and out of the securing blocks.
 - 6) The column section was placed onto the floor.

Transfer to and from the Wheelchair with the PAU Attached

Apparatus Configuration

The test wheelchair was opened with the brakes locked to avoid unwanted motion. The armrest closest to the subject was rotated to the up position so that it did not interfere with the transfer. The subject initially transferred to a secure, padded bench to provide a stable surface from which to transfer. The subject instructed the experimenter how to position the test wheelchair in order to provide optimal access to the wheelchair for transfer. Prior to performing the transfer task, the subject prepared his/her body position for transfer. Footrests were left off of the test wheelchair to avoid unnecessary obstacles. The column section of the PAU was placed in the transfer position, i.e., rotated forward to its farthest point.

Procedure

- 1) The subject moved from the padded bench into the test wheelchair so that both legs were between the column of the PAU and the bench.
- 2) The inside leg was lifted over the column section of the PAU.
- 3) The subject shifted such that the most secure and comfortable position was obtained (this generally involved shifting back against the backrest).
- 4) The experimenter then instructed the subject when to begin transferring out of the test wheelchair.
- 5) The subject shifted to the optimal position for transfer.
- 6) The leg on the side opposite of the transfer was lifted over the column section of the PAU.
- 7) The subject moved from the test wheelchair to the padded bench.

Some subjects chose to perform the transfer into the wheelchair task by placing the leg over the column before shifting into the wheelchair.

Switching Column Modes

Apparatus Configuration

The subject was seated in the test wheelchair with the brakes locked while performing this task. Initially, the column section of the PAU was placed in the transfer position,

i.e., rotated forward to its farthest point. In this position, the motor and drive wheel section of the column were located to the anterior side of the column in relation to the wheelchair.

Procedure

- 1) The upper part of the column was lifted (rotated toward the subject) until the upper crossbar met the upper securing blocks. The inner column was rotated as it was lifted such that the motor and drive wheels were located at the posterior of the column in relation to the wheelchair. As the upper crossbar met the upper securing blocks, the right hand grasped and squeezed the release mechanism (finger grips) on the upper crossbar. This motion permitted the upper crossbar to slide into the upper securing blocks on the wheelchair.
- 2) The upper section of the column was pulled with the left hand while the finger grips were squeezed with the right hand until the column came to rest at the first detent in the securing blocks. This was the manual-drive column position.
- 3) The finger grips were released and an attempt was made to move the column forward and back. This motion tested to determine if the crossbars were locked securely into the detents. If motion occurred, it stopped when the column was moved into the correct position for the detent.
- 4) The experimenter acknowledged this position and indicated to the subject to proceed to the next column position.
- 5) The finger grips were squeezed on the upper crossbar with the right hand and the upper section of the column was pulled back into the second detent position with the left hand. This was the power-drive position. Again, the finger grips were released and an attempt was made to move the column back and forth. If the crossbars had not locked into the second detent, the column moved until it was secured.
- 6) The experimenter acknowledged this position and indicated to the subject to proceed to the next column position.
- 7) The finger grips were squeezed with the right hand while the upper section of the column was pushed forward out of the securing blocks with the left hand. The column was allowed to come to a rest in the forward position; the inner column rotated downward as the motor came to a rest under the column.

Driving Tasks

Apparatus Configuration and Pretask Practice

The subject was permitted a five-minute practice session with the wheelchair prior to the practice trials of the tasks. This first practice session took place in the center of the testing room as far from walls and obstacles as possible. During this time the subject was asked to practice with the controls and maneuvering of the test wheelchair with the PAU in the center of the room. The purpose of the session was to familiarize the subject with the apparatus prior to exposure with walls or defined tasks. This was a safety precaution implemented to minimize the likelihood of collisions or unwanted motions.

The subject was seated in the test wheelchair throughout these tasks. The PAU was placed in the power-drive mode and the brakes were released. A course was marked

with tape in the testing room according to Figure 14. The test wheelchair was placed at the point marked as the "Starting Point" before the first task was initiated.

Following completion of the driving tasks, two, circular plastic obstacles were placed in line with the starting point at the distances defined in Figure 14. These obstacles were utilized in the maneuvering tasks outlined in the following section.

