A Process for Usability Testing of Lifesaving Equipment

William Porter, Carin Kosmoski, Rohan Fernando
National Institute for Occupational Safety and Health
Pittsburgh Mining Research Division
626 Cochrans Mill Road
Pittsburgh, PA 15236 U.S.A.

Usability for any product, and especially for a lifesaving device, is critical in that the users will be interacting with the device in a highly stressful and complex environment. This study examined self-contained breathing apparatus (SCBAs) and conducted a usability assessment of these SCBAs with refill stations as it pertains to mine escape. Data was collected examining three usability topic areas (effectiveness, efficiency, and satisfaction) and eight constructs within these topic areas (completeness, accuracy, time requirements, overall satisfaction, discomfort, ease of use, system performance, and user preference). This paper documents the usability framework adopted and the methodology used to answer the research questions of the study and includes sample results and discussion. The methodology presented can be modified and used to test other lifesaving technologies to compare the usability of the devices and to estimate the ability of the devices to function as expected in a lifesaving situation.

INTRODUCTION

Usability for any product, and especially for a lifesaving device, is critical in that the users will be interacting with the device in a highly stressful and complex environment. An underground mine—where the mine atmosphere can rapidly become toxic in the event of a fire—represents one such environment where users must rely on lifesaving devices to survive in an emergency. This study examined the major commercially available US manufacturers' self-contained breathing apparatus (SCBAs) and conducted a usability assessment of these SCBAs with refill stations as it pertains to mine escape. This research addressed the following questions: (1) Can a person don and operate current commercially available SCBAs in a mining environment? (2) Can a person refill the current commercially available SCBAs from an SCBA refill stations in a mining environment? (3) Do users show a preference between current commercially available SCBAs or SCBA refill stations when used in a mining environment? The purpose of this paper is to document the usability framework adopted and the methodology used to answer the stated research questions. Sample results and discussion on the overall study results are also included.

Usability Framework

The usability testing process as established by the International Organization for Standardization (ISO) was chosen as an overall structure for this research. ISO recommends the testing of three key parameters to fully assess a products usability: effectiveness, efficiency, and satisfaction. ISO 20282-2 defines the process with which testing is to be completed and the type of measures to be collected. Within this guidance some flexibility exists as to the type of information and the way in which the information is to be collected. This paper details how the guidance from ISO was further developed into working definitions, instrumentation, and data analysis procedures specific to lifesaving technology.

For this research, effectiveness was separated into two constructs of completeness and accuracy, efficiency was examined using the construct of time requirements, and satisfaction was separated into five constructs of overall satisfaction, discomfort, ease of use, system performance, and user preference. Table 1 provides working definitions for each portion of this usability framework, and Figure 1 depicts the relationship of topic areas, constructs, and the associated measures used during this research study.

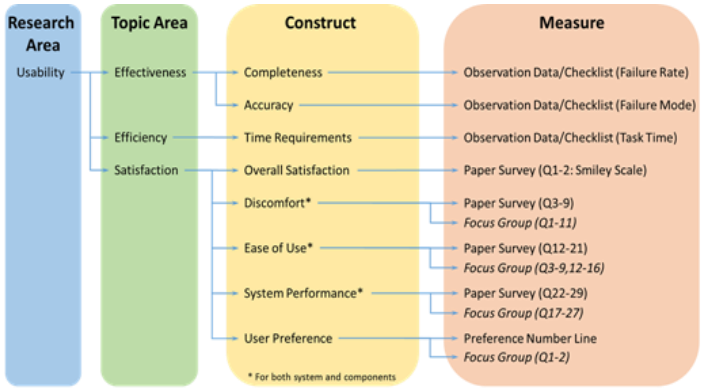

Figure 1: Structure of usability testing. Quantitative measures are shown in plain font and qualitative measures are shown in italics.

Several decisions were made to allow for the use of this usability framework in addressing the research questions posed. First, the products being examined are commercially available SCBAs and SCBA refill stations available to mines in the United States for use during mine escape. Second, the intended users of these products are United States underground coal miners. Third, the specified goal of the product is to allow users to escape from an Immediately Dangerous to Life or Health (IDLH) environment by isolating their lungs from the potentially toxic environment. This includes the process of refilling the SCBA while still isolating users' lungs from the environment. Lastly, the context of use of the product is an underground mine emergency which has produced an IDLH environment and requires miners to self-escape from the mine.

