APPENDIX A.   PILOT STUDY

A.1   BACKGROUND

The success with which non-usability experts can use the critical incident technique may be linked to their ability to properly recognize and report critical incidents. This same assumption was made by Castillo (1997), and supports the notion of participants undergoing some form of critical incident training.   In this pilot study, a critical incident training tool was designed and validated to determine its effectiveness with respect to training non-usability experts.

 The most appropriate medium with which to disseminate the training program was an important issue. Castillo used a variety of media: videotape and verbal instructions. Due to its accessibility to an international network of users and its ability to rapidly distribute software at low cost, the World Wide Web (WWW) is an ideal medium upon which to develop a training program (Hilbert and Redmiles, 1998).   It would assure equal accessibility to both laboratory-based and remote users and permit easy integration of audio, video, and animation material (if deemed useful).

Another decision to be made was the specific approach by which to implement the training program. In order to accommodate remote users and the use of the WWW, a self-paced training approach was selected as most feasible. This would allow users to undergo training even when spatially and temporally separated from the experimenter. Therefore, the ideal transmission method and medium for remote evaluation via the critical incident technique was a self-paced on-line training tool.

A.1.1   Needs Assessment

The most important task in designing a training tool is to determine what goals are to be achieved and how. The main objective of the critical incident training program was to teach participants how to identify a critical incident and how to describe it so as to create a meaningful and accurate critical incident report. Training the user to perform these two tasks requires that development of a certain set of knowledge, skills, and abilities (KSAs). The KSAs deemed relevant for each task are listed in Table A-1.

Table A-1: Questions to be Addressed During Training

Task 1: Identify a critical Incident	Task 2: Report a critical incident
1. Knowledge of the definition of a critical incident.	1. Understanding of the purpose of a critical incident report form.
2. Knowledge of the definition of the critical incident technique.	2. Ability to access the critical incident form.
3. Understanding of why critical incidents are important and the ability to express these reasons.	3. Ability and skill in responding to questions contained on the critical incident report.
4. Knowledge of, and ability to recognize, the tell-tale features of a critical incident.	4. Knowledge of when to report a critical incident.
5. Ability to classify the critical incident as positive or negative.	5. Knowledge of how, and skill in submitting a critical incident report.

User participants exposed to the training program will primarily be non-usability experts, with minimal (if any) knowledge of, and experience with, the critical incident technique. The

assumption of minimal levels of each KSA listed in Table A-1 was considered accurate for the user participant population. The goal of the training was to provide the trainees with sufficient training such that they could demonstrate the following abilities during the remainder of the usability evaluation:

1. Correctly identify a critical incident that occurred during an interaction.
2. Access the critical incident training report form during or immediately following the completion of the task during which the critical incident occurred.
3. Properly identify the type of critical incident that occurred.
4. Properly identify the location in which the critical incident occurred.
5. Sufficiently describe the task and critical incident such that the experimenter could reconstruct the incident as it occurred.

A.1.2 Training Content Design Approach

The identification of two primary tasks, each having its own specific set of KSAs, suggested a natural division of training content into two modules, each one addressing one of the tasks. The next step required consideration of how to design the content within each of the modules. Guidelines for training material design were consulted. In keeping with Castillo's (1997) critical incident training program, the minimalist approach was adopted. This particular approach requires adherence to the principles and heuristics listed in Table A-2.

Table A-2: Principles and Heuristics for Designing Minimalist Instruction (van der Meij and Carroll, 1995)

	Principle		Heuristic
1	Choose an action-oriented approach.	1.1	Provide an immediate opportunity to act.
		1.2	Encourage and support exploration and innovation.
		1.3	Respect the integrity of the user's activity.
2	Anchor the tool in the task domain.	2.2	Select or design instructional activities that are real tasks.
		2.3	Components of the instructions should reflect the task structure.
3	Support error recognition and recovery.	3.2	Prevent mistakes whenever possible.
		3.3	Provide error-information when actions are error-prone or when correction is difficult.
		3.4	Provide error-information that supports detection, diagnosis, and recovery.
4	Support reading to do, study, and locate.	4.2	Be brief; don't spell out everything.
		4.3	Provide closure for chapters.

Supplementing these principles and heuristics are several suggestions by Harrison (1995). Based on empirical-based research, these suggestions are aimed at increasing the effectiveness of on-line help for procedural tasks, particularly in light of the inclusion of visuals within instructional text. The following suggestions fit within the context of critical incident identification and reporting:

1. Explanations providing goal structure or organizational information are presented at the beginning of the procedural task.

2. Procedural instructions are divided into small steps with each step labeled and presented in the order they are to be executed.

3. Information within each step is explicitly stated, and the most important information is presented first.
4. Visuals accurately depict the procedural step since users often prefer to follow the visual examples rather than the instructions and visuals tend to help eliminate orientation errors.
5. Visuals are accompanied by some form of written or spoken instructions in order to cue the user to the important aspects of the visuals.
6. Instructions are segmented to reinforce the concept of chunk or steps for completing the procedure.

By following these guidelines, a more accurate mental representation of the procedure and better connections from referential links to physical objects should be constructed.

A.1.2.1 Practice Exercises

The opportunity to practice concepts learned was considered as an important aspect of the training program content. According to Wiedenbeck et al. (1995), the most of effective way of implementing hands-on practice is via a well-conceived set of exercises, as opposed to exploration (wherein the choice of the practice is left to the discretion of the trainee). In the context of critical incident training, exercises would provide the user with exposure to critical incidents in the context of actual usage, thereby making the concepts learned easier to relate to task performance. However, as cautioned by Wiedenbeck et al.(1995), the benefits of exercises can only be realized if they are carefully designed and tested.

Thus, structured exercises were included at the end of each module. Four exercises were designed per module. Participants were required to complete each exercise and often demonstrate a certain level of performance (as defined by a criterion – for example, obtaining a certain number of correct responses) before continuing to the proceeding exercises. If participants were unable to meet the performance requirement, they were instructed to repeat the exercise a second time. Exercises were ordered in increasing complexity to increase user confidence. A more detailed description of each exercise is provided in Table A-3.

Table A-3: Description of Critical Incident Training Exercises

Module	Exercise Description	Purpose of Exercise	Performance Criterion
1	1. Users are asked to complete the following three exercises, each of which is designed according to a true/false format: a) Choose **1** of 3 definitions that most accurately describes a critical incident. b) Identify which of **8** critical incidents are positive or negative. c) Identify which of **10** signs and symptoms are better associated with a positive or negative critical incident.	To demonstrate: • knowledge of the definition of a critical incident. • knowledge of, and ability to recognize, the tell-tale features of a critical incident. • ability to classify the critical incident as positive or negative	Must receive a score of 10 or more out of 19 to proceed to next exercise.
	2. Users are presented with five scenarios, each of which describes a situation in which a person is interacting with a particular piece of software or hardware. The user is asked to identify which of the scenarios includes a critical incident.	• Knowledge of, and ability to recognize, the tell-tale features of a critical incident.	Must receive a score of 3 or more to proceed to the next exercise
	3. Users are asked to fill out an information form on a fictitious web site. The response format for the graduation date field is purposefully not specified (negative critical incident). When the form is submitted an error message is presented indicating that the wrong date format is used. Users are then asked to complete the information form a second time using the proper format. A message then appears informing the user that a critical incident just occurred and why it should be considered as such.	To force user to experience a negative critical incident such that they are better able to identify one when it occurs.	N/A
	4. Similar to the previous exercise. Users are asked to find the current temperature in Blacksburg (in °F). A list of links provided to obtain this information is designed such that a red blinking arrow appears in front of whatever link is currently active (i.e. under the cursor). This helps increase the salience of the selected link (positive critical incident). Once the user submits the temperature, they are informed that a critical incident occurred and why it should be considered as such.	To force user to experience a positive critical incident such that they are better able to identify one when it occurs.	N/A

Module	Exercise Description	Purpose of Exercise	Performance Criterion
2	1. Users are asked to complete the following three exercises, each of which is designed according to a true/false format: a. Identify which of **7** types of information should be provided when describing the task being performed. b. Identify which of **7** types of information should be provided when describing the critical incident. c. Choose amongst **3** possible responses for when a critical incident should be reported.	To demonstrate: • Ability and skill in responding to questions contained on the critical incident report. • Knowledge of when to report a critical incident.	Must receive a score of 12 or more to proceed to the next exercise.
	2. Participants were asked to repeat Exercise 3 from Module 1. This time, however, they were required to access the correct critical incident using the appropriate sequence of steps and correctly respond to the questions. Sample responses for the non-rating questions were provided to assist the user.	To demonstrate: • Ability to access the critical incident form. • Ability and skill in responding to questions contained on the critical incident report. • Knowledge of how, and skill in submitting a critical incident report.	
	3. Participants were asked to repeat Exercise 4 from Module 1. This time, however, they were required to access the correct critical incident using the appropriate sequence of steps and correctly respond to the questions. Sample responses for the non-rating questions were provided to assist the user.	Same as above	
	4. Participants were asked to perform certain calculations using an on-line calculator called Magi-Calc. This application was designed purposefully to cause participants encounter critical incidents. Participants are instructed to report at least two critical incidents that occur as accurately as possible. Sample responses were posted at the end of the exercise to allow users to evaluate their performance	Same as above plus: • knowledge of, and ability to recognize, the tell-tale features of a critical incident. • ability to classify the critical incident as positive or negative	

A.1.3   Training Tool Description

The design approach described in the previous section provided a set of requirement according to which a web-based training tool could be implemented. It was decided that the training tool would be based on a modular approach for reasons described earlier and would contain four main web page types: Introduction, Module 1, Module 2, and Conclusion.   Each of these web page types is described in the sections below.

<u>Introduction Web Page</u>
The Introduction web page is presented as Figure A-1. It serves as the home page for the training tool and presents information regarding the purpose, rationale, and overall content of the training program. To assist the user unfamiliar with web documents, it also gives instructions on how to interact with the tool. The Introduction Page links directly to Module 1.
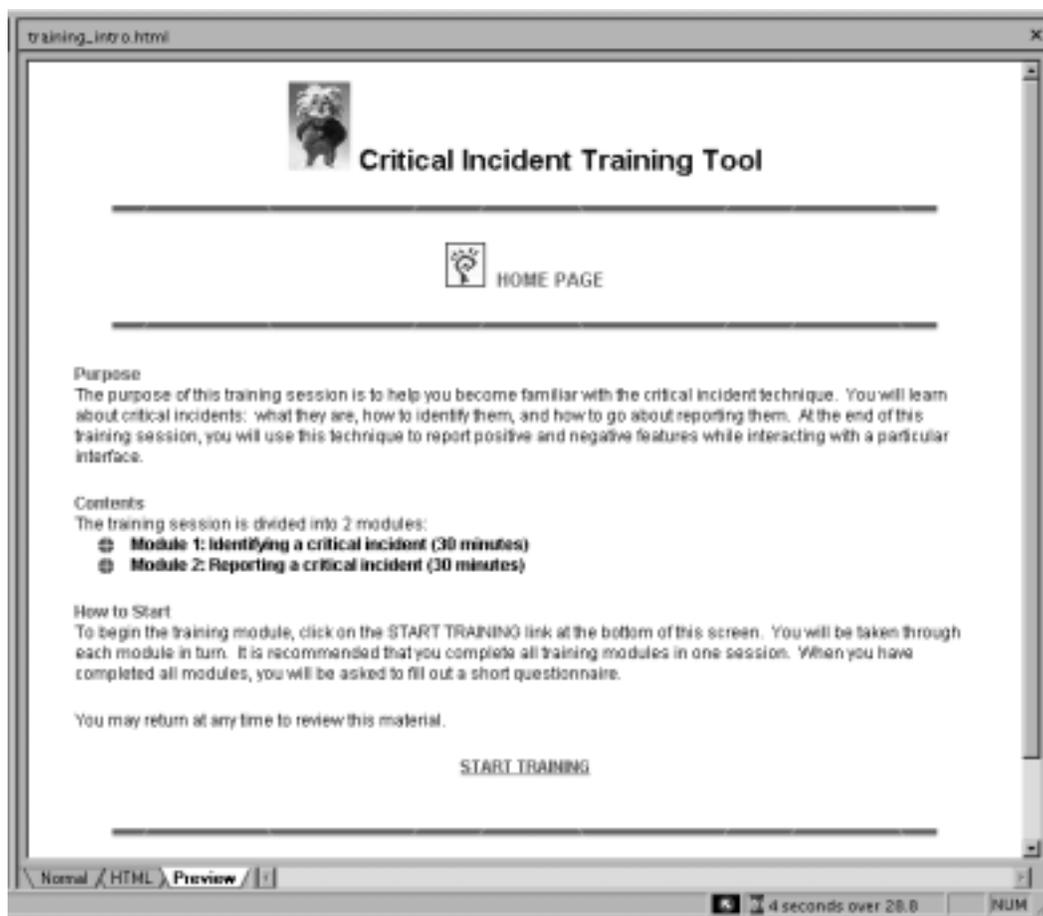
Figure A-1. Training Tool Introduction Screen