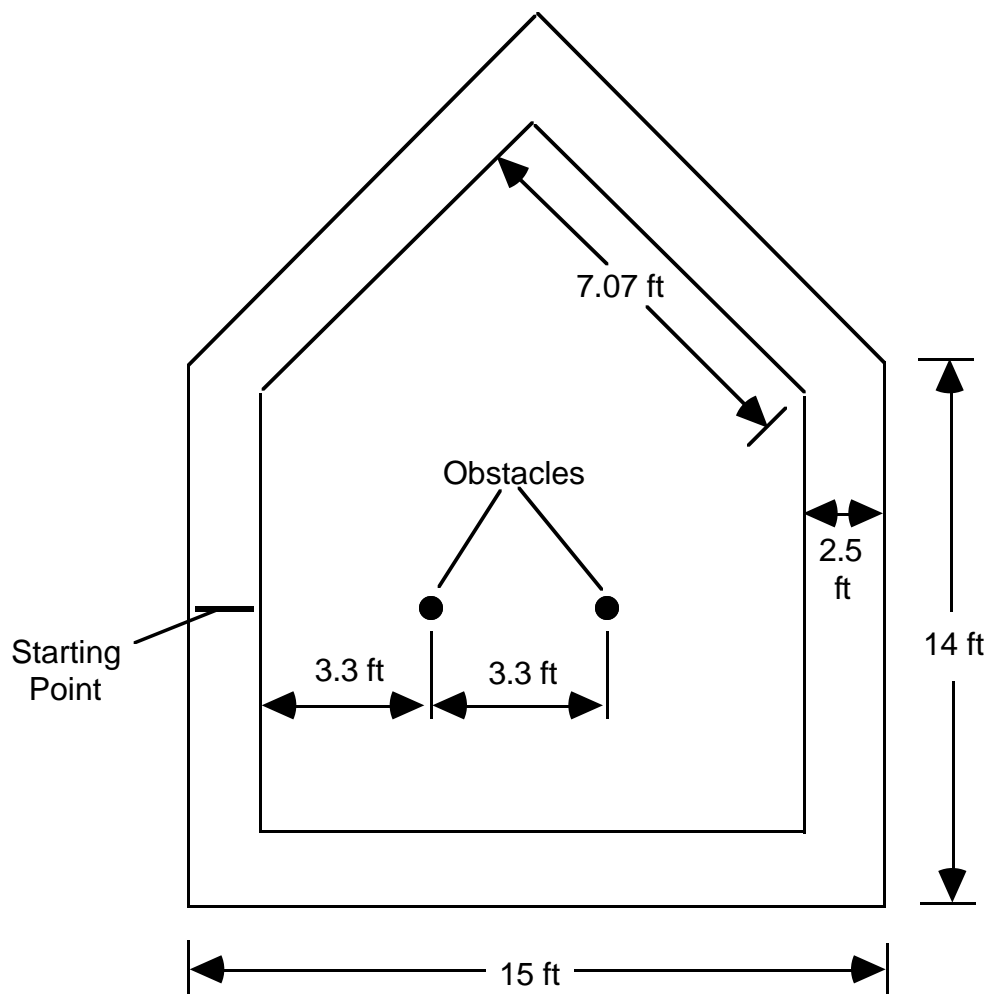


Figure 14. The driving and maneuvering task course (adapted from Lefkowicz, 1993).

Procedure

- 1) The subject maneuvered the wheelchair clockwise through the defined course attempting to stay within the course boundaries. The wheelchair was operated at a subjectively comfortable speed and stopped once the entire course was traversed once.
- 2) When the experimenter gave the instruction, the wheelchair was turned such that it was facing the opposite direction on the course (counterclockwise). The subject maneuvered the wheelchair so that it was at the starting point facing the new direction.

- 3) When the experimenter gave the instruction, the subject maneuvered the wheelchair counterclockwise through the defined course attempting to stay within the course boundaries. The wheelchair was operated at a subjectively comfortable speed and stopped once the entire course was traversed once in the counterclockwise direction.

Maneuvering Tasks

Apparatus Configuration

The subject remained seated in the test wheelchair throughout these tasks. The PAU was placed in the power-drive mode prior to initiating the first maneuvering task and the brakes were released. The test wheelchair was placed at the point marked as the "Starting Point" facing into the center of the course. In this position, the wheelchair was in line with the plastic obstacles which identified the maneuvering course (see Figure 14). The wheelchair was propelled in power operating mode around the obstacles in a figure-eight pattern as defined in the procedure.

Procedure

- 1) The subject maneuvered the wheelchair so that it was at the starting point facing into the center of the course.
- 2) The subject maneuvered the wheelchair around the two obstacles within the course in a figure-eight pattern. The wheelchair was operated at a subjectively comfortable speed and stopped when it again reached the marked starting point (the wheelchair facing away from the center of the course).
- 3) When the experimenter gave the instruction, the subject maneuvered the wheelchair, going in reverse around the two obstacles in a figure-eight pattern (traveling in reverse throughout the pattern). The wheelchair was stopped when the marked starting point was reached.

Questionnaire

After the tasks were completed and reviewed, the subject was given a questionnaire to complete. The purpose of the questionnaire was to collect data for a user evaluation of the new PAU. The questionnaire and associated answers provided by the subjects are listed in Appendix K: User Evaluation Questionnaire and Subject Responses.

When finished with the questionnaire, the subject was paid for his or her time and thanked for participating.

Pretesting

Three pretests were conducted for the usability evaluation. It was the purpose of the pretests to determine problems with the evaluation which needed to be corrected prior to actual data collection. There were several problems identified during pretest evaluations which involved a variety of evaluation components. These components were altered as needed in order to maximize the success of the evaluation. In addition, the pretests

provided an estimate of the time which was required to complete the tests as well as an opportunity for the investigator to practice.

The first two pretests involved completion of the usability evaluation with able-bodied subjects. Able-bodied subjects were chosen for the initial pretests so that the pool of subjects available for data collection was not diminished. Information collected from the first and second pretests was used to improve the evaluation design prior to the subsequent pretests.

The third pretest was conducted with a wheelchair user who successfully answered the subject screening questions. Only minimal changes to the evaluation were necessary following the third pretest, therefore, the data collected during the pretest were incorporated into the evaluation data.

Results and Data Analysis

Initial use characteristics of the user-PAU system were investigated by analyzing the user performance data of the defined tasks. Results have been analyzed to provide quantitative information on the time required to complete user performance tasks and on the estimated times users would be willing to spend performing the tasks. Qualitative analyses of critical incident and observation data produced two lists of identified design problems which were prioritized in terms of frequency and criticality. In addition, results of the user evaluation questionnaire were compiled along with additional verbal subject remarks to produce a report of subject design suggestions, usability comments, and an ease-of-operation rating for the PAU.

Removal of Data

The data from two subjects were removed from the usability evaluation analysis. Subject three and subject four were found to not fit into the predefined subject category (potential consumers of the PAU product in its current form). Either through misrepresentation or misunderstanding, each subject was not appropriately screened out of the evaluation. While subject three was able to complete four of the thirteen tasks, it was necessary to stop the testing due to safety concerns. Also, the performance during the four completed tasks did not represent the intended user population.

Subject four was not able to complete any of the specified tasks. After spending several hours with this subject, it was the experimenter's opinion that the subject was not able to fully comprehend the screening questions.

The investigator included any useful comments provided by the two subjects, and observations concerning their performance, in an auxiliary analysis. This analysis included useful design suggestions and comments provided by the able-bodied subjects during the first two pretests as well. These data were not included in the analysis of data from the remaining subjects.