Table 1: Usability Framework for the Purposes of This Study

Term	Definition	Source
usability	extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use	ISO 9241‑210:2010 definition 2.13
usability testing	evaluation that involves representative users performing specific tasks with the system to enable the measurement of efficiency, effectiveness, and/or user satisfaction	ISO 25060:2010 definition 2.17
effectiveness	the extent to which users are able to achieve the intended goals of the system	adapted from ISO 20282-2:2013 Introduction, and ISO 9241-11:1998 Definition 3.2
completeness	the extent to which the tasks/subtasks (as defined by the task analysis) are completed by the user	adapted from ISO 20282-2:2013 7.5.2.2
accuracy	the frequency of errors, frequency of assists to the user, and frequency of access to help or documentation by the user during the task/subtasks	adapted from ISO 25062:2006 5.4.4.1
efficiency	the resources (time) required by users to achieve the goals of the system	adapted from ISO 20282-2:2013, and ISO 9241-11:1998 Definition 3.3
time requirements	the time required to achieve critical tasks	adapted from ISO 25060:2010
satisfaction	the extent to which users are satisfied with their experience with the system	adapted from ISO 20282-2:2013, and ISO 9241-11:1998 Definition 3.4
overall satisfaction	the extent to which users are satisfied with their experience with using the system in the specified context of use	adapted from ISO 20282-2:2013, and ISO 9241-11:1998 Definition 3.4
discomfort	the degree to which the user is satisfied with physical comfort (lack of discomfort) of the system in the specified context of use	adapted from ISO 20282-2:2013 4.14, E.2, and ISO 25010:2011
ease of use	the users' feelings towards their interaction with the overall system and specific components of the system	developed by study researchers
system performance	the users' feelings toward overall design of the devices and interaction of the system components	developed by study researchers
user preference	the user's overall preference for the different SCBAs and refill stations	developed by study researchers

METHODS

Data collection for this study took place underground in the National Institute for Occupational Safety and Health (NIOSH) Safety Research Coal Mine (SRCM) in Bruceton, PA. The SRCM facility was chosen because of easy access, its close proximity and similarity to a real mining environment.

Participants

Four employees (3 males and 1 female) were recruited and participated in this study. The study participants were federal employees located at the Bruceton, PA, research campus. One participant wore glasses and had short facial hair at the time of testing, which is relevant when using a SCBA because it can affect fit. All participants gave their informed consent in writing prior to testing after being made aware of the study requirements and potential risks. The participants had a mean age of 29 years (SD ±1.8), a mean height of 70 in. (SD ±3 in.), and a mean weight of 193 lbs. (SD ±34 lbs.).The participants were generally in very good health and physical condition and had limited to no familiarity with SCBAs or other breathing apparatus as self-reported during participant screening.

Procedure

To accomplish the objective of this research, a comprehensive human subjects experiment was undertaken. During this experiment, a group of participants were asked to interact with SCBAs and SCBA refill stations in an underground coal mine. The SCBAs examined during this testing were chosen because they represent the major models that are commercially available to mining companies for use during mine escape. Six SCBA and two refill stations were tested. The SCBAs had varying face mask/hood designs, location of refill ports, strap/harness design, etc. The refill stations had varying hose length, hose spacing, activation mechanism, signage, etc.

This experiment occurred in two phases. Phase one allowed participants to become familiar with all of the SCBAs and SCBA refill stations in an above ground setting, while phase two required the participants to use each SCBA and each SCBA refill station in an actual coal mine environment. During phase two, participants were observed and timed during all testing by NIOSH researchers to document any usability concerns with the devices. After interacting with each device, participants completed paper surveys. These surveys included rating of the level of overall satisfaction, discomfort they experienced, perceived ease of use and system performance. After testing, the participants rated their preference for each SCBA and refill station. Additionally, any feedback from the participants on the SCBA/refill station and their perceptions of interactions between the devices were noted with special attention being paid to the usability of the devices during mine escape. At the completion of testing, a moderator-led focus group was held with all participants to further capture any issues or concerns they may have had with the SCBAs or refill stations use during testing.

To test the SCBA refill stations at full capacity, four participants performed the study together, with the same four grouped together for both phases. At the start of testing for both phases, participants were supplied with and asked to don essential safety gear including a hard hat, mine belt, metacarpal gloves, cap lamp, and boots.

During phase 1, participants were led through an introduction to an SCBA and refill station training program by the researchers during which the components and functionality of the devices were demonstrated. At the completion of the training, participants were allowed to interact with the SCBAs and refill stations for as long as they needed to feel comfortable with the devices. The participants were then required to don at least one SCBA and then refill the SCBA from at least one refill station.

Phase two, underground testing, occurred on two days one-and-one-half weeks after phase one. The participants were again required to don all essential safety gear and once ready were led underground to the testing location. Once

underground the participants were given time to become comfortable with the surroundings and ask any questions they had on the devices being used during testing.