<u>Module 1</u>
Module1, as illustrated in Figure A-2, presents training content relevant to identifying critical incidents. It is comprised of three major sections: Introduction, Instruction, and Exercises. The purpose of the Introduction section is to present an overview of the content to be addressed in the form of a hypertext list of major topics.   Trainees can use this list to jump to sections appropriate to their interests. The Instruction section presents the training material relevant to identifying a

critical incident. Visuals supplement instructional text where permissible in accordance with Harrison (1995). Harrison found that the use of visuals within on-line help instructions for computer-based procedures enabled adult subjects to perform more procedural steps in less time and with fewer errors than subjects who received no visuals within on-line help instructions. Screen shots and working samples of objects used during the reporting task, such as the report button, information dialogue boxes, and critical incident report forms, are examples of visuals selected for inclusion.

The third section of the module contained the hands-on exercises described in Section A.1.2.1. Successful completion of these exercises was necessary to proceed to Module 2, although opportunities to link back to any one of the sections in Module 1 (ex. Instruction or Exercises) are also provided.

Figure A-2. Module 1 Web Page

Module 2
Module 2 was designed with similar layout and formatting as Module 1 to ensure consistency and flow amongst the module web pages. Unique to this module is the inclusion of review of Module 1 material, the purpose of which is to review key topics from the previous module. This module is also broken down into three sections and successful completion of the hands-on exercises is required to progress to the training conclusion page. Depicts a section of the Module 2 web page.
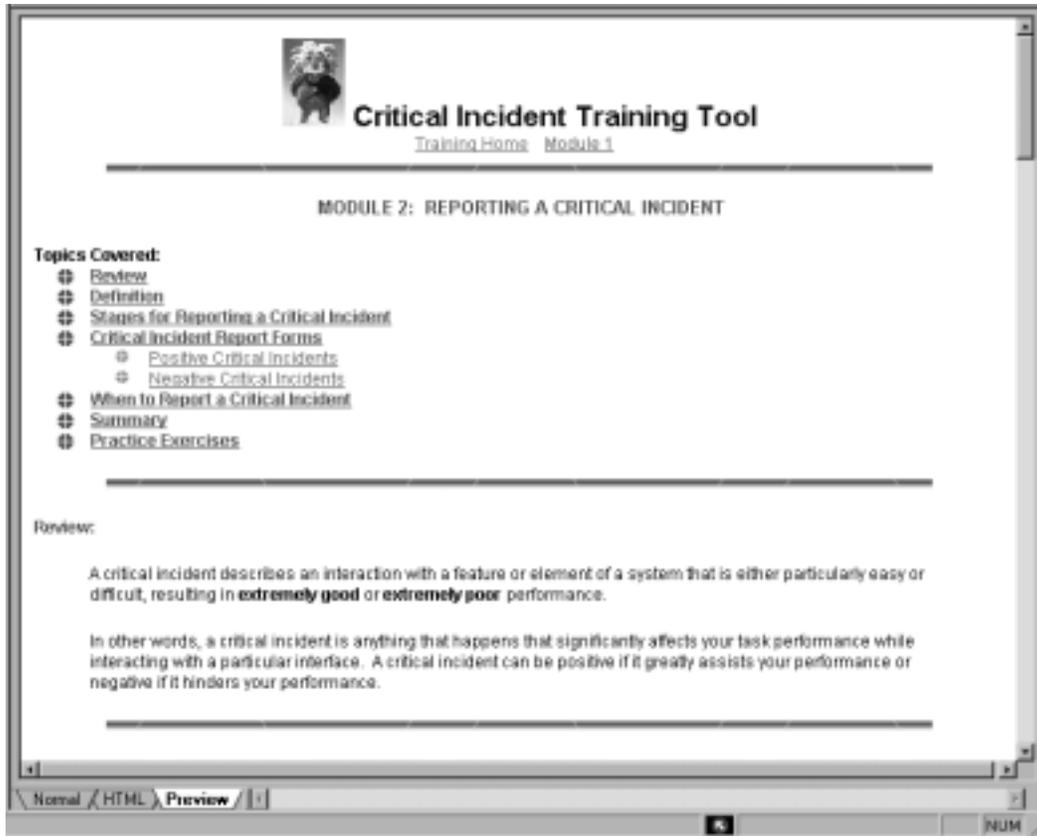
Figure A-3. Module 2 Web Page

Conclusion

The Conclusion Page provides an opportunity to review of key concepts associated with both modules and to provide the trainee with some final suggestions for ways in which to identify critical incidents (especially positive critical incidents). Figure A-4 provides a snapshot view of the Conclusion web page.

After reviewing the material on the conclusion page, the user will be given the opportunity to fill out a short questionnaire regarding the usability of the training program. A copy of this questionnaire is presented in Appendix E of this report. The questionnaire, once completed and submitted, is designed such that its contents are sent to the evaluator via email. Not only will this questionnaire help elicit subjective input from the user regarding the effectiveness and perceived usefulness of the training, but it can be used to track the effect of implementing design changes to enhance the training tool. Once the questionnaire has been submitted the user will be declared trained. It should be noted that a time limit will not be imposed at any point during the training program. The only requirement is that participants complete the modules in order and in one sitting.

155
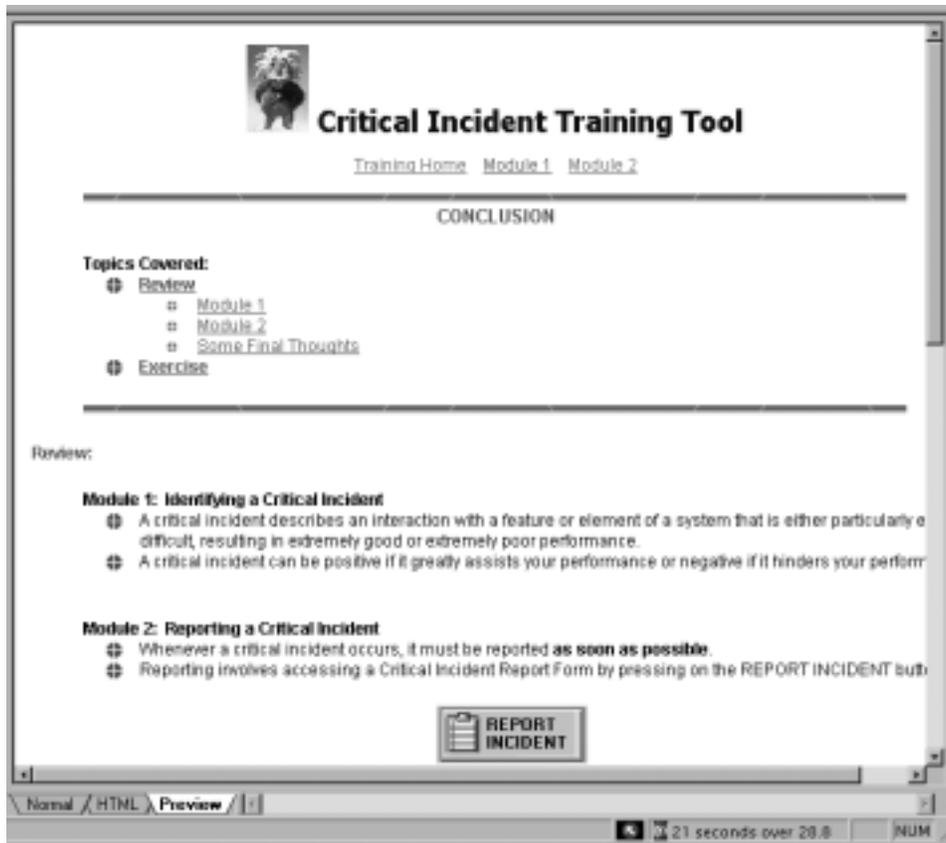
Figure A-4. Conclusion Web Page

A.2   OBJECTIVE

A pilot study was conducted to determine the effectiveness of the critical incident training tool design, and specifically, to determine the effectiveness of the training with hands-on exercise versus training with instructions only. Effectiveness was measured in terms of performance in identifying and reporting critical incidents during a web site usability evaluation. A secondary objective of this study was to conduct a formative evaluation of the critical incident report forms and in general, validate the design of the training and reporting tools.

A.3   EXPERIMENTAL DESIGN

The experiment was designed as a single-factor between-subject design. The independent variable of interest was the type of training: instruction only or instruction with exercise. The dependent variable was level of performance in identifying and reporting critical incidents during the interaction with a fictitious web site called Virtual Bouquet. The purpose of Virtual Bouquet was to allow on-line shoppers a means of creating their own customized flower bouquets. The web site was designed with five intentional critical incidents (ICI) embedded with its design and task structure. Performance of each training group was evaluated based on the number of these intentional critical incidents identified and reported. Additional measures were the total number of critical incidents reported (ICI plus user-defined), time spent performing the evaluation, time spent reporting either positive or negative critical incidents, and number of times in which help is accessed.

Post-test questionnaire data was collected to support a formative evaluation of the on-line critical incident report forms. This data included the ease with which the user was able to describe and rate the task and critical incident using the critical incident report form. Additional questions required the participant to rate the appropriateness of the ratings scales, the ease of accessing the report form, and the understandability of the questions. Also collected will be participant likes and dislikes (i.e. critical incidents) with regards to the on-line critical incident training tool and critical incident repot forms.

A.4   TEST HYPOTHESES

The following hypotheses were established for testing:

I.   Training with instruction and exercises leads to shorter evaluation times in comparison with training with instruction only.

II.   Training with instruction and exercises leads to shorter reporting times in comparison with instruction only.

III.   Training with instruction and exercises leads to a greater number of critical incidents reported in comparison with instruction only.

IV.   Training with instruction and exercises leads to fewer times in which a help function is accessed in comparison than that with instruction only.

V.   Participants given training with instruction and exercises will be more likely to score three or more (i.e. obtain a passing score) on a practice evaluation than those given training with instruction only.

A.5   PARTICIPANTS

Ten participants (8 males, 2 females) were recruited to participate in this pilot study. Age range was 21-38 years of age. Participants were considered eligible if they had minimal or no prior knowledge or experience with the critical incident technique, but were fluent in English and had at least three or more years experience using a computer. At least some experience using Microsoft Internet Explorer was also required. Participants were randomly assigned to one of two experimental conditions: instruction with exercise (I/E) and instruction only (I).

A.6   EQUIPMENT

Participants used a Gateway 2000 desktop PC computer with a 21" monitor and E-4200 processor. Testing was conducted on-line, using the Microsoft Internet Explorer web browser to access the critical incident on-line training tool, which was stored on the lab server (discus.ise.vt.edu).

A.7   INTERFACE DESCRIPTION

There were two interfaces with which the participants were required to interact during the experiment: the training system and Virtual Bouquet.