Time to Complete User Performance Tasks

Data collected on the time to complete defined tasks were analyzed to demonstrate:

- which tasks, specifically, the subjects were or were not capable of performing with the PAU.
- the amount of time the subjects required to perform the specifically defined tasks with minimal training and exposure to the PAU (initial use).
- the percentage of subjects who performed the defined tasks within predicted task performance time ranges.
- estimates of the amount of time required to perform multi-task scenarios.
- how the amount of time the subjects required to perform the tasks compared to the amount of time the subjects estimated that they would be willing to spend performing the tasks.

Time to complete task data from all subjects tested are shown in Table 9: Time to Complete User Performance Tasks. For purposes of providing a complete record, data from subjects one, two, three, and four are included in the table but not in the calculations. From this table, comparisons have been made between the collected data and the usability specifications. These comparisons are represented by the calculated percentage of subjects falling into each of the three performance categories defined here. The categories are:

- not able to complete task.
- task time is greater than the "Planned Target Level" specification.
- task time is equal to or less than the Planned Target Level specification. This category represents the predicted time range for subject task performance.

The sum of percentages across the three categories may sum to a number slightly different from 100 percent due to rounding. Table entries labeled "not tested" represent a task which the subject was not given the opportunity to complete. This situation occurred three times during the evaluation when the investigator determined that it was not safe for the subject to proceed. In some cases, the not able to complete entry, NC, represents a task which the subject chose not to attempt because the subject did not believe that he or she was capable of completing the task.

The average time required to complete a task across subjects five through eleven was calculated and recorded in Table 9 under the column titled "S5-S11 Avg. Time." Only those subjects successfully completing the task were included in the calculation.

Reverse Figure-Eight Maneuvering Task. Twelve of the thirteen defined tasks produced average subject performance times less than the Planned Target Level times (established with the heuristic based on design team performance times). The Reverse Figure-Eight Maneuvering Task produced an average subject time of one minute and 28.8 seconds, while the Planned Target Level time for this task is one minute and 28 seconds. This result is basically equivalent to the Planned Target Level for the task, and the poor performance as compared to the other tasks reflects the difficulties experienced by the subjects with the task (analyzed in the impact analyses). More subjects were also not able to complete the Reverse Maneuvering task than any other task. Two out of the six subjects tested for the task (33.3 percent), failed to complete the course successfully. Two other subjects completed the course, but failed to do so within the Planned Target Level. Only two of the subjects attempting the course, were successful within the defined Planned Target Level.

Successfully Completed Tasks. 100 percent of the subjects tested were able to complete five of the defined tasks within the specified Planned Target Level. Performance on the five defined tasks: Transfer into the Wheelchair, Column Switching from Manual to Power Mode, Clockwise Driving, Counterclockwise Driving, and Forward Figure-Eight Maneuvering; represents the effectiveness of the PAU when driving forward, switching from manual to power operating mode, and transferring into the wheelchair while the unit is attached. One person did require an extended time to transfer out of the wheelchair with the PAU attached (time greater than the Planned Target Level). However, the transfer was successful and all other subjects completed the task within the specified Planned Target Level.

Battery and Column Attachment/Detachment Tasks. Subject six was the only subject out of seven attempting the tasks who was not able to complete the battery and column attachment and detachment tasks. The subject chose not to attempt the tasks based on a concern for loss of balance. These tasks each involved retrieving or placing objects on the ground, and the subject did not feel safe while leaning over to such an extended position. The subject did explain that he thought he might be capable of performing the tasks if they were not low to the ground. This is a significant design consideration, the low position of the battery forces the user to lean over. The column section, however, attaches at a more workable height and this problem may have been alleviated by redefining the task. For example, the column section may have been retrieved and replaced onto a bench positioned at the attachment height.

For the battery and column procedures, the attachment sequences required more than twice the time required for detachment. The average time required for battery attachment was 39.8 seconds and the average time required for detachment was 14.7 seconds. All subjects completing the battery detachment task did so within the Planned Target Level. While there was one subject completing the battery attachment task with a time of 1 minute, 18 seconds, which is greater than the Planned Target Level of 52 seconds. The same type of pattern is seen with the column attachment/detachment procedures. Only one subject completing the column detachment procedure produced a time greater than the Planned Target Level, while two subjects fit in this category with the column attachment task.

Switching to/from Transfer Mode Tasks. Subject six had difficulty switching the column unit to and from transfer mode. The impact analyses identified the reasons for this problem and they are carefully explained in that section. The subject was not able to complete the power mode to transfer mode switching task, and completed the transfer mode to manual mode switching task with a time of 20 seconds which is greater than the Planned Target Level of 18 seconds.

Subject five was the first wheelchair user to perform the tasks and was not able to complete the transfer mode to manual mode switching task. This was due to the weight distribution of the subject in the wheelchair -- primarily at the front edge of the wheelchair seat. The subject was not able to move back against the wheelchair seatback due to his size, and therefore caused the sides of the wheelchair frame to pull together at the front edge of the seat. This resulted in a decrease in the width between the securing blocks for the upper crossbar which in turn made the task of pulling the crossbar into the blocks quite difficult. This problem was alleviated for subsequent subjects (numbers three

through eleven) by placing a board and cushion on top of the sling seat of the wheelchair. The board served to distribute the subject's weight more evenly across the wheelchair frame. This was the only change to the evaluation arrangement after this subject. Therefore, the data from this subject were included in the evaluation.

Estimated Times the Subjects Would be Willing to Spend on the Tasks

Pre- and post-task time estimates of the time the subjects would be willing to spend on a defined task are listed in Tables 10 and 11. The time estimates represent the subject's verbal response to the question "As a potential consumer, how much time would you be willing to spend performing this task?" The post-task estimate also included a statement telling the subject how long he or she required to perform the task.