While underground for phase two, a total of 12 trials were undertaken in the SRCM. A total of eight trials were completed on day one and four trials were completed on day two. For each trial, the participants (all four at one time) were asked to don a randomly pre-selected SCBA. Once they had donned the SCBA they walked as a group for 10 minutes around the mine along a predetermined path ending at the refill stations; this simulated a group of miners escaping a mine disaster. Researchers were present and guided the participants along the path during all testing. When the participants reached a randomly selected refill station (one of the two being tested), they were instructed to refill their SCBAs (all participants at the same time). Participants were asked to not assist one another until asked for assistance by the other participant. If assistance was requested it was recorded per the task failure modes. Regardless of the participants having different randomly preselected SCBAs, all participants used one refill station for each trial. Once the participants believed their SCBA was refilled, they were instructed to disconnect and return to the starting location and doff the SCBA. The participants then filled out a paper survey and rated their preference for the SCBA used during the trial on a number line. For the number line, participants were asked to assign a preference value for the SCBA and refill station combinations by placing cards with pictures of the SCBAs on a number line from 0 to 100 (zero being worst and 100 being best). Participants were allowed to move and shift their placement of each SCBA on the number line as they experienced each additional SCBA.

The participants completed the above process for all combinations of the six SCBAs and two refill stations tested (total of 12 randomized trials for each participant). At the end of all of the trials, participants were asked to revisit their preference number line and assign each SCBA a final preference rating. Once all trials and surveys were completed, the participants were led out of the mine and released until the focus group.

On the afternoon of the second day of phase two, a moderator-led focus group was conducted. During this focus group, all four participants were asked to identify any issues or concerns they might have with the SCBAs or SCBA refill stations tested use during mine escape. The focus group occurred in a standard conference room with each of the SCBAs on display for the participants to refer to, if needed, during the course of the discussion. The focus group was audio recorded and detailed notes were taken by the moderator for subsequent data analysis.

Data Collection Instruments

Data was collected examining three usability topic areas (effectiveness, efficiency, and satisfaction) and eight constructs within these topic areas (completeness, accuracy, time requirements, overall satisfaction, discomfort, ease of use, system performance, and user preference). The relationship between these topic areas and constructs are derived from the ISO definitions for usability and usability testing (see Figure 1).

The first step in examining these constructs was to develop a detailed task analysis of the steps required to successfully use the devices in a mine escape scenario. This task analysis was undertaken by researchers on the project and verified by subject matter experts prior to starting the research. The task analysis was then examined to identify critical elements relative to each construct, such as failure mode for the completeness construct. The instruments detailed below for data collection were determined by comparing the critical elements within each construct and task, such as time to don the SCBA within the time requirements construct, with the measures suggested by ISO.

Three main instruments were used during data collection: first, an observation checklist used to document the measures being collected by the research observer; second, a paper survey used to document the measures being collected directly from the participants at the conclusion of each trial; third, a focus group guide used to facilitate the moderator-led focus group at the conclusion of all testing (see Figure 1). An example question from the paper survey for the construct of Ease of Use was “Indicate the level of ease/difficulty you experienced with each of the following activities using the response options: very difficult (1), difficult (2), somewhat difficult (3), somewhat easy (4), easy (5), very easy (6), and unable to perform the task.”

To aid in interpreting the quantitative data collected during the surveys and observations, qualitative data within each construct was collected during the focus group. During the focus group, participants were asked to comment about each component of the SCBAs in general and their overall preference for the SCBAs—that is, questions did not guide the participants to comment on specific usability concerns, but allowed participants to comment on any aspect of each specific component. During the focus group, the moderator probed into the more specific areas when appropriate given the responses from the participants to the general preference questions.

Data Analysis

Each construct had a slightly different data analysis procedure used to analyze the information collected during testing. Inside each construct, researchers were looking for trends or outliers (both positive and negative) for the attributes being examined. By following these data analysis procedures, researchers were able to evaluate the SCBAs and SCBA refill stations based on each construct individually, as follows.

Completeness

Completeness was explored using descriptive statistics compiled based on the number of participants who completed each task/subtask (i.e. percentage of participants who completely achieved each task/subtask).

Accuracy

Accuracy was explored using descriptive statistics compiled based on the number of participants who made at least one error, required assistance, and/or accessed help or documentation while performing the task/subtask.

Time requirements

The time required to complete the tasks was explored using descriptive statistics (mean and range).

Overall satisfaction

Information on participants' overall satisfaction was explored using descriptive statistics (mean and range) for the survey questionnaire items pertinent to this specific construct (mean Smiley scores and standard deviations across participants).