A.7.1   On-line Critical Incident Training Tool

A detailed description of the on-line critical incident training tool is provided in Section A.1 of this report. The contents of this tool were modified for the purpose of this pilot study to create an instruction-only training tool. The creation of this modified tool involved eliminating all hands-on exercises. Retained was all instructional text and a report incident practice opportunity, in which participants were asked to fill out a positive or negative critical incident report form based on a fictitious critical incident. This practice opportunity was not considered practice exercises

since performance level criterions were not enforced nor were feedback mechanisms in place for the user to evaluate their performance. In addition, it was considered important that all participants be able to demonstrate proficiency in executing the steps required to report a critical incident.

A.7.2 Virtual Bouquet

The interface evaluated by participants was a fictitious web site called Virtual Bouquet. The objective of Virtual Bouquet was to allow customers to create their own bouquet of flowers on-line and then ship them to someone anywhere in the world. The web site was comprised of four web pages: a home page (see Figure A-5), a bouquet creation page (see Figure A-6), an order form page (see Figure A-7), and a confirmation screen (see Figure A-8).

Figure A-5. Virtual Bouquet Home Page

Figure A-6. Virtual Bouquet Creation Web Page
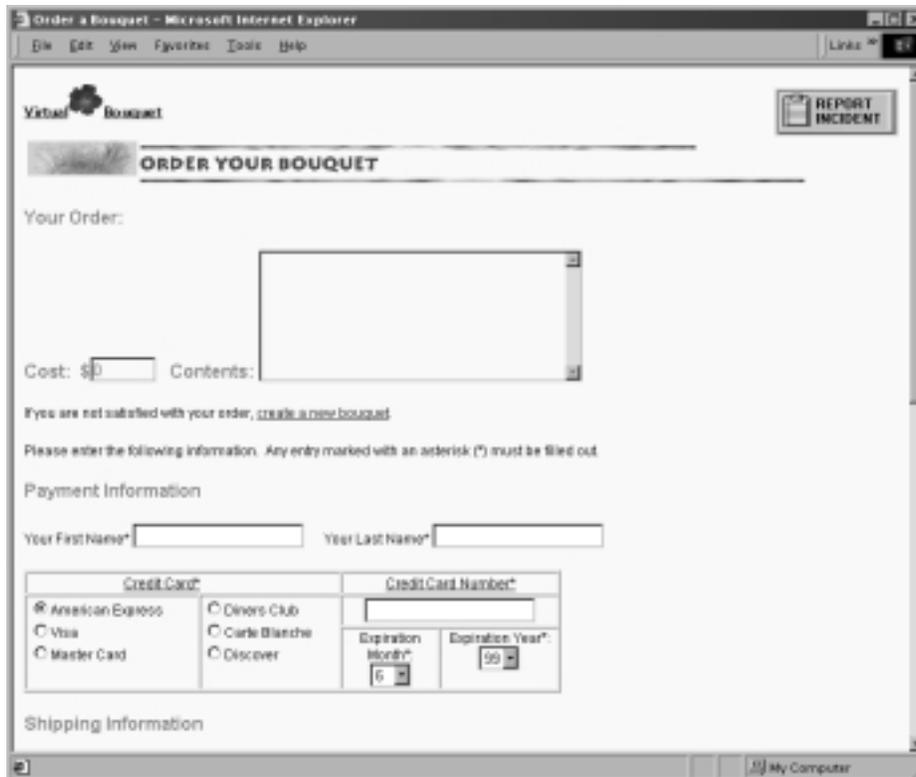
Figure A-7. Virtual Bouquet Order Form Web Page
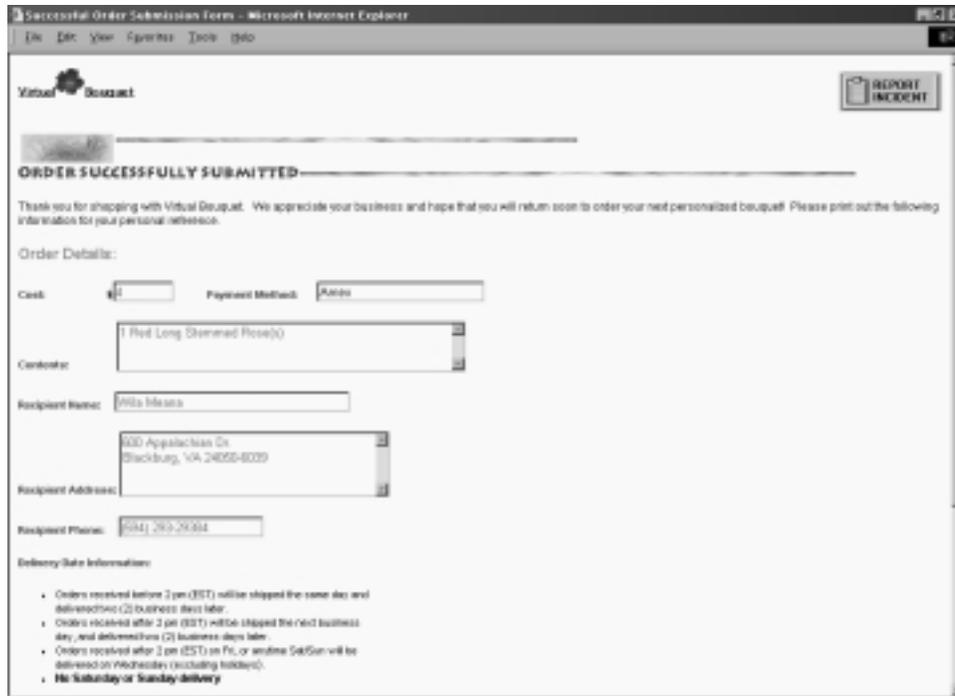
Figure A-8. Virtual Bouquet Order Confirmation Web Page

This web site was designed such that a user would encounter a particular set of critical incidents. The following 5 ICIs (4 negative and 1 positive) were embedded within the web site design and task structure:

1. Instruction font on bouquet creation page written in small font and had poor contrast with background color. (NEGATIVE)
2. It is not possible to make a change to the contents of the bouquet. The user must instead clear the entire contents of the bouquet and re-start the process. (NEGATIVE)
3. The method by which the user increments the number of a particular flower type is inconvenient: a mouse must be used to click on counter buttons (+ or -). There is no means by which to enter the desired number directly. (NEGATIVE)
4. The browser window in which the order form page initially appears is too small and non-resizable, making it difficult to view a significant portion of the web page contents (see Figure A-9). (NEGATIVE)
5. A prompt is given to verify the user that the order has been successfully submitted. (POSITIVE)

Figure A-9: Non-resizable Browser Window

For each participant, a score was tabulated based on the number of ICIs identified and reported. A score of 3 or above was considered a passing score.

A.8 TASK SCENARIOS

Participants were asked to undergo the training program to which they were assigned. At the completion of this training session, they were asked to evaluate a new interface using the critical incident technique. During this evaluation, participants were asked to consider the following scenario: "You would like to send your friend a bouquet of flowers for her birthday". They were then requested to complete the set of tasks listed below by interacting with the Virtual Bouquet web site as required and reporting any critical incidents encountered.

Table A-4: Summary of Scenario Tasks Assigned

Task	Description
1	Create a bouquet consisting of the following flowers to be delivered in a glass vase: • yellow carnations • 4 red tulips • sunflowers • 1 baby's breath
2	You decide that the bouquet is too expensive. Eliminate the two sunflowers.
3	The bouquet is now to your satisfaction. Indicate that you would like to submit the order.
4	Enter in your current address.

Task	Description
5	Indicate that you would like to pay by MasterCard: Number: 1234 5678 9012 3456 Expiration date: 08/99
6	Send the bouquet to the following address care of your friend: Joanne Friend 500 Sunridge Dr. Blacksburg, VA 24060 US Phone Number: (540) 222-3333
7	Select a birthday card and leave the message "Happy 21st Birthday, Joanne."
8	Place the order.

A.9   PROCEDURE

Upon arrival to the Human-Computer Interaction Lab (Whittemore 530), participants were asked to read an informed consent form (see Appendix E), and if in agreement with the terms and conditions, asked to sign the form to indicate their voluntary consent to participate. Participants were then given a short questionnaire designed to gather information regarding skill levels with respect to computers, web browsers, and the critical incident technique. A copy of the questionnaire is provided in Appendix E of this report.

Participants were then taken into a separate room containing a computer workstation. This workstation was comprised of an Internet-accessible PC desktop computer and 21" monitor. The on-line critical incident training program to which the participant was assigned was already open on the desktop (via the Microsoft Internet Explorer web browser). The participant was provided a short overview of the program and then asked to go through each module in sequence, with the option of reviewing the material or taking a break at any time.   Time to complete each module was recorded. Time required to complete each exercise and the score obtained on each exercise were recorded for participants assigned to the training with exercise program.

At the completion of the training session, participants were given the option of a short break, after which time they were asked to evaluate a web interface using the critical incident technique.   An on-line description of the web interface and of the tasks to be performed using that interface, was presented to the participant. Participants were told to complete the set of tasks as best as possible and that no more than ten critical incidents should be reported. It was emphasized that this limit was being imposed for time purposes and not because there were ten critical incidents to be reported.   The following data were recorded:

- time required to complete each critical incident report form
- number of times help was accessed during the reporting process
- time required to complete the evaluation
- total number of intentional critical incidents identified

A post-test questionnaire was handed out at the completion of the practice evaluation. The purpose of this questionnaire was to solicit subjective information regarding the participant's interaction with the training tool and the critical incident report forms. Participants were then reimbursed and thanked for their participation in the study. On average, the test session lasted one hour for instruction without exercise participants and two hours for those with exercises.

A.10 RESULTS

Data was gathered from three sources: the pre-test questionnaire, the evaluation session, and the post-test questionnaires. Results compiled from each data source are presented below. A level of significance of 0.05 was used for all analyses unless otherwise noted.

A.10.1 Pre-Test Questionnaire Results

All participants indicated having used computers for three or more years. Of interest was to determine whether the familiarity of, and experience with, the critical incident technique varied systematically amongst the groups. A one-way ANOVA was conducted to investigate this issue. An ANOVA was considered a suitable analysis approach since a true Likert rating scale was used to collect both familiarity and experience rating data.   In this analysis, a result of no difference is of interest. Accordingly, a level of significance of 0.2 will be used.

Table A-5. ANOVA Summary Table for Familiarity with the CIT Prior to Training

Source	DF	SS	MS	F	P
Condition	1	0.10	0.10	0.08	0.784
Error	8	10.00	1.25		
Total	9	10.10			

Since p=0.784>> 0.2, it can be concluded that there was no difference in the level of familiarity of the training groups.

Table A-6. ANOVA Summary Table for Experience with the CIT Prior to Training

Source	DF	SS	MS	F	P
Condition	1	0.400	0.400	1.60	0.242
Error	8	2.000	0.250		
Total	9	2.400			

Since p=0.242>0.2, it can be concluded that there was no difference in the level of experience of the training groups.

A.10.2 Evaluation Performance Results

All participants were required to apply the critical incident technique towards the evaluation of a fictitious web site in which ICIs were embedded.   Differences in performance between the two training groups were of interest. Several measures of performance were used, the results from which are provided below. A level of significance of 0.05 is used for analyses unless otherwise mentioned.

A.10.2.1 Time to Complete Evaluation

The time to complete the evaluation was recorded. A one-way ANOVA was used to determine if a significant difference in time existed between the two training conditions.   The results of this analysis are presented in Table A-7. There are no significant results in this analysis. There is insufficient evidence that a significant difference exists in the time required to complete the test evaluation, with a level of significance of 0.05.

Table A-7. ANOVA Summary Table for Evaluation Completion Times

Analysis of Variance for Time Eva					
Source	DF	SS	MS	F	P
Condition	1	14516	14516	0.05	0.835
Error	8	2496438	312055		
Total	9	2510954			

A.10.2.2 Time to Report Critical Incident

The time to report a critical incident was also recorded. Timing began when the participant pressed the REPORT INCIDENT button was pressed and ended when the SUBMIT button on the critical incident report form was pressed. Times for all critical incidents reported by a participant were averaged to give a single average reporting time. A one-way ANOVA was used to determine if a significant difference in average reporting time existed between the two training conditions. The results of this analysis are presented in Table A-8.