Table 10. Pre-Task Time Willing to Spend Estimates

Table 11. Post-Task Time Willing to Spend Estimates

Of the seven subjects, only one produced task performance times which exceeded the estimates of the time the subject would be willing to spend on the task. The time required to complete four of the tasks was greater than the pre-task time willing to spend estimates. This number dropped to only one task exceeding time willing to spend with the post-task estimates. The maneuvering in reverse task performance time was two minutes, 10 seconds, while the post-task time willing to spend estimate was 45 seconds. In addition, there were a total of eight task performances which resulted in a "not able to complete" declaration (see previous section). These instances also exceeded the pre-task time willing to spend estimates. Overall, 72 task performances produced times which were equal to or less than both the pre- and post-task time willing to spend on the task estimates.

Statistical Comparisons. It was hypothesized that the subject time willing to spend estimates would be altered by the knowledge of initial use performance time for a task. This hypothesis was based on the assumption that the subjects would not provide an accurate pre-task estimate, due to a lack of a reference point. Given a reference point, the knowledge of initial use performance time, the post-task estimate would be altered to a truer value representing time willing to spend on a task. It was further hypothesized that the observed task performance times would be equal to or less than both the pre- and post-task time willing to spend estimates. Statistical tests were conducted to determine if these hypotheses were supported with the collected data. Alpha levels for the tests were set at 0.10 based on the subjectivity of the data. It was anticipated that time estimates provided by the subjects would have a high variance. Collected data supported this assumption and therefore the 0.10 alpha level was maintained.

For each of the thirteen defined tasks, a single factor analysis of variance was performed on the resulting times for the three factors: pre-task time estimate, observed performance time, and post-task time estimate. The full analysis of variance tables are located in Appendix L: Time Data Statistical Comparisons. A significant difference at the 0.10 alpha level was found between the times for the following tasks: battery attachment, battery detachment, column attachment, column detachment, transfer into the wheelchair, transfer out of the wheelchair, driving straight and clockwise, and driving straight and counterclockwise tasks.

No significant difference was found between the three times for any of the column switching tasks. Results from the transfer to manual column switching task produced a p -value of 0.24, and both the manual to power and power to transfer switching tasks produced p -values of 0.21. The forward and reverse maneuvering tasks also did not have significantly different time results at the 0.10 level. P -values for the forward maneuvering and reverse maneuvering task times were 0.14 and 0.54 respectively.

Paired comparisons were performed on the tasks found with significantly different times. Tukey's post hoc comparisons were conducted with a family alpha level of 0.10 and individual alpha level of 0.042. In all eight tasks exhibiting significantly different times (battery attachment, battery detachment, column attachment, column detachment, transfer into the wheelchair, transfer out of the wheelchair, driving straight and clockwise, and driving straight and counterclockwise tasks), the observed performance times were significantly less than the pre-task time estimates. Also, for all thirteen tasks, no significant difference was found between the performance times and the post-task time estimates. The pre-task time estimates were found to be significantly greater than the post-task time estimates for both column attachment and detachment tasks. However, no difference was found between pre- and post-task time estimates for the remaining eleven tasks.

Discussion. Based on these results, it can be concluded that the observed performance times for all thirteen tasks are equal to or less than both the pre- and post-task time willing to spend estimates. This finding can be interpreted to demonstrate that PAU performance concerning time to complete these thirteen tasks, falls within consumer satisfaction levels. These analyzed data did not include the eight trials out of 84 which resulted in a "not able to complete" declaration. Design deficiencies responsible for the failed task attempts are identified and analyzed in a later section.

There is a lack of evidence for a statistically significant difference between the pre- and post-task time willing to spend estimates provided by the subjects. Only two out of the thirteen tasks provided evidence of a significant difference between the time estimates (column attachment and detachment tasks). Based on these findings, it cannot be concluded that knowledge of initial use task performance time influences estimation of the time a person would be willing to spend performing the task.

It should be noted that these statistical comparisons were based on very small sample sizes (average of n equal to five) with large sample variances. The data collected are extremely subjective and this contributed to the high variation. In addition, only one subject was responsible for five of the "not able to complete" task performances and provided the only performance times which exceeded pre- and post-task time willing to spend estimates. Following the usability testing, this subject stated that the PAU "does not fit my needs" and did not feel that he would be a possible consumer. It is difficult to determine if the results of the study influenced this statement, or if the non-consumer status influenced the performance. Future studies may consider emphasizing subject screening procedures which identify consumers with a need for the product.

Estimated Multi-Task Scenario Times

The average task times were used to estimate user performance times for sequences of tasks. For example, the time required for a wheelchair operator to attach the PAU, transfer into the wheelchair, and place the PAU in power-drive mode can be estimated by summing the average times calculated for the following tasks:

- Battery Attachment.
- Column Attachment.
- Transfer Into Wheelchair.
- Column Switching from transfer position to manual-drive position.
- Column Switching from manual-drive position to power-drive position.

Time estimates were calculated for the following scenarios by summing the average task times for the tasks listed after each scenario. The summed times have been rounded to the nearest second.

- While seated next to the wheelchair, the time required for the wheelchair operator to attach the PAU, transfer into the wheelchair, and place the PAU in the power-drive mode:

Task	Average Time to Complete	Average Time Willing to Spend
Battery Attachment	39.8 sec	1 min 58 sec
Column Attachment	29.0 sec	3 min 48 sec
Transfer Into Wheelchair	8.7 sec	2 min 2 sec
Column Switching from transfer position to manual-drive position	10.4 sec	1 min 47 sec
Column Switching from manual-drive position to power-drive position	4.0 sec	1 min 28 sec
Sum of all task times	1 min 32 sec	11 min 3 sec

The estimated time required for a wheelchair operator to attach the PAU, transfer into the wheelchair, and place the PAU in the power-drive mode, is 1 minute, 32 seconds. This is much faster than an average time of 11 minutes, 3 seconds which the subjects estimate they are willing to spend on this set of tasks. The time is calculated by averaging the estimated times the subjects are willing to spend on the task (post-task estimates) and summing across all of the tasks in the set. Only estimates provided by subjects who were able to complete the tasks are included.