Discomfort/Ease of Use/System Performance

Discomfort, Ease of Use, and System Performance were examined similarly using multiple approaches. First, participants' general perceptions experienced while walking/moving while wearing an SCBA and while connected to a refill station were examined. Additionally, experiences with the specific components of the SCBAs were collected. The resulting information was examined using descriptive statistics (measures of central tendency and variance) by examining the survey questionnaire items pertinent to these specific concepts.

Following the analysis of the quantitative data from the paper surveys, the qualitative data from the focus groups was examined to further enrich or explain the trends, and/or outliers that were found during the quantitative data analysis. For example, researchers may have found (from analysis of the quantitative data) that there is a difference in the amount of discomfort experienced by participants caused by the shoulder straps of one of the SCBAs as compared to the other SCBAs. The qualitative data gained from the focus group discussions could then be used to provide explanations for why that difference was found and what, specifically, was causing the difference in discomfort level (e.g., strap thickness/thinness, padding of the straps, etc.). Finally, the qualitative data from the focus groups was also examined holistically and without the survey questionnaire items in mind, in an effort to identify any information or themes that researchers did not originally anticipate emerging.

RESULTS

Given space constraints, only the results of one construct will be discussed to provide an example of how the procedures were implemented for one construct.

Ease of Use

Participants' evaluation of the overall ease of use of each of the SCBAs during the donning process is presented in Figure 2. SCBA 6 received "somewhat easy" ratings for turning on the SCBA and adjusting the shoulder straps, but "somewhat difficult" to "difficult" ratings for the remaining tasks. The focus group discussion indicated that the participants preferred SCBA 2 overall (3 of the 4 participants preferred SCBA 2 and 1 preferred SCBA 5). This was supported by statements such as, "I was most comfortable with the [SCBA 2]. It was very wearable. I felt the weight distribution was good considering that I was ducking over a lot and rarely able to stand up straight. I felt the facemask was snug and easy to adjust. The connector piece for the refill line extended out so I could connect the refill line without getting help from others." Another example statement expressing SCBA preference: "The straps were easy to adjust. The [SCBA 5] refill was easy to connect because it was right on your shoulder."

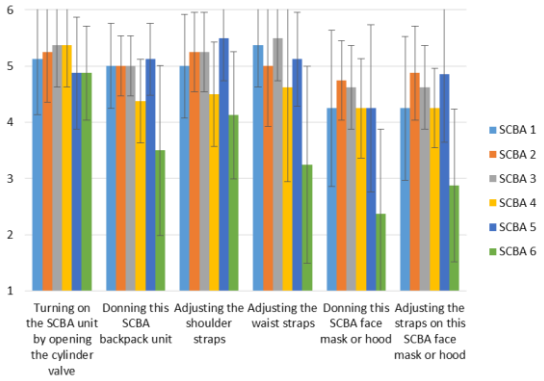

Figure 2: Overall ease of use during donning (mean+-SD) [Scale: (1) very difficult, (2) difficult, (3) somewhat difficult, (4) somewhat easy, (5) easy, (6) very easy].

The focus group data provided further explanation for the strong negative ratings for the hood of SCBA 6 with statements such as the following: "The big negative of [SCBA 6] is the hood. I was constantly focused on trying to keep my hat from falling off. My hat did not fit over the hood. This caused me to use more energy. I doubt whether I could get out of the mine on my own wearing this unit. I felt like I was practically an invalid." "I would not be able to self-escape. This SCBA would just add to my panic." "I became very disoriented wearing it." Most of the participants (3 of the 4 participants) preferred refill station 1 over the refill station. During the focus group, the participants reported that refill station 1 was more intuitive and easier to use. For example, participants made statements such as, "I really liked that, on the [refill station1], once you are connected you are good to go. Less chance of user error."

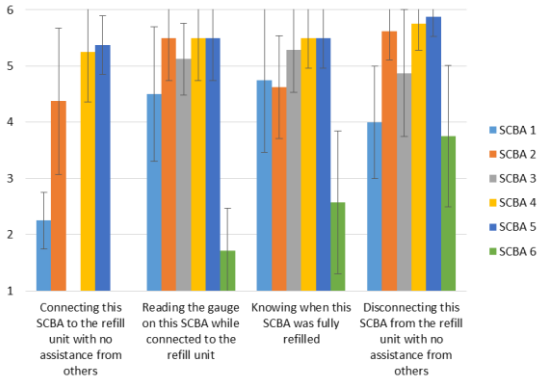

Figure 3: Overall ease of use during refill (mean+-SD) [Scale: (1) very difficult, (2) difficult, (3) somewhat difficult, (4) somewhat easy, (5) easy, (6) very easy].