Table A-8. ANOVA Summary Table for Average Reporting Times

Analysis of Variance for Time Spent Reporting					
Source	DF	SS	MS	F	P
Condition	1	3518	3518	1.18	0.308
Error	8	23776	2972		
Total	9	27293			

There are no significant results in this analysis. There is insufficient evidence that a significant difference exists in the average time required to report a critical incident, with a level of significance of 0.05.

A.10.2.3 Total Number of Critical Incidents

Also recorded was the total number of critical incidents reported by each participant. A one-way ANOVA was conducted to determine if a significant difference existed in the total number of critical incidents reported by each training group. The results of this analysis are provided in Table A-9.

Table A-9: ANOVA Summary Table for Total Number of Critical Incidents

Source	DF	SS	MS	F	P
Condition	1	0.40	0.40	0.07	0.797
Error	8	45.20	5.65		
Total	9	45.60			

There are no significant results in this analysis. There is insufficient evidence that a significant difference exists in the total number of critical incidents reported, with a level of significance of 0.05.

A.10.2.4 Negative Critical Incident Reports

Of interest is to examine critical incidents at the positive and negative level to determine if any differences exist in the types of incidents identified and submitted between the training groups. The total number of negative and positive critical incidents reported by each participant were

recorded and analyzed separately via one-way ANOVA. The results of these analyses are presented in Table A-10 and Table A-11 respectively.

Table A-10.   ANOVA Summary Table for Total Number of Negative Critical Incidents

Source	DF	SS	MS	F	P
Condition	1	0.10	0.10	0.09	0.771
Error	8	8.80	1.10		
Total	9	8.90			

Table A-11.   ANOVA Summary Table for Total Number of Positive Critical Incident Reports

Source	DF	SS	MS	F	P
Condition	1	0.90	0.90	0.36	0.565
Error	8	20.00	2.50		
Total	9	20.90			

There are no significant results in these analyses. Hence, there is insufficient evidence that a significant difference exists in the total number of negative or positive critical incidents reported, with a level of significance of 0.05.

A.10.2.5 Number of Times Help Was Accessed

Not only were the numbers of critical incidents reported by each training group of interest, but so too were the times that help was accessed during the reporting process. Accessing help may indicate a lack of the appropriate level of knowledge or ability to answer the questions appropriately. A one-way ANOVA was conducted to determine if the training program had a significant effect on the number of times a participant accessed a help page. The results of this analysis are presented in Table A-12.

Table A-12.   ANOVA Summary Table for Total Number of Times Help was Accessed

Source	DF	SS	MS	F	P
Condition	1	36.1	36.1	2.43	0.158
Error	8	118.8	14.9		
Total	9	154.9			

A.10.2.6 Evaluation Test Score

The number of intentional critical incidents correctly identified and reported by the participant determined his or her overall test score. As described previously, five critical incidents were intentionally embedded within the Virtual Bouquet Web Site. A pass/fail system was adopted whereby a score of three or more correctly identified intentional critical incidents was treated as a pass. A score of less was treated as a failure. A Chi-Square Test of Independence was conducted to determine if the score obtained by the participants was independent of the training program to which they were exposed.

Statistical Hypothesis Test:
H_o: O = E
H_a: O ≠ E
$\alpha = 0.05$
Decision Rule: I reject H_o if $\chi^2_{obs} > \chi^2_{tab}$.

Table A-13. Contingency Table

	Evaluation Test Score		
	Pass	Fail	
Instruction + Exercises	2	3	5
Instruction Only	3	2	5
	5	5	10

Table A-14. Joint Probability Table

	Evaluation Test Score	
	Pass	Fail
Instruction + Exercises	2.5	2.5
Instruction Only	2.5	2.5

$\chi^2_{obs} = \Sigma\Sigma(O_{ij} - E_{ij})^2/E_{ij} = 0.4$
$\chi^2_{tab} = [df=(2-1)(2-1)=1] = 3.84$

Since $\chi^2_{obs} < \chi^2_{tab}$, fail to reject the null hypothesis. Therefore, the score obtained by participants is independent of the training program to which they were exposed.

A.10.3 Post Test Questionnaire Results

The post-test questionnaire was comprised of a series of questions that required the participant to rate some aspect of the training program or of the critical incident on-line reporting tool. The questions can be grouped in two categories related to the type of information they provide: training evaluation information versus training tool and reporting tool design information.   The former type is useful to investigate additional differences between the experimental conditions; that latter type is of greater use for supporting a formative evaluation of the training and reporting tools.

Questions pertaining to the knowledge, skills, and abilities gained via the training program are of interest to determine if participants assigned to different training conditions differed regarding their subjective evaluation of the program (differences that may not exist in objective performance measures). A five-point Likert-rating scale (ranging from 1=strongly disagree to 5=strongly agree) was used for all questions, allowing responses to be analyzed as interval data. A one-way ANOVA was conducted for each post-test question at a significance level of 0.05. The results of these analyses are presented below.

Question 1: The training helped me learn to **identify positive** critical incidents.

Table A-15. ANOVA Summary Table for Question 1

Source	DF	SS	MS	F	P
Condition	1	0.050	0.050	0.09	0.775
Error	7	3.950	0.564		
Total	8	4.000			

Question 2: The training helped me learn to **identify negative** critical incidents.

Table A-16.   ANOVA Summary Table for Question 2

Source	DF	SS	MS	F	P
Condition	1	0.56	0.56	0.56	0.480
Error	7	7.00	1.00		
Total	8	7.56			

Question 3: The training helped me learn to **report positive** critical incidents.

Table A-17.   ANOVA Summary Table for Question 3

Source	DF	SS	MS	F	P
Condition	1	1.606	1.606	2.85	0.136
Error	7	3.950	0.564		
Total	8	5.556			

Question 4: The training helped me learn to **report negative** critical incidents.

Table A-18.   ANOVA Summary Table for Question 4

Source	DF	SS	MS	F	P
Condition	1	0.050	0.050	0.18	0.685
Error	7	1.950	0.279		
Total	8	2.000			

Question 5: I feel better prepared to **identify** critical incidents after going through the training tool.

Table A-19.   ANOVA Summary Table for Question 5

Source	DF	SS	MS	F	P
Condition	1	0.022	0.022	0.07	0.798
Error	7	2.200	0.314		
Total	8	2.222			

Question 6: I feel better prepared to **report** critical incidents after going through the training tool.

Table A-20.   ANOVA Summary Table for Question 6

Source	DF	SS	MS	F	P
Condition	1	1.422	1.422	3.56	0.101
Error	7	2.800	0.400		
Total	8	4.222			

The above analyses do not show significant differences between the ratings given by the I/E group versus those by the I group regarding whether the training tool helped in the learning of, and preparation for, identifying and reporting a critical incident. These results are consistent with those obtained through analysis of objective performance measures.

The other type of information gathered via the post-test questionnaires was designed to support a formative evaluation of the design of the training tool and critical incident report forms. In this way, problems could be fixed prior to the thesis experiment. Accordingly, the responses allocated for each question were averaged across all training participants. Table A-21 presents

the responses pertinent to the training tool and Table A-22 presents those pertinent to the critical incident report form.

Table A-21. Average Ratings - Formative Evaluation of the Critical Incident Training Tool

	SUBJECT											
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10		
Question	I/E Group					E Group					AVERAGE	ST DEV
Simple to Use	5	4	5	5	4	5	4	5	5	4	4.60	0.52
Easy to Follow	4	4	5	5	4	5	4	5	5	4	4.50	0.53
Sufficient Info	5	3	N/A	5	4	5	4	4	4	3	4.11	0.78
Easy to Understand	4	4	5	5	4	5	4	4	5	3	4.30	0.67
Clear Organization	4	4	5	5	4	5	4	4	5	4	4.40	0.52
Liked Interacting	4	3	4	4	4	5	2	4	N/A	3	3.67	0.87

Table A-22. Average Ratings - Formative Evaluation of the Critical Incident Report Form

	SUBJECT											
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10		
Question	I/E Group					E Group					AVERAGE	ST DEV
Like Idea of Reporting	4	3	4	4	4	5	4	4	4	4	4.00	0.47
Ease of Accessing	4	1	4	5	5	5	5	5	4	4	4.20	1.23
Ease of Reporting	3	2	4	4	4	5	3	5	4	2	3.60	1.07
Questions Understandable	3	2	2	2	5	4	4	5	4	2	3.30	1.25
Questions Detailed	4	2	4	2	4	5	4	4	5	4	3.80	1.03
Motivated to Report +	4	1	3	5	3	5	4	5	4	4	3.80	1.23
Motivated to Report −	3	1	3	4	3	1	2	4	4	3	2.80	1.14
Easy to Describe Task	4	1	4	4	4	4	3	4	4	2	3.40	1.07
Easy to Describe CI	4	1	4	4	4	5	3	4	5	2	3.60	1.26
Ease of Rating Impact on Task Performance	4	4	4	5	3	5	3	4	4	3	3.90	0.74
Ease of Rating Impact on Satisfaction	5	4	4	5	3	5	4	4	4	4	4.20	0.63
Ease of Rating Error Severity	4	4	5	5	3	5	3	4	4	4	4.10	0.74
Scales Appropriate	4	4	4	4	4	5	4	4	5	4	4.20	0.42

Also gathered were participant likes and dislikes regarding the critical incident technique training tool.

Table A-23. Negative Aspects Reported for the Critical Incident Training Tool

Description	Frequency (/17)
Lots of text; wordy	2
Duplicate information given.	1
Boring to go through	1
Lack of confidence regarding the ability to identify a critical incident.	1
Help information should be displayed within or in the same window as the critical incident form.	1
Information difficult to remember	1
More simplistic wording to describe tools needed.	1
Positive critical incident example not good.	1
Provision of a top navigation bar not appropriate if a particular sequence of web pages desired.	1
Excessive scrolling	1
Integrate description of steps with implementation of the steps.	1
Return to Training Link did not work.	1
More examples needed.	1
Multiple windows make it difficult to follow and keep track of material.	1
Lack of feedback when filling out critical incident report form.	1
Inconsistency in the way in which help information is presented (i.e. pop-up window versus message box).	1

Table A-24. Positive Aspects Reported for the Positive Critical Incident Training Tool

Description	Frequency (/16)
Examples were helpful.	4
Easy to understand; intuitive	4
Ease of reviewing material (using links).	2
Easily accessible	2
Easy to follow	1
Information covered quickly enough so as not to get boring.	1
Feedback was useful.	1
Good organization and content.	1

A.11 DISCUSSION

In this section, the results are discussed with respect to their implications towards the on-line critical incident training and reporting tool designs.

A.11.1 Pre-test Questionnaire Results

The questionnaire results indicate that there were no systematic differences amongst the groups with respect to computer experience, level of familiarity with the critical incident technique, and experience in using the technique. These results indicate that training groups were initially equivalent. Equivalence ensures that systematic differences in measures taken from each group are more likely to be attributed to differences in training programs.

A.11.2 Performance Measures

Analyses were conducted to determine if the incorporation of exercises in an on-line critical incident training program results in difference in performance, where performance was measured in the context of a (fictitious) web site usability evaluation. Specifically, the time spent evaluating the interface and reporting the incidents, the number of incidents generated (total, positive, and negative), the number of times help was accessed during the reporting process, were analyzed, and the score obtained. The results of the analyses failed to demonstrate significant differences for any of these measures. Therefore, the following conclusions can be made with 95 % confidence:

1. There is insufficient evidence of a significant effect of exercises on the time spent reporting total, positive, and negative critical incidents while performing a test evaluation
2. There is insufficient evidence of a significant effect of exercises on the time spent performing a test evaluation.
3. There is insufficient evidence of a significant effect of exercises on the average time spent reporting a critical incident.
4. There is insufficient evidence of a significant effect of exercises on the number of times help was accessed during the reporting of a critical incident.
5. Insufficient evidence of a significant effect of exercises on the score obtained during the test evaluation.