- While seated in the wheelchair with the PAU in power-drive position, the time required for the wheelchair operator to place the PAU in the transfer position, transfer out of the wheelchair, detach the column section, and detach the battery:

Task	Average Time to Complete	Average Time Willing to Spend
Column Switching from power-drive position to transfer position	9.2 sec	1 min 46 sec
Transfer Out of the Wheelchair	10.6 sec	1 min 44 sec
Column Detachment	14.3 sec	3 min 12 sec
Battery Detachment	14.7 sec	1 min 39 sec
Sum of all task times	49 sec	8 min 21 sec

The estimated time required for a wheelchair operator to place the PAU in the transfer position (from power mode), transfer out of the wheelchair, detach the column section, and detach the battery, is 49 seconds. In comparison, the average estimated time that subjects are willing to spend on this set of tasks is 8 minutes, 21 seconds.

- The wheelchair operator seated next to the wheelchair with the PAU in place (transfer mode). The time required for the wheelchair operator to transfer into the wheelchair and place the PAU in the power-drive position:

Task	Average Time to Complete	Average Time Willing to Spend
Transfer Into Wheelchair	8.7 seconds	2 min 2 sec
Column Switching from transfer position to manual-drive position	10.4 seconds	1 min 47 sec
Column Switching from manual-drive position to power-drive position	4.0 seconds	1 min 28 sec
Sum of all task times	23 seconds	5 min 17 sec

The estimated time required for a wheelchair operator to transfer into the wheelchair and place the PAU in the power-drive position (from transfer mode), is 23 seconds. The summed average times the subjects are willing to spend on this set of tasks is 5 minutes, 17 seconds.

- While seated in the wheelchair with the PAU in power-drive position, the time required for the wheelchair operator to place the PAU in the transfer position and transfer out of the wheelchair:

Task	Average Time to Complete	Average Time Willing to Spend
Column Switching from power-drive position to transfer position	9.2 seconds	1 min 46 sec
Transfer Out of the Wheelchair	10.6 seconds	1 min 44 sec
Sum of all task times	20 seconds	3 min 30 sec

The estimated time required for the wheelchair operator to place the PAU in the transfer position (from power mode) and transfer out of the wheelchair, is 20 seconds. This time

is less than one tenth of the average estimate of time willing to spend on the tasks of 3 minutes, 30 seconds.

Discussion. These time estimates represent "initial use" performance only and do not include every aspect of the actual task sequences. For instance, the times do not reflect the action of placing the PAU parts into their positions on the floor prior to the attachment effort. However, it is anticipated that these rough estimates calculated for actual wheelchair user task performance with the PAU can be a useful reference.

In each of the four scenarios, the average observed initial performance times were far less than the averaged estimated times that the subjects would be willing to spend on the tasks. These estimated times were provided by the subjects after they had been told the length of time his or her performance of the task required. Therefore, the subjects each had a frame of reference upon which to base his or her time estimates. This finding indicates that the time required to perform the attachment, detachment, and operation tasks associated with the PAU, is well-within the time a potential consumer is willing to spend on the tasks.

It should be reemphasized, however, that only times provided by subjects who were able to complete the tasks, are included in the analysis. The reasons for noncompletion of a task are analyzed within the Critical Incident and Critical Observation Impact Analyses. Identification and prioritization of these design deficiencies and others is the focus of the analyses and this information should be utilized as a starting point for the next iteration of design changes.

Critical Incident Impact Analysis

The two impact analyses proposed for the usability evaluation were adapted from the analysis developed by Del Galdo et al. in 1986. The following procedures were used to analyze the data collected from the critical incident and verbal report tools.

- 1) For each critical incident, the "critical incident description" data collected from the verbal reports and the "features of the design leading to the critical incident" data collected from the critical incident tool, were examined.
- 2) From these data, the problem with the PAU design or task procedures which led to each critical incident, was defined.
- 3) The list of defined problems from each subject was consolidated into one list for all of the subjects tested.
- 4) The total number of critical incidents across all subjects resulting from each individual problem (the same problem may have caused different types of critical incidents) was summed. Each subject contributed a maximum of one frequency count per problem. When the same problem was identified more than once by the same subject, the maximum criticality rating provided by the subject was assigned to the problem for the consolidated analysis.
- 5) The criticality ratings for all critical incidents which resulted from each individual problem were averaged. This produced one criticality rating for each problem in the list.
- 6) The list of problems was placed in a table with the associated frequency and average criticality in order of decreasing frequency. Those problems with the same frequency were ordered by decreasing average criticality.

This process produced a list of problems with the PAU which are prioritized in terms of both frequency and criticality. The list developed from the critical incident analysis is seen in Table 12.

Table 12. Problems Identified by Critical Incident Analysis

Critical Observation Impact Analysis

The same impact analysis procedure was followed to produce a separate list of prioritized problems from the critical observation data. Two sets of criticality ratings for the critical observation problem list are presented in Table 13. The average criticality ratings in column 4 are the ratings produced by the investigator who reviewed the usability evaluation notes, video tapes and audio tapes. A separate rating was provided for each observation of the encountered problem, and the final number is the average across all observations of the problem. The criticality of the problem was defined based on its impact on an individual subject's performance. Therefore, the same problem produced a variety of ratings across subjects. Each rating reflected how critical the experimenter found the problem to be with respect to the negative impact that it had on user performance and satisfaction. This was completed with the underlying assumption that a decrease in user performance or satisfaction would negatively impact the user's decision to buy the product.

Table 13. Problems Identified by Critical Observation Analysis

The ratings listed in columns 5 and 6 were provided by two additional design team members who were familiar with the usability evaluation procedures. They reviewed the summarized problems, as listed in a randomized version of the table, and offered a criticality rating for the specific problem. The team members did not review any material other than the table. One of the team members (column 5) was present throughout the usability evaluations with the subjects. Ratings from all three design team members were averaged and placed in column 7 of the table in an attempt to alleviate some of the bias present with a sole experimenter.