The overall ease of use during the refill process is presented in Figure 3. Several of the tasks presented in Figure 3 are for a limited number of trials due to participants not being able to perform the task as required, see Table 2. For example, SCBAs 1, 3, and 6 had 50 to 88% of "unable to perform" ratings for connecting to the refill station with no assistance from others. In addition, the task of disconnecting from the refill station without assistance had a 50% "unable to

perform" rating for SCBA 6. Observation during testing and focus group data discussion showed that this inability was due to the refill port being located on the back of the SCBA for these devices. SCBAs 1 and 3 had "somewhat easy" to "easy" ratings for all refill tasks except for the connection tasks. SCBA 6 received "difficult" ratings for reading the gauge and knowing when the SCBA is full during the refill process. Examining the focus group data, it was found that participants rated SCBA 6 this way due to their inability to see out of the hood due to fogging of the face mask.

SCBAs 2, 4, and 5 received an overall positive ease of use rating as no major ease of use concern was found for the SCBAs. SCBAs 3 and 4 did not receive either a positive or negative rating for ease of use as the main ease of use concern was the location of the refill port on the SCBA; outside of the refill port location no ease of use concern was identified. SCBA 6 received an overall negative rating for ease of use due to the two major design issues identified: the hood and the location of the refill port.

Table 2: Unable to perform task (number of trials).

	SCBA					
	1	2	3	4	5	6
Connecting this SCBA to the refill unit with no assistance from others	4	0	7	0	0	7
Reading the gauge on this SCBA while connected to the refill unit	0	0	0	0	0	1
Knowing when this SCBA was fully refilled	0	0	1	0	0	1
Disconnecting this SCBA from the refill unit with no assistance from others	1	0	0	0	0	4

DISCUSSION

The study results show that a person can don and operate current commercially available SCBAs in a mining environment. Second, the results demonstrate that a person is capable of refilling the current commercially available SCBAs from a refill station in a mining environment when the refill port of the SCBA is located on the front of the SCBA. Lastly, the participants in this research did show a preference between current commercially available SCBAs and refill stations.

Several key findings from the research should be considered when developing mine escape SCBAs and SCBA refill stations. For example, the system performance of the SCBA must also allow for refill by the user alone. To accomplish this, the SCBA must have the refill port located on the front of the unit.

The results of this study can also be used to examine the existing SCBAs and refill stations tested to identify specific areas in which manufacturers could improve their devices for use during mine emergencies. In general a manufacturer could first examine the overall usability ratings to determine which usability construct should be addressed to improve the overall usability of the device. Once the area which needs attention is identified, the manufacturer could examine the findings specific to that construct to identify why the device received such a rating. When the reason for the rating is understood, the manufacturer could complete a design review of the device with specific attention to any negative findings and identify potential design modifications to address the concerns raised. This process will allow for the identification of critical issues which need to be addressed and potential modifications to improve the usability of the device during mine escape.

Only four participants were part of this study. In usability testing it is believed that 4 or 5 evaluators will find 80% of usability problems, adding more evaluators will provided diminishing returns, and that the most severe usability problems will be identified by the first few subjects (Virzi, 1992). Virzi also found that nearly all high-level and most of the medium-level concerns could be identified after only four evaluators. Given this, four participants would be sufficient to achieve the stated conclusions of the research question posed.

CONCULSION

The usability framework presented and the methods utilized during this study were able to answer the research question posed as part of this research. This methodology could be modified and used to test other lifesaving technologies to compare the usability of the devices and to estimate the ability of the devices to function as expected in a lifesaving situation. While the overall approach utilized is established in standards the exact details of how this testing was carried out could be useful to others who plan to carry out similar usability testing.

DISCLAIMER

The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health (NIOSH). Mention of any company or product does not constitute endorsement by NIOSH.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contributions of the following NIOSH researchers to the completion of this research effort: Jon Hummer for his assistance during data collection and preparation of the SCBAs. Gary Buchan, Alan Mayton, and Mary Ellen Nelson for their assistance during data collection and setup of the experimental equipment. Mahiyar Nasarwanji for his assistance during data collection. Kyle Stanyar and Bob Peters for their assistance conducting the focus group and analyzing the data. Jacquie Jansky for her assistance with data entry of the survey data collected.

REFERENCES

ISO 20282-2:2013. "Usability of consumer products and products for public use--Part 2: Summative test method." International Organization for Standardization.

Virzi, R.A. (1992). Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough? *Human Factors*, 34(4), 457-468. doi: 10.1177/001872089203400407