The implication of these conclusions is that comparable levels of performance can be obtained without hands-on exercises. This result counteracts the expectation that exercises aid learning of skills related to identifying and reporting critical incidents. The conclusion of no difference, however, has several practical ramifications with respect to the thesis experimental protocol and the nature of critical incident technique training. First, it justifies the use of an instruction-only training program, which is much shorter in duration than that with exercises (i.e. ½ hour on average versus 1 hour). This will allow the introductory sessions to be completed in less time. Second, it implies that the skills, abilities, and knowledge necessary to implement the critical incident technique are such that practice to achieve mastery is not needed. In other words, the technique does not require a very extensive background/education to be learned and applied; it is not exclusive to those with a formal background in usability engineering. This provides support for its utility in remote evaluation wherein users (and not usability experts) are the reporters.

A.12 POST-TEST QUESTIONNAIRE RESULTS

The questionnaire responses were analyzed for one of two reasons: 1) to determine if the incorporation of exercises results in difference in subjective evaluation of the training program and 2) to identify usability problems (or successes) with which to drive a redesign of the training and reporting tools.

A.12.1 Subjective Measure Analysis

Not only were performance measures of interest, but so too were subjective measures regarding the effectiveness of the training program. The results of the analyses failed to demonstrate significant differences for any subjective measures. Therefore, it can be concluded, with 95 % confidence, that there is insufficient evidence of a significant effect of exercises on the subjective ratings of the training program, with respect to the extent to which skills of identifying and reporting were learned and the ease with which those skills could be implemented. These results are consistent with those obtained in performance measure analysis, and hence, further substantiate the use of an instruction-only critical incident training tool.

A.12.2 Critical Incident Training Tool Formative Evaluation

Ratings below a level of 3.5 (out of 5) were highlighted as potential problem areas. All rating responses presented above exceed this minimum threshold, failing to indicate any particular problematic areas. The lowest rating allocated was that in response to the statement "I like interacting with the training tool". This result was expected due to the lack of perceived relevancy of the training content. Relevancy, a key means by which to foster learner motivation and interest, was difficult to convey to pilot study participants for whom there was no requirement to use the technique beyond the training session.   The thesis experiment will likely present a more realistic application for the training tool, since participants will be asked to apply the technique over a period of five days.

The negative aspects reported by the participants covered a broad range of issues, including navigation, feedback, and training material content. Regarding the latter, users stated that the content was long, repetitive, difficult to remember, and boring to go through. Many of these comments may again be linked to a lack of perceived relevance, which affected trainee motivation. These comments also help to demonstrate one of the disadvantages of using primarily print material to convey training content; that is, it can bore readers if too lengthy or written using non-user centered terminology.

Many participants were in agreement with respect to the positive aspects of the training tool. For instance, many participants stated that the examples were helpful. This finding is incongruent with performance measures, which indicted that examples and other exercises had little effect. This finding may suggest that the major benefit of having examples and exercises is to boost trainee satisfaction and confidence, and in turn increase acceptance of the tool. It is interesting to note that there are several instances in which the positive and negative aspects conflict with each other (ex. "Lots of text; wordy" versus "Information covered quickly enough so as not to get boring"). These discrepancies demonstrate the extent to which individual differences can effect the perception and evaluation of the training tool.

A.12.3 Critical Incident Report Form

The response ratings for the critical incident report form indicate several areas that fall below or very close to the criterion level of 3.50. The lowest ranking statement was "I was motivated to report positive critical incidents." This lack of motivation has been observed elsewhere. For example, Castillo (1997) observed that non-usability expert user-subjects trained to recognize critical incidents during their own usage of a software application rarely, if ever, identified positive critical incidents. Contributing to this lack of motivation may be the expectation that successful task performance should be the norm, rather than the exception.

Other lower ratings were assigned to the ease of reporting, and more specifically to describing the task and critical incident, and the understandability of the questions. These issues are interrelated and, from post-test discussions with the participants, were a by-product of the way in which the critical incident report form was designed. For instance, task and critical incident description input fields were designed each as a single multi-text input field, as illustrated in Figure A-10 and Figure A-11. Within each field, keywords were written to denote the various topics that should be addressed within the description.

```
INTERFACE:

STATUS (ex. Help Page open):

INTENTION:

ACTIONS PERFORMED:
```

Figure A-10. Task Description Input Field Design

```
WHEN:

WHAT:

WHY:

EXPECTATION:

RECOVERY:
```

Figure A-11. Critical Incident Description Input Field Design

Two problems with the input field design existed. First, participants demonstrated a tendency to use the TAB button to move from keyword to keyword. This strategy was ineffective since the keywords were embedded within a text field rather than in separate input fields. Rather than move to the next keyword, cursor control went to the proceeding input field. User frustration often resulted.

The second problem encountered was that participants had difficulty in the proper interpretation of the keywords.   This was particularly true of those given in the critical incident input field. For example, participants often interpreted "Why" as *why* the critical incident occurred versus *why* they considered it to be a critical incident. The former interpretation is undesirable since it carries the implication that the *user* caused the critical incident to occur rather than the interface. Another, more direct, means of eliciting this information is necessary. Also confusing was the "When" keyword. Participants often responded to this by describing the task that they were performing. This information was already elicited from the task description input field, causing the participants to perceive the questions as being highly redundant. The intended use of the "When" keyword was to explain where in the task the incident occurred. It was observed, however, that the scope of most task descriptions only included the task during which the critical incident occurred, making the inclusion of a "When" keyword unnecessary.

Problems with task description keywords were also observed, albeit to a lesser extent. For example, participants were unfamiliar with the use of the term "interface" to denote a software application or web page. A more user-centered term or an example illustrating the term is necessary. The "Status" keyword also caused confusion, again most likely due to a lack of familiarity with the terminology. The purpose of this keyword was to encourage participants to describe the state of the interface at the time of the critical incident, including the exact web page,

command, dialogue box, etc. that was active (i.e. being used). Post-test discussions with participants revealed that a more clear and intuitive keyword phrase might have been "Where in Interface".

A.13 ON-LINE CRITICAL INCIDENT TRAINING AND REPORTING TOOL REDESIGN

The results obtained in this pilot study supported a redesign of both the critical incident training tool and report form.

A.13.1 Training Tool Redesign

The most significant change made to the training program was the elimination of practice exercises. This change was substantiated by the lack of significant performance differences amongst participants who received and did not receive the practice exercises. Changes were implemented by simply linking the instructional components of each module together.

Due to the number of positive comments made regarding the use of examples, it was felt that at least one example should be retained from the practice exercise set. The purpose of this example would be to force the user to experience a critical incident and then to follow through the necessary action sequence to report that incident using the on-line critical incident tool. Only one example could be retained due to time constraints. The negative critical incident example was selected since comments indicated that this example provided the highest added value of the two. The exercise was placed at the start of the conclusion module for review purposes.

Many participants experienced confusion when required to manipulate multiple browser windows. This problem was resolved by sequencing all web pages in a linear fashion so that only one window was needed for the whole training session. Links were provided to allow the user to backtrack to various points in the training program if review of past material was deemed necessary or worthwhile.

Links or buttons embedded within the body of a web page also generated confusion. For example, participants often voiced uncertainty about whether or not to follow an embedded link, citing concerns that they would not be returned to the original page. To minimize this ambiguity, a page break was inserted after each embedded link, creating in multiple web pages per module. The benefits of this approach are that it forces users to follow each link in order to progress through the training tool and helps retain the sequential flow of web pages. The process by which embedded buttons (and links) were eliminated is depicted in Figure A-12.
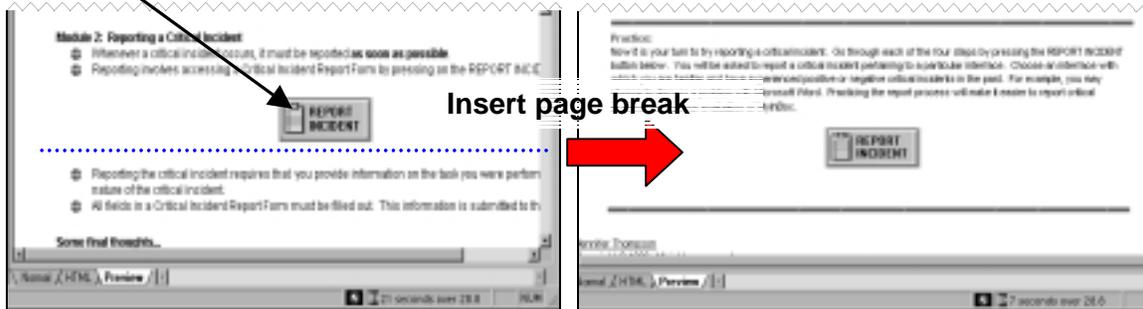
Figure A-12. Page Break Insertion to Eliminate Embedded Button

Training material corresponding to the reporting process was also modified. Originally, an explanation of the process was given and then the participant asked to report an incident. Comments indicated that a more intuitive approach would be to explain each step as the participant went through each one. The implementation of this redesign is illustrated in Figure A-13.
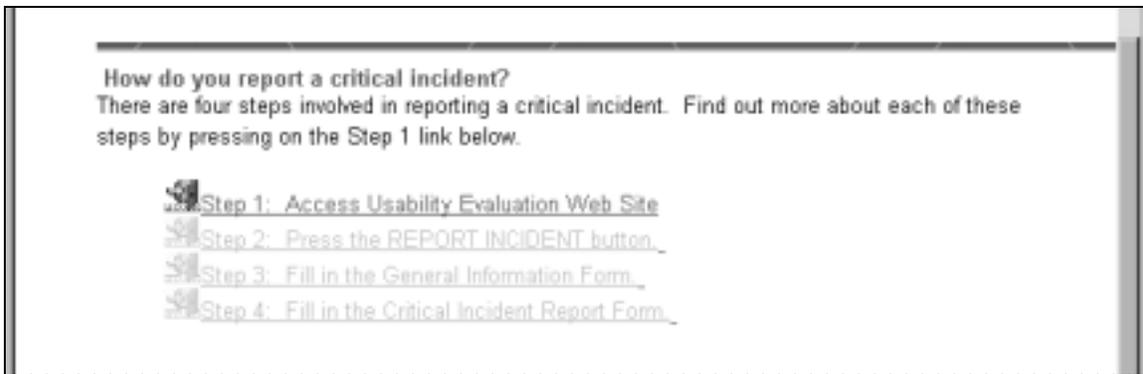

Figure A-13. Implementation of Combined Explanation and Action Approach to Learning how to Reporting Critical Incidents

All of the aforementioned changes were adopted as described. The modified training tool was then used for training of experimental user-subject and usability expert participants.

A.13.2 Critical Incident Report Form Redesign

Discussions with participants revealed that more explicit and direct questions, versus keywords, and examples were necessary to accommodate novice-level skills in usability evaluation and the critical incident technique. These suggestions, together with the observed need to support TAB key navigation, led to a redesign of the critical incident report forms. The major changes made are described below.

The first change made was to abandon the embedded keyword approach. Instead, the questions "Describe the task you were performing when the critical incident occurred" and "Describe the critical incident in detail" were broken down into a series of questions designed to elicit the most valuable information without overlap without keywords.

For example, the task description section was broken down into four input fields: the interface being used by the reporter, his or her location in the interface, the overall task objective, and an explanation of how the task was carried out. Figure A-14 illustrates these four input fields. The critical incident description section was also broken down into four input fields, as shown in Figure A-15. To expedite the reporting process, all input fields were designed to be as quick to fill out as possible without limiting the user's ability to accurately describe a variety of tasks and critical incidents. Examples were included directly on the critical incident report form, as recommended by the pilot study participants. The information button concept was retained to provide a more detailed set of instructions on how to answer each question. When the button is pressed, a pop-up window appears which the user can keep open or can close once its contents have been reviewed. Finally, questions were written to be as descriptive and as user-centered as possible and a table format was selected to better distinguish one input field from the next.