Discussion. Design problems listed in the critical observations table include those observable problems experienced by subjects which were not expressed in the verbal report. The list also includes some difficulties which the experimenter detected that were not noticeable by the participant, in addition to the observable problems which were identified by the subject. The list does not include cognitive problems which subjects may have experienced, because these were not observable to the experimenter. For this reason, it was important to include the critical incident identified problems which do include cognitive difficulties expressed in the verbal report process.

The two lists serve as a starting point from which to consider design changes for the next iteration of the PAU. The prioritized design deficiencies are discussed further in a later section entitled Design Recommendations.

User Evaluation Questionnaire

The data from the questionnaire were analyzed for innovative ideas, a description of the product's usefulness to the subjects as potential consumers, identification of problems, and determination of avenues which may serve to solve the problems.

Responses to the questions have been consolidated and summarized here. The discussion has been augmented with comments and suggestions offered by the subjects throughout the evaluation (not specifically written on the questionnaire).

Particularly Difficult Aspects of the PAU. A majority of the subjects (four out of seven) expressed a concern for the operation and control of the PAU when driving in reverse. Comments such as "difficult in trying to control what direction it was going in," and "in reverse, when the handle is turned to more than 45 degrees it tends to want to pull out of your hand and turn you 90 degrees" were used to describe the problem. Additional comments varied from subject to subject and included statements indicating (but not explaining) that the transferring with the PAU attached, loading and unloading the battery, and squeezing the crossbar release mechanism, were all processes which were particularly difficult. One subject commented that "I would forget to lean the control down to lock it in place [rotation of the inner column when switching to transfer position]" and also that "the control is fixed too close to the body." The subject who made this statement was relatively large and the extension of the handle back toward his body became uncomfortable when he was required to reach forward.

Suggestions for Future Design Changes. Subjects provided a range of suggestions for improving the current PAU design. Their ideas are listed here.

- Implement a velcro fastener for the battery sling which could permit attachment and detachment with only one hand. Another subject suggested designing the sling so that only one clasp is required to secure the battery, instead of the two clasps currently required.
- Relocate the battery on the wheelchair so that the wheels of the PAU drive unit do not hit the battery. [The drive wheels currently come to rest against the battery as the column unit is rotated into transfer position. It is the battery which keeps the column from rotating further down toward the ground (alternatively, the cross frame section of the wheelchair would eventually stop the column before reaching the ground). Removal of the battery may provide an easier transfer process. However, it may also introduce a problem by requiring the user to reach farther forward to pull the column up into a driving position.]

Another subject suggested hanging the battery on the back of the chair (from the push handles) or placing it on a sliding rack under the chair. On the sliding rack the battery could be placed on the rack and then pushed into position so that it would not be necessary to reach up under the chair to attach and detach. [The rack arrangement introduces a new attachment and detachment problem in itself. Such a system would require secure attachments which would need to be removed for the wheelchair to fold.]

- Move the top crossbar release mechanism up onto the column so that it would be easier to switch between column positions. Several subjects introduced this idea and one subject also suggested utilizing longer releases to increase gripping surface area.
- Relocate and redesign the lower crossbar release mechanism. One subject suggested using a large single button release placed on the lower side of the top column/crossbar attachment block. This position is based on the optimal handling configuration of the column unit when lifting for attachment. The unit is balanced and easily controlled when held at the underside of the top column/crossbar attachment block. One hand can hold the unit at this location

and manipulate the entire column unit into position. Figure 15 points to the placement of the release button on a photograph of the column unit.



Figure 15. Suggestion for placement of the lower crossbar release mechanism (PICT, 133 k).

- Model the lower crossbar release mechanism after the quick-release wheelchair wheel design concept. This would permit placement of the column unit onto the wheelchair without actuating the release mechanism (simply drop it into position) which would only be utilized when detaching the unit.
- Design the handle and controls such that the distance from the user's body can vary depending on the user's torso size. [This type of variability may also serve to optimize fit for users with different arm lengths as well.]
- Place the forward/reverse switch in a position that is easier to manipulate. The subject suggested a button on the handle body oriented perpendicular to the side of the handle. The investigator's interpretation of the control idea is shown in Figure 16.

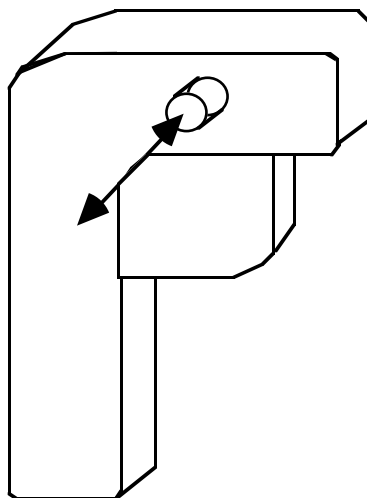


Figure 16. Forward/reverse control suggestion.

Product Usefulness and Consumer Potential. Four of the seven potential consumer subjects stated that the new product would be useful to them. Explaining that it would be less than an electric chair [expected to cost less], it would be easier to transport [than an electric chair], and that it would serve to increase independence by getting "me places that I would have to have help." One subject pointed out that it would not only be very useful going up inclines, but that it would be less strain on the body [than a manual wheelchair].

Two of the three potential consumer subjects who did not think the product would be useful to them, explained that they are too active at this stage in their lives to use such a product. Both were young and very athletic and indicated that they did not have a problem at all getting around. The third subject who did not think that the product would be useful, simply explained that it "does not fit my needs." This subject primarily operates a manual wheelchair when at home, and an electric scooter when traveling outside the home.