Figure A-14. Redesigned Input Fields for Task Description

Figure A-15. Redesigned Input Fields for Critical Incident Description

Also modified were the rating scale descriptors for the following questions: 1) Rate the impact of the critical incident on your task performance (positive critical incident report form) and 2) Rate the impact of the critical on your satisfaction with the system or interface (positive and negative critical incident report forms). In both of these questions, the term "significant" was used to denote the highest ranking of impact on task performance or satisfaction, with "major", "minor", "negligible" and "none" denoting progressively lower rankings. However, participants expressed difficulty in distinguishing between "major" and "significant".   It was decided that "extreme" was a more appropriate and meaningful descriptor to denote the highest impact ranking. All of the aforementioned changes were adopted as described. The modified training report form was then used for the collection of experimental data.

APPENDIX B.   **TASK SCENARIOS AND EMAIL MESSAGES**

B.1 DAILY SCENARIO DESCRIPTIONS

B.1.1 Day 1 Daily Scenario

Your scenario for Day 1 asks you to register as a VEMS user and configure your account for the remainder of the evaluation. At the completion of this scenario, you should be familiar with both the VEMS web site and voice system. When you have completed the scenario, press the SCENARIO IS COMPLETE button to return to the main scenario page.

If more convenient, print out a copy of this page for your reference.

TASK 1: Register Yourself as a VEMS User

DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS

1. Access the VEMS web site.

2. Go to the Registration web page.

3. Register in a Corporate Account.

4. The corporate account number is ########.

5. Fill in the appropriate information. Choose an account number that you can easily remember – it is suggested to use your home phone number. You will be required to use this account number throughout the experiment.

6. Copy down your account number, pass code, and email address. You must give your account number to the experimenter.

ASK 2: Set up Personal Profile Page

DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS

1. Proceed by clicking on the Setup Email Handling Link. However, you do not need to make any changes to your email handling.

2. Go to your Personal Profile Page.

3. Add Mark Lillehammer (your supervisor) to your Address Book. His email address is your_supervisor@techpointer.com.

4. Add Jennifer Thompson at jethomp6@vt.edu (your experimenter) to your Address Book.

5. Prioritize emails from Jennifer Thompson (jethomp6@vt.edu).

6. Your co-worker, Madeline Finch, has just requested that she be added to your Address Book. Use the ExpressLane to QuickAdd Madeline Finch (madeline_finch@techpointer.com) to your personal directory. NOTE: The ExpressLane feature is located on your Personal Profile Page.

TASK 3: Test out your VEMS Telephone Account

DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS

Now that you have set up your VEMS account, it is important that you practice using the telephone account. Recall that the telephone service uses Automatic Speech Recognition to interpret your voice commands and uses Speech Synthesis to read you your email messages.

Here are some helpful hints:

- When speaking to the system, it is important that you speak clearly (but do not over-emphasize your words) and at a normal rate.

- If you hear a double-beep, the system has not understood your voice command and is requesting you to repeat the command.

- You can always exit the system by saying goodbye. Interrupt the system by pressing the # key.

Instructions:

1. Dial into the VEMS voice system using the following number.

2. Say your account number and password. DO NOT ENTER THESE NUMBERS USING THE KEYPAD.

3. Listen to any new messages in your mailbox.

B.1.2   Day 2 Daily Scenario

Today is your second trial day for using the system. This is just a reminder that you should be reporting any critical incidents that you encounter. Remember that these are aspects of the voice and/or web interfaces that result in **extremely poor or extremely good performance**. To report critical incidents simply press the REPORT INCIDENT button.

When you have completed today's scenario, press the SCENARIO IS COMPLETE button to return to the main scenario page.

If more convenient, print out a copy of this page for your reference.

TASK 1: Read Email

DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS

1. Dial into the VEMS voice system using the following number: **1-888-694-6269**. If you are in the lab, remember to dial "9" first.

2. Read all email. If the sender of an email requests that you perform a certain action, perform it. Do not delete any email messages unless instructed to do so.

3. Prioritize messages from Graham Roeburg (stay on the phone to do this).

4. Dictate a new message to your supervisor (Mark Lillehammer - your_supervisor@techpointer.com). Tell him the hotel at which your meeting in Chicago is taking place (TIP: Graham Roeburg sent you this information in his email).

TASK 2: Find out more about the VEMS Web Site

<p align="center">**DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS**</p>

1. Go to your web page account.

2. Add Graham Roeburg (**graham_roeburg@techpointer.com**) to your personal Address Book.

TASK 3: Prepare for your Meeting

<p align="center">**DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS**</p>

You have just found out that you are to attend a meeting in Chicago tomorrow. Using your VEMS web account, do the following:

1. Add a new reply.

2. Give the reply a suitable name.

3. This reply should say that you will be attending a meeting in Chicago but that you reply to email upon your return.

B.1.3   Day 3 Daily Scenario

Today is DAY 3 of your evaluation of VEMS. You are to pretend that you are attending a business meeting in Chicago.

REMINDERS:
- Report any critical incidents that you encounter. Remember that these are aspects of the voice and/or web interfaces that result in **extremely poor or extremely good performance**. To report critical incidents simply press the REPORT INCIDENT button.
- When you have completed today's scenario, press the SCENARIO IS COMPLETE button to return to the main Daily Scenario page.
- Print out a copy of this page if you are having difficulties alternating between windows.

TASK 1: Manage Email

<p align="center">**DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS!**</p>

It's lunchtime and you finally get a chance to leave the meeting room and check your messages.

1. Phone your VEMS account. Remember to dial "9" first if you are in the lab.

2. Read all email. If the sender of an email requests that you perform a certain action, perform it. Do not delete any email messages unless instructed to do so.

3. Reply (do not dictate) to messages from Donna Hannun and Mark Lillehammer. Use the reply from your reply list that you created yesterday (i.e. the one that indicates that you are at a conference and will be away for the day).

4. You are getting tired of receiving Travel Information emails. Using phone commands, exclude emails from that sender (Travel Agent, travel_info_99@yahoo.com).

B.1.4   Day 4 Daily Scenario

Today is DAY 4 of your evaluation of the VEMS system. You are to assume that you have arrived back from your meeting in Chicago and are back in Blacksburg.

REMINDERS:
- Report any critical incidents that you encounter. Remember that these are aspects of the voice and/or web interfaces that result in **extremely poor or extremely good performance**. To report critical incidents simply press the REPORT INCIDENT button.
- When you have completed today's scenario, press the SCENARIO IS COMPLETE button to return to the main Daily Scenario page.
- Print out a copy of this page if you are having difficulties alternating between windows.

TASK 1: Manage Email

DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS!

You're back in the Blacksburg after a productive meeting in Chicago. You decide to check your email at a payphone while waiting at the drycleaners.

1. Phone into your VEMS account. Remember to dial "9" first if you are in the lab.

2. Read all email. If the sender of an email requests that you perform a certain action, perform it. Do not delete any email messages unless instructed to do so.

3. You notice that you have a lot of Weather Information email building up. To help you clean up your inbox, remove all existing email messages from **weather_info@usa.com**.

4. A co-worker of yours, Graham Roeburg, will be taking over the meeting regarding the UpLink contract. You promised Graham that you would forward him the email regarding the meeting. This email was sent to you from your supervisor (Mark Lillehammer) yesterday with subject header "Meeting Details". Find this email and forward it to Graham Roeburg.

5. Reply to the email from Donna Hannun using the reply called "Thank You" from your reply list. This is a default reply that you think is appropriate to send.

TASK 2

DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS!

You finally get to work after waiting for an hour at the drycleaners.

7. Go to your web page account.

8. Delete the reply that you created regarding the conference in Chicago.

9. You just found out that the Jennifer Thompson's email address has been changed to **jenn_syde@hotmail.com**. Make the appropriate changes to your address book.

10. Prioritize email messages from your supervisor, Mark Lillehammer.

B.1.5 Day 5 Daily Scenario

Today is DAY 5 of your evaluation of VEMS. It is the last day of the evaluation - CONGRATULATIONS!

- Report any critical incidents that you encounter. Remember that these are aspects of the voice and/or web interfaces that result in **extremely poor or extremely good performance**. To report critical incidents simply press the REPORT INCIDENT button.
- When you have completed today's scenario, press the SCENARIO IS COMPLETE button to return to the main Daily Scenario page.
- Print out a copy of this page if you are having difficulties alternating between windows.

TASK 1: Manage Email

DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS!

You're working at home today on some proposals. You decide to call into your VEMS account to see if you have received any emails.

1. Phone into your VEMS account. Remember to dial "9" first if you are in the lab.

2. Read all email. If the sender of an email requests that you perform a certain action, perform it. Do not delete any email messages unless instructed to do so.

3. You realize that you have a lot of email building up in your inbox. Remove all messages that are more than 1 day old.

4. Send the reply called "Great Idea" from your reply list to your best friend Jesse (**yourbest_friend@usa.net**).

TASK 2: Configure Web Page

DO NOT FORGET TO REPORT YOUR CRITICAL INCIDENTS!

1. Go to your web page account.

2. Since this is the last day of the study, you are now able to configure your email account to your liking. Try setting up your email handling such that it will poll your Virginia Tech account. You can change these settings after you have completed this scenario.

TASK 3: Fill out Questionnaire

Since this is the last day of your evaluation, I ask that you complete a questionnaire regarding your feedback about VEMS. When you are done the questionnaire, you will be returned to this page.

B.2 EMAIL MESSAGES

B.2.1 Day 2 Email Messages

TOTAL NUMBER SENT: 5

Number	
Sender	Jennifer Thompson,
Subject	Daily Reminder from the Experimenter
Body	Dear Participant, Thank you again for participating in this study. Your efforts are greatly appreciated and will be of great help in improving the VEMS interface. Today is Day 2 of your evaluation. Also, will you please forward this message on to Mark Lillehammner, your work supervisor, to alert him of your participation in this study? Thank you. Sincerely, Jennifer Thompson
Tasks Requested	**Forward** message to Mark Lillehammer

Number	
Sender	Travel Agent, travel_info_99@yahoo.com
Subject	Travel Info: Ninety-Nine Dollar Fares for Fall
Body	Dear Travel Information Recipient. Major airlines, led by America West, are offering one-way cross-country fares for as little as $99 between select cities this fall. The fares are good for travel between August 16 and December 6, 1999. Tickets must be purchased by July 28, 1999. Examples of fares available for $99 (one-way, based on required round-trip purchase) include: Columbus to Los Angeles Tampa to Phoenix Ft. Lauderdale to Los Angeles Austin to Seattle To qualify for the low fares, an overnight stay is required, and tickets must be purchased 14 days in advance. Some blackout dates apply. Fares are subject to change. Seats are limited and may not be available on all flights. Tickets are non-refundable. Taxes, surcharges, and passenger-facility charges of $3-$12 per round trip apply, depending on itinerary. Other restrictions and rules may apply. Make your fall travel plans now with Preview Travel!
Tasks Requested	none

Number	
Sender	Billy Bob, friend_billy@writeme.com
Subject	Quick Hello
Body	Hey there! It has been awhile since we last talked, so I thought I would email you and see how you are doing. Still enjoying Blacksburg? How is the job going? I hear you're participating in a voice usability study. That's pretty cool. Anyways, life in Montana is wonderful. The people here are so down to earth and relaxed. Nothing like New York. I'm so glad I decided to move out here. Well, have to run. I'll get in touch with you soon. In the meantime, hope everything is going well. Bye, Billy Bob
Tasks Requested	none