Ease of Operation. The responses to question four, "How easy was the power assist to operate?," were compiled and presented in a bar graph which illustrates the frequency with which each level of the answer was selected. Figure 17 shows the results of the ease of operation ratings provided by the subjects. The seven point scale graphically presented bipolar endpoints of "Easy" at the "1" endpoint, and "Difficult" at the "7" endpoint.

On the scale of one to seven, where one is "Easy" and seven is "Difficult," the average ease of operation rating was 2.3. Six of the seven potential consumer subjects rated the PAU ease of operation level to the left (easy side) of the scale midpoint of four. Only one subject provided a rating on the difficult side of the scale at level five. This subject also explained that the product would not be of use to him due to his active lifestyle. Interestingly, the experimenter observed that this subject was able to perform all tasks with more apparent ease than any of the other subjects.

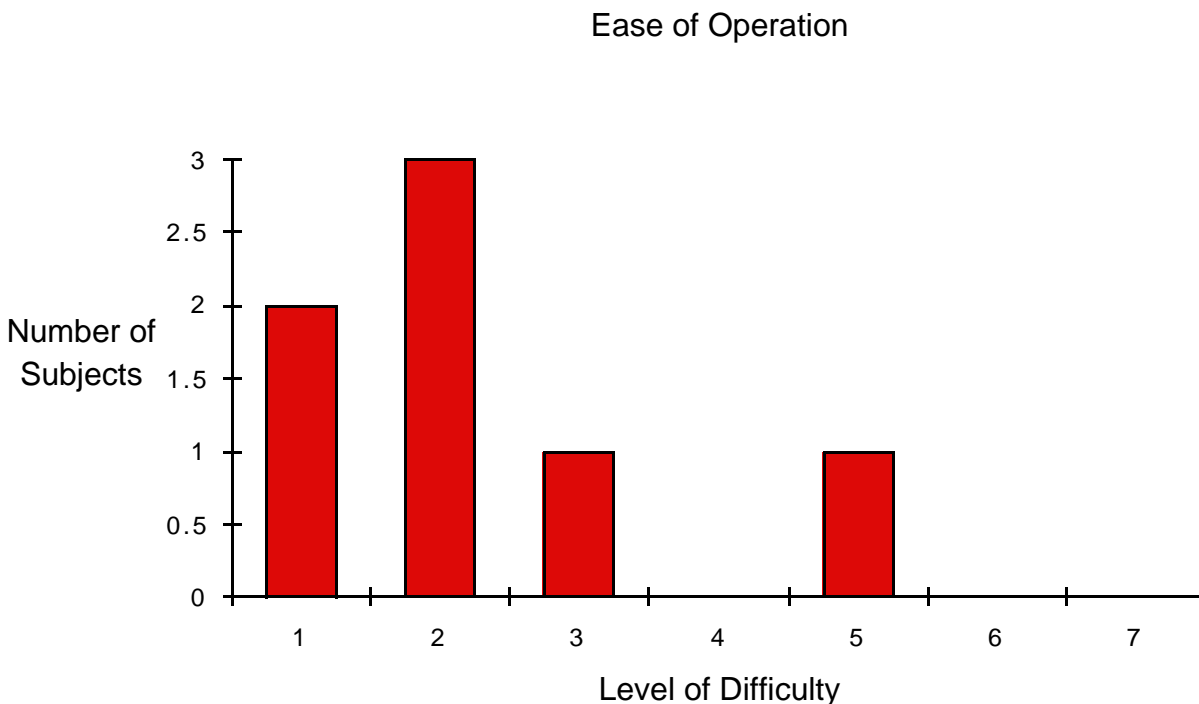


Figure 17. Ease of operation questionnaire results.

Auxiliary Analysis -- Able-Bodied and Dropped Subjects

The four additional participants in the usability study, two able-bodied pretest subjects and two dropped subjects, did provide some helpful information.

Subject One -- Able-Bodied Pretest. An important design deficiency was identified in the first pretest with an able-bodied subject. It was discovered that a pinch point exists in the juncture at the top of the outer column and the extended handle arm (component which reaches back toward the operator). The space between the two components gaps open when the unit is placed into power operating mode. This gap then closes again when the user switches the column out of power mode by rotating the drive wheels off of the ground. If the user handles the column at the gapped area while switching column positions, the hand can be pinched between the two metal pieces. This is an important safety consideration and the design should be altered. The first subject was also a design team member and he has provided an ease of operation rating for the prototype of 2 on the difficulty scale of 1 to 7.

Subject Two -- Able-Bodied Pretest. The second pretest subject was not a member of the design team. This subject found many of the problems identified by the wheelchair subjects such as: loud motor whine, difficulty in placing the column unit into the securing blocks, too much resistance in the finger trigger, and not going in the intended direction when driving in reverse. In addition, this subject identified a problem not introduced by the wheelchair operator participants. She described the crossbar release mechanism as having a "pinchy type feel, rough on the edges." Subject two provided an ease of

operation rating of 4 to describe forward operation, and 6 to describe operation in reverse.

Subject Three -- Dropped Subject. Subject three was a quadriplegic wheelchair operator and therefore did not match the capability levels defined for potential consumers of the new invention. His participation did provide some information concerning the potential future market of quadriplegic consumers. The new PAU would require adaptations designed to accommodate users with more limited hand and arm capabilities as well as decreased upper body stability. Changes to the design as a result of recommendations from this usability evaluation will incorporate more universal design features. This shift to a more universal design will expand the possible consumer base of the product. In its current form, subject three rated the ease of operation as a 5.5 on the difficulty scale of 1 to 7.

Subject three recommended making the adaptations for the quadriplegic market. He was capable of completing the battery attachment and detachment sequences. There was some difficulty with the securing clasps resulting in an attachment time of 3 minutes, 40 seconds, and a detachment time of 4 minutes, 5 seconds. Subject three was also able to transfer in and out of the test wheelchair with the PAU attached and commented "That would not be a problem," and mentioned that he did take extra care with balance.