Number	
Sender	Weather Info - weather_info@usa.com
Subject	Blacksburg Weather – [date]
Body	*Note: the contents of this email change on a daily basis to reflect the current weather in Blacksburg* LOCAL FORECAST FOR BLACKSBURG Tonight...cloudy with areas of drizzle. A chance of rain after midnight. Low 55 to 60. East wind 5 to 10 MPH. Chance of rain is 50 percent. Wednesday...cloudy and continued cool with a chance of light rain. high in the middle 60s. Southeast wind 5 to 10 MPH. Chance of rain is 50 percent. Wednesday night...cloudy with a chance of rain. Low around 60. Chance of rain is 30 percent A LOOK AHEAD On Thursday the forecast calls for a high of 81 and a low of 59. Conditions for the day will be showers/sun. On Friday the forecast calls for a high of 83 and a low of 62. Conditions for the day will be partly cloudy.
Tasks Requested	none

Number	
Sender	Graham Roeburg - graham_roeburg@techpointer.com
Subject	Meeting in Chicago confirmed
Body	[name], Looks like the meeting in Chicago has been given the go-ahead by Lillehammer. The plan is to leave tomorrow morning at 8 AM to catch the 10 AM flight out of Roanoke. The meeting is taking place at the Sheraton Lodge. The number there is (242) 555-2344, in case someone needs to get in touch with us. We'll probably arrive back in Blacksburg sometime tomorrow night. I'll see you tomorrow at 8. Graham
Tasks Requested	none

B.2.2 DAY 3 EMAIL MESSAGES

TOTAL NUMBER SENT: 7

Number	
Sender	Jennifer Thompson,
Subject	Daily Reminder from the Experimenter
Body	Dear Participant, Just a reminder that this is Day 3 of the evaluation. You should receive 7 new emails including this one. Please let me know if you do not receive all 7 emails. And don't forget to complete your daily scenario. Just a thought: If any of you know anyone else who would be interested in participating in this study, please dictate me a message to let me know. Thanks again, Jennifer
Tasks Requested	**Dictate** message to experimenter (optional)

Number	
Sender	Travel Agent, travel_info_99@yahoo.com
Subject	Microbrew Festival in Oregon
Body	Dear Travel Information Participant, Plan a trip to Portland, Oregon for a Microbrew Festival No, there aren't 99 bottles of beer on the wall at the annual Oregon Brewers Festival in Portland, but almost. Portland's waterfront parks fill with lights and people for several annual summer festivals. This year, there will be 72. Yes, that's 72 different beers—one from each of 72 breweries around the United States. The annual gathering, now in its 12th year, is held on the last weekend of July in Tom McCall Waterfront Park, alongside Portland's Willamette River. This year the fest runs 23–25 July.
Tasks Requested	none

Number	
Sender	Brenda Donnaway (brenda_donnaway@hotmail.com)
Subject	(none)
Body	*Note: this message was initially sent to the user as a dictation created in a VEMS account. However, difficulties in sending this message resulted in having this message sent as text from a hotmail account.* Hi! This is Brenda from Human Resources. I understand that you have a meeting coming up regarding a contract of some sort. Could you dictate a reply to me and tell me with what company this contract is being arranged? Thank you.
Tasks Requested	**Dictate** reply with company name (Uplink Corporation)

Number	
Sender	Billy Bob, friend_billy@writeme.com
Subject	Big News!
Body	Hey [name]! It's me again. You won't believe what I just found out. Remember Patty Oakland? Well, her and Mike Oswald are getting married. Crazy, isn't it? Remember in high school? Those two used to hate each other. Anyways, thought you would be interested in hearing the news. Hope work is going well. Talk to you later, Billy Bob
Tasks Requested	none

Number	
Sender	Weather Info - weather_info@usa.com
Subject	Blacksburg Weather – [date]
Body	*Note: the contents of this email change on a daily basis to reflect the current weather in Blacksburg*
	LOCAL FORECAST FOR BLACKSBURG
	Tonight...cloudy with areas of drizzle. A chance of rain after midnight. Low 55 to 60. East wind 5 to 10 MPH. Chance of rain is 50 percent.
	Wednesday...cloudy and continued cool with a chance of light rain. high in the middle 60s. Southeast wind 5 to 10 MPH. Chance of rain is 50 percent.
	Wednesday night...cloudy with a chance of rain. Low around 60. chance of rain is 30 percent
	A LOOK AHEAD
	On Thursday the forecast calls for a high of 81 and a low of 59. Conditions for the day will be showers/sun.
	On Friday the forecast calls for a high of 83 and a low of 62. Conditions for the day will be partly cloudy.
Tasks Requested	none

Number	
Sender	Mark Lillehammer, your_supervisor@techpointer.com
Subject	Meeting Details
Body	[name],
	I would like to schedule a meeting with you and the other members of the R & D department to discuss our upcoming contract with UpLink Corporation. The meeting will take place in Conference Room A July 15 at 10:00. I will open the meeting with a brief overview of the contract details and would like you to give a 15-minute presentation covering the role that your department will play. I am considering inviting a representative from UpLink to provide input on our planned strategy.
	Please confirm that you will be able to attend this meeting.
	Thank you,
	Mark Lillehammer
Tasks Requested	none
Number	

Sender	Donna Hannun, donna_hannun@yahoo.com
Subject	Digital Camera
Body	[name], I was informed that you signed out one of the digital cameras from the Virginia Tech MultiMedia Lab two days ago. I will be needing this camera for a site visit on Monday. Could I arrange a time to get the camera from you? Thanks, Donna
Tasks Requested	none

B.2.3   DAY 4 EMAIL MESSAGES

TOTAL NUMBER SENT: 5 (but participant only received 4)

Number	1
Sender	Jennifer Thompson,
Subject	Daily Reminder from the Experimenter
Body	Dear Participant, Just to let you know, today is Day 4 of your evaluation. You will be receiving 4 emails. If you do not receive these emails, please let me know. If you are still receiving email from "Travel Agent", please exclude email from this sender before you hang up. If you are having any problems accessing either the VEMS web site or the Usability Evaluation Web Site, please let me know. Otherwise, keep up the great work. Jennifer
Tasks Requested	**Exclude** email from Travel Agent (travel_info_99@yahoo.com) if this was not done on Day 3

Number	2
Sender	Donna Hannun, donna_hannun@yahoo.com
Subject	Digital Camera Issue Resolved
Body	Hi [name]. I just found out from the Virginia Tech multimedia lab that there is a second digital camera available for signing out. So, no worries about giving back the camera. I also let them know that you would like to extend your loan until Monday, as you mentioned yesterday. Donna
Tasks Requested	none

Number	3
Sender	Madeline Finch - madeline_finch@techpointer.com
Subject	Reimbursement
Body	[name], I am putting together your reimbursement forms for the trip to Chicago. Can you send me a message regarding the name of the hotel that you stayed at during your conference? I believe Graham Roeburg provided you with this information two days ago. If you no longer have access to this information, please let me know and I will contact Graham. Thanks, Madeline
Tasks Requested	**Dictate** reply with name of hotel in Chicago (Sheridan Lodge).

Number	4
Sender	Travel Agent, travel_info_99@yahoo.com
Subject	A night in 1920s Paris
Body	*Note: this message should not be placed in the user's VEMS account, provided they properly excluded mail from this sender on Day 3* Dear Travel Information Recipient, This summer, relive the charm of 1920s Paris at a newly restored vintage hotel in the heart of the city's storied "Golden Triangle." To celebrate its grand reopening, the Millennium Commodore Paris is offering a special introductory rate during the peak European tourist season. The introductory rate includes a complimentary upgrade from a standard room to a larger "Club Room," a full breakfast, and a bottle of wine. Originally opened in 1927, the Millennium Commodore has been restored to period condition, including the addition of an authentic Parisian brasserie with a bar set up under the hotel's original glass dome. It is located within walking distance of major Paris landmarks such as the Louvre, the Arc de Triomphe, the Eiffel Tower, and the Galeries Lafayette.
Tasks Requested	none

Number	5
Sender	Weather Info - weather_info@usa.com
Subject	Blacksburg Weather – [date]
Body	*Note: the contents of this email change on a daily basis to reflect the current weather in Blacksburg* LOCAL FORECAST FOR BLACKSBURG Tonight...cloudy with areas of drizzle. A chance of rain after midnight. Low 55 to 60. East wind 5 to 10 MPH. Chance of rain is 50 percent. Wednesday...cloudy and continued cool with a chance of light rain. high in the middle 60s. Southeast wind 5 to 10 MPH. Chance of rain is 50 percent. Wednesday night...cloudy with a chance of rain. Low around 60. Chance of rain is 30 percent A LOOK AHEAD On Thursday the forecast calls for a high of 81 and a low of 59. Conditions for the day will be showers/sun. On Friday the forecast calls for a high of 83 and a low of 62. Conditions for the day will be partly cloudy.
Tasks Requested	none

B.2.4   DAY 5 EMAIL MESSAGES

TOTAL NUMBER SENT: 5

Number	1
Sender	Jesse (Your Best Friend) – yourbest_friend@netaddress.com
Subject	Hello there
Body	Hi! So you're finally done the voice email usability study! What are you going to do with all the money. I also heard that you get to keep that VEMS account of yours. Maybe I should sign up. Still up for a movie this weekend? Let me know. See you later, Jesse
Tasks Requested	none

Number	
Sender	Brenda Donnaway – brenda_donnaway@hotmail.com
Subject	Uplink representative visit
Body	Hello. I just wanted to thank you for the information regarding the Uplink contract. The representative from Uplink will be arriving in Blacksburg tomorrow at 9:00 AM. I have someone going to pick him up at the airport at that time. The contact's name is Henry Well. Could you forward this message to Graham Roeburg and Mark Lillehammer? I believe they would be interested in Mr. Well's arrival details. Sincerely, Brenda
Tasks Requested	**Forward** message to Graham Roeburg and Mark Lillehammer

Number	
Sender	Madeline Finch - madeline_finch@techpointer.com
Subject	Reimbursement Complete
Body	[name], Just wanted to let you know that I have completed the reimbursement forms and have submitted them to accounting. You should receive payment in a week's time. Madeline
Tasks Requested	**Dictate** reply with name of hotel in Chicago (Sheridan Lodge).

Number	
Sender	Jennifer Thompson,
Subject	Daily Reminder from the Experimenter
Body	Dear Participant, Just to let you know, today is Day 5, the last day of your evaluation. Congratulations! You will be receiving 5 emails including this one. If you do not receive these emails, please let me know. Also, please make sure to fill out the Final Questionnaire before you finish your scenario. I appreciate all the input you can provide. I would just like to say thank you for participating in the study. Your efforts will go towards improving the VEMS system. Jennifer
Tasks Requested	none

Number	
Sender	Weather Info - weather_info@usa.com
Subject	Blacksburg Weather – [date]
Body	*Note: the contents of this email change on a daily basis to reflect the current weather in Blacksburg* LOCAL FORECAST FOR BLACKSBURG Tonight...cloudy with areas of drizzle. A chance of rain after midnight. Low 55 to 60. East wind 5 to 10 MPH. Chance of rain is 50 percent. Wednesday...cloudy and continued cool with a chance of light rain. high in the middle 60s. Southeast wind 5 to 10 MPH. Chance of rain is 50 percent. Wednesday night...cloudy with a chance of rain. Low around 60. chance of rain is 30 percent A LOOK AHEAD On Thursday the forecast calls for a high of 81 and a low of 59. Conditions for the day will be showers/sun. On Friday the forecast calls for a high of 83 and a low of 62. Conditions for the day will be partly cloudy.
Tasks Requested	none

Number	
Sender	Donna Hannun, donna_hannun@yahoo.com
Subject	Digital Camera Issue Resolved
Body	Hi [name]. I just found out from the Virginia Tech multimedia lab that there is a second digital camera available for signing out. So, no worries about giving back the camera. I also let them know that you would like to extend your loan until Monday, as you mentioned yesterday. Donna
Tasks Requested	none

APPENDIX C.   IRB APPROVAL PACKAGE

C.1   IRB PROTOCOL
EXPERIMENTAL PROTOCOL
Investigating the Effectiveness of Applying the Critical Incident Technique to Remote Usability Evaluation

Justification of Project
The purpose of this research is to evaluate the feasibility of the critical incident technique in the context of remote usability evaluation. The need for feasible remote evaluation techniques is growing on account of the increased number of distributed and international user bases, making laboratory testing costly and impractical. Moreover, remote evaluation provides a means by which to capture usage patterns pertinent to actual work environments.