In response to the questions, "As a potential consumer, would this product be useful to you? Why or why not?," he said, "Yes, absolutely. Many would like to maintain the independence of the manual wheelchair and the temporary nature of this product works well to do this."

Subject Four -- Dropped Subject. Subject four was unable to perform any of the defined tasks, but was familiarized with the new PAU prototype through a demonstration and offered some comments. Concerning the battery sling, she suggested utilizing only one clasp instead of two so that one hand could be used to secure the battery. She also suggested that it would be easier to pull a single crossbar release mechanism, instead of squeezing the two finger grips together. Subject four preferred a joystick for powered operation and expressed a concern with balance if she was to attempt attachment or detachment procedures. In response to the questions "As a potential consumer, would this product be useful to you? Why or why not?," she said, "Yes, if it was designed a little better." Subject four was not able to transfer into the wheelchair to attempt operation.

Evaluation Structure

The unique structure of the usability evaluation was tailored to optimize information collection concerning user-PAU interaction characteristics. The successful identification and prioritization of design deficiencies and solicitation of design recommendations demonstrates the effectiveness of such usability testing methods with a mechanical prototype. Several aspects of the evaluation structure have been identified which may be improved in the future. The experimenter suggests consideration of the following recommendations before utilizing the evaluation structure with future applications.

Reverse Figure-Eight Maneuvering Task

The Reverse Figure-Eight Maneuvering Task was not representative of a realistic product performance requirement. One subject commented on the task by saying "I can't see any practical way [reason] you would need to go for a long stretch [in reverse]."

The task did serve to identify significant design concerns with the PAU steering control system. However, the extensive nature of the task may have placed an unrealistic importance on these findings. Only one third of those tested were able to successfully complete the task within the Planned Target Level. This produced a feeling of inadequacy with several participants and the experimenter was required to carefully explain that the results were due to a design problem (following the video review procedure). Each of the six subjects attempting the task clearly demonstrated the ability to reverse the wheelchair in a controlled manner for a very short distance (length of the wheelchair). Such a performance may be adequate for user needs such as backing the wheelchair so that it will face a different direction.

Videotape Review Intervals

The videotape review procedure was scheduled too often within the evaluation. The review session after each of the six task groups made procedural flow discontinuous. Given the same circumstances, the tasks should be grouped into three categories so that only three review sessions are required. The battery and column attachment/detachment procedures can be grouped together, followed by the transfer and column switching procedures, then finally the driving and maneuvering procedures.

There was no apparent decrease in information collected from tasks performed earlier on the videotape than later (early tasks within the group versus those performed right before the tape was reviewed). Therefore, it is not anticipated that information would be lost by consolidating the tasks into larger groups. The use of the videotape review is recommended for data collection. One subject commented "It really helps to be able to see what you did."

Verbal Report Procedure

All subjects tested had some difficulty with the verbal report task to "think aloud." Despite a demonstration by the experimenter, a practice session, and feedback from the experimenter, the subjects did not express all thoughts while reviewing the videotape. Several subjects were quite reserved and required repeated prompts from the experimenter to speak. Though the task did provide good data on the user-PAU interface, it appeared that the unnatural request to think aloud made several of the subjects uncomfortable. Subjects may be more responsive to direct questions such as "Did you experience any problems with this procedure?" or "Did the device react the way that you were expecting it to react?" These types of questions may make it easier for some subjects to express their thoughts.

The extensive use of the verbal report tool in usability testing with human-computer interfaces has proved successful. The types of tasks involved with computer interface testing may be more cognitively based than the evaluation of motor tasks with a mechanical prototype. Such cognitive tasks require a method to express thought

behaviors such as the verbal report tool. Mechanical prototyping, however, is more manual skill-oriented and may not require the extensive evaluation of the user's thought processes. For this reason, consideration should be given in the future to retrieving critical incident data with a more direct approach; such as direct questions posed to the testing subject concerning the performance.

Formal Testing Atmosphere

Much effort was required on the part of the experimenter to make the subjects feel at ease with the overall testing situation. The formal laboratory environment, informed consent form, highly structured experimental procedure, and the recording of the performances on video and audio tape, all added to the formal atmosphere. This was not conducive to eliciting uninhibited comments and creative suggestions.

Future application of these usability evaluation procedures should involve a method to vary subject control based on the individual needs of the participant. For example, some subjects were willing to follow conversation tangents into creative design suggestions and even unique task attempts. One subject asked if he could attempt to attach and detach the column unit while sitting in the wheelchair. The experimenter took advantage of this opportunity and learned new information about user-PAU performance. If the initial structure of the evaluation was not so rigid, more opportunities such as this may have been presented.

An example of such a lost opportunity was discussed earlier with subject number six and the column attachment/detachment procedure. The subject was unable to attempt the task due to a balance concern with leaning over to retrieve and place the column unit on the ground. He felt that otherwise he would not have a problem with performance of the task. The experimenter had not anticipated any benefit from straying from the predefined evaluation procedure, and therefore moved onto the next set of tasks. In retrospect, it may have been more beneficial to readjust the task by placing the column unit on an elevated surface (such as a bench) and allowing the subject to proceed from there.

Time Estimates

Many of the subjects had difficulty providing pre- and post-task estimates on the time that they would be willing to spend to perform the tasks. In some cases, so much thought was put into the answer that the flow of the evaluation was disrupted by the time needed to formulate an answer. Numbers were sometimes provided which did not have apparent reasoning behind them. For example, one subject provided the answer of five minutes to ten out of the twelve post-task time estimate questions he was asked.

It is also difficult to determine how accurate the estimates are with respect to the length of time required to reach a point when the value of the product is no longer worth the effort required. One subject provided a pre-task estimate of ten minutes (time she would be willing to spend performing the task) to complete the Reverse Figure-Eight Maneuvering Task. She was then observed to be visibly frustrated within two minutes of attempting to perform the task and was asked to stop her attempt soon after. This may give some indication of the inability of some people to estimate these time intervals.