The success of the critical incident technique has already been demonstrated in the evaluation of computer and paper interfaces. This work will indicate whether or not this technique is equally successful when applied to a web and voice interface.   Previous studies have generated preliminary results supporting the use of user reporters of critical incidents. Experimental data to support or refute this claim will be gathered during this study by comparing the number and quality of critical incident reports generated by users to those created by usability experts. A portion of these user reporters will participate as remote subjects, a technique that has not yet been carried out beyond a simulated remote environment. Finally, the changes in critical incident identification and reporting that take place with repeated exposure to the system will also be assessed. It is expected that this data will support the development of guidelines regarding the minimum amount of time needed to obtain the maximum amount of critical incident data.

Procedure
The following is a description of the procedure to be followed in this research.

Participants
Thirty (15 males, 15 females) participants will be recruited to participate in this study. All participants will be Virginia Tech volunteer users of VEMS with no prior experience using the system. The participants will be required to have at least a year of experience using computers and the Web, but must be unfamiliar with usability testing methods. Once recruited, participants will be randomly assigned to one of the three conditions, ensuring that an equal number of males and females are assigned to each. Two additional participants will be recruited to act as trained observers of users interacting with VEMS. These participants must have conducted at least 5 human factors experiments and must have completed the equivalent of two semesters of human factors-related coursework. All participants will be compensated for their participation in the testing.

Equipment
Data will be collected in both a laboratory and remote setting. All users will be provided with a VESMS account. Laboratory participants will be provided with a desktop PC-compatible computer with Internet access. The computer will be located at a workstation also equipped with a telephone that has dial-out capabilities. Remote users will be required to gain access on their own to the above equipment (computer and telephone).

195

Experimental Tasks
This study is comprised of an introductory session and 4 subsequent mini sessions. On each day of the evaluation participants will be requested to complete a series of tasks consisting of basic emailing functions, such as replying, forwarding, etc. Instructions detailing these tasks will be obtained from the Usability Evaluation Web site, an on-line evaluation tool. Participants will be required to complete tasks assigned on a particular day on that day, although no time constraints will be imposed. While performing these tasks, some participants will be asked to evaluate VEMS by applying the critical incident technique.

During the introductory session, participants will be asked to complete a background questionnaire. The experimenter will then provide a short introduction to VEMS and to the Usability Evaluation Web Site. Some participants will be asked to go through the Critical Incident Technique training tool, which involves reading instructional material pertaining to the critical incident technique. The training tool is divided into two main modules, corresponding to the two major components of the technique. Each module will be covered in sequence, with the option of reviewing the material or taking a break at any time. The test session will finish with a questionnaire regarding the user's interaction with the training tool. The entire training session should last approximately 30 minutes.

During the remainder of the introductory session, participants will be asked to carry out their first set of tasks that involves setting up a VEMS account and learning how to use VEMS. While performing these tasks, some participants will be reminded to apply the critical incident technique to evaluate the system.

Laboratory-based participants will then be asked to return to the lab each day for the next four days, at times arranged with the experimenter. During these sessions, participants will be asked to perform a set of tasks outlined on the Usability Evaluation Web Site and, in some cases, to evaluate VEMS by applying the critical incident technique. Remote participants will not be requested to return to the lab. Instead, they will be asked to perform daily scenarios using any computer, provided it has an internet connection, access to Microsoft Internet Explorer (preferably version 5 or higher), and is located in close proximity to a phone. Mini sessions for all participants (laboratory and remote) are expected to last approximately 15-30 minutes.

Participants will be requested to only use VEMS in the laboratory and at times arranged with the experimenter. At the evaluation period, participants will be asked to complete a satisfaction questionnaire.

Informed Consent
Any research involving human subjects conducted at Virginia Tech must be reviewed and agreed to by the Institutional Review Board (IRB), a body dedicated to the review of such experiments performed at Virginia Tech. The IRB insures that federal guidelines (Code of Federal Regulations 46.102) are met where appropriate, that participants receive ethical treatment, and are not exposed to undue risk. A copy of the informed consent form for this experiment can be found in Section A.1.

Risks and Benefits
There will be minimal risk in this experiment. That is, there will be no increased chance of being exposed to a hazard for a participant in this experiment than he would normally be exposed to during everyday life while using a personal computer.

Confidentiality and Anonymity

The results of this study will be kept strictly confidential. At no time will the investigators release the results of the study to anyone other than individuals working on the project without the written consent of the participant. The information provided by each participant will be labeled using a participant numbering system instead of actual participant names. The numbering system will be used for all statistical analyses and written reports of the research. The only individuals with access to the data will be the students and their course professor/faculty sponsor.

Biographical Sketches of the Investigators

The following is a brief background for each of the investigators.

Jennifer Thompson, BASc

Ms. Thompson is a second year Masters student in the Human Factors Option of the Department of Industrial and Systems Engineering at Virginia Tech. She is currently a member of the Human-Computer Interaction (HCI) Laboratory and is serving as a Graduate Teaching Assistant for Introduction to Human Factors. She received her Bachelor of Applied Science degree in Systems Design Engineering from the University of Waterloo (Ontario, Canada) in May of 1998. Ms. Thompson is a member of both the Human Factors and Ergonomics Society and the Human Factors Association of Canada, and is soon to be inducted into the Alpha Pi Mu society. She is a recipient of the Natural Sciences and Engineering Research Council of Canada Post-Graduate Scholarship. Her research interests in human factors include HCI, usability, web accessibility, speech synthesis, and virtual environments.

Robert C. Williges

Robert C. Williges is Ralph H. Bogle Professor of Industrial and Systems Engineering at Virginia Polytechnic Institute and State University. He is also a Professor of Psychology and Computer Science. He is the Director of the Human-Computer Interaction Laboratory and Director of the Usability Methods Research Laboratory. He has over 25 years experience in managing and directing human factors engineering research dealing with topics including human-computer interaction, computer-based training procedures, and human factors research methodology. Before joining the faculty at Virginia Tech, Dr. Williges was on the faculty of the Department of Psychology at the University of Illinois at Urbana-Champaign and associate head of the Aviation Research Laboratory. He received the M.A. and Ph.D. degrees in engineering psychology from The Ohio State University and the A.B. degree in psychology from Wittenberg University. He is a fellow of the Human Factors and Ergonomics Society and the American Psychological Association and is a member of the Institute for Industrial Engineering and the Association for Computing Machinery. He has served as president of the Human Factors and Ergonomics Society and as president of Division 21 of the American Psychological Association. Dr. Williges is a past editor of *Human Factors*. He has over 275 scientific publications and has made over 190 technical presentations at national and international scientific meetings. His current research interests include developing methods for usability evaluation, improving hypermedia information presentation, designing computer-based assistive technology for computer users with disabilities, human factors engineering in manufacturing, and emerging human-computer interfaces involving visualization and computer-based conferencing.

A.1 INFORMED CONSENT FORMS

C.1.1 Informed Consent Form for Remote/Reporting Treatment Group

Informed Consent for Participants

Title of Project: Evaluation of a Voice Email Service

Investigators: Jennifer A. Thompson and Robert C. Williges

I.   Purpose of the Research

You are invited to participate in a study that involves experimentation for the purpose of evaluating and improving a voice email messaging service (VEMS). VEMS provides voice access to email, thereby allowing its users to access and listen to email from any phone, anywhere in the world. The service uses speaker-independent Automatic Speech Recognition (ASR) technology so that voice commands can be used to navigate through and interact with the system. Text to Speech (TTS) technology allows the system to read out email messages. To use the system, users must dial a toll-free number, and can then access their email by reading out an account number and password. To configure an account, users are required to access a web page account, again by means of an account number and password.

While interacting with VEMS, you will be asked to evaluate it using the critical incident technique. The critical incident technique is a method that allows for the collection of observations pertaining to events, called critical incidents, associated with features that lead to extremely good or extremely poor performance. In order to assist you in learning the critical incident technique, you will be trained using an on-line training tool. The objective of this is to improve your ability to successfully apply the critical incident technique to the evaluation of VEMS.

II.   Procedures

This study is comprised of an introductory session and 4 subsequent mini sessions. Specifically, each day of the evaluation you will be asked to complete a series of tasks consisting of basic emailing functions, such as replying, forwarding, etc. Instructions detailing these tasks can be obtained from the Usability Evaluation Web site, an on-line evaluation tool. You will be required to complete the tasks assigned to you on a particular day on that day, although you may take as much time to do so as you need. Simply perform the tasks to the best of your ability.   While you perform these tasks, you will be asked to evaluate VEMS by applying the critical incident technique. It should be noted that we are not evaluating you or your performance in any ways: you are helping us to evaluate our system. All information that you help us attain will remain anonymous.

During the introductory session, you will be asked to complete a background questionnaire. The experimenter will then provide you with a short introduction to VEMS and to the Usability Evaluation Web Site. You will then be asked to go through the Critical Incident Technique training tool, which involves reading instructional material pertaining to the critical incident technique. The training tool is divided into two main modules, corresponding to the two major components of the technique. You will go

through each module in sequence, with the option of reviewing the material or taking a break at any time. The test session will finish with a questionnaire regarding your interaction with the training tool. The entire training session should last approximately 30 minutes.

During the remainder of the introductory session, you will be asked to carry out your first set of tasks that will involve setting up your VEMS account and learning how to use the VEMS service. While performing these tasks, you will be obligated to apply the critical incident technique to evaluate the system.

You will then be asked to perform tasks outlined on the Usability Evaluation Web Site on each day for the next four days. You will not be required to return to the lab to perform these tasks. Rather, you can use any computer, provided it has an Internet connection, access to Microsoft Internet Explorer (preferably version 5 or higher), and is located in close proximity to a phone. An ideal place would be at home. If you do not have access to this equipment, please inform the experimenter now. The completion of these tasks should take approximately 20-30 minutes, and you are obligated to evaluate the VEMS system by applying the critical incident technique. It is advised that you complete your tasks between 8 AM and 10 PM on the day that they are assigned. After the evaluation period, you will be asked to complete a satisfaction questionnaire about your usage of the system.

III.   Risks

There are no known risks to the participants of this study. The tasks specified by the experimenter are not tiring, but during the training and test sessions, you are welcome to take rest breaks as needed.

IV.   Benefits of this Research

Your participation in this project will provide information that may be used to improve the usability of VEMS. No guarantee of benefits has been made to encourage you to participate. You may receive a synopsis summarizing this research when completed. Please leave a self-addressed envelope with the experimenter if you wish a copy of the results to be sent to you.

V.   Extent of Anonymity and Confidentiality

The results of this study will be kept strictly confidential. Your written consent is required for the researchers to release any data identified with you as an individual to anyone other than personnel working on the project. The information you provide will have your name removed and only a subject number will identify you during analyses and any written reports of the research.

The screen actions of your usage in the laboratory will be videotaped and your voice will be recorded. These tapes will be stored securely, viewed only by the experimenters (Jennifer Thompson, Dr. Williges, Dr. H. Rex Hartson, and Dr. Kleiner), and erased after 3 months. If the experimenters wish to use a portion of your videotape for any other purpose, they will get your written permission before using it. Your signature on this form does <u>not</u> give them permission to show your videotape to anyone else.

VI.   Compensation

You will receive $10 per hour for participation in the introductory study, which should last approximately 1 – 1.5 hours. You will receive a flat rate of $5 for each subsequent mini session. Payment will be made immediately after you have completed all evaluation sessions.

VII.   Freedom to Withdraw

You are free to withdraw from this study <u>at any time</u> for any reason. You will be compensated for your participation up to the point of withdrawal.

VIII.   Approval of Research

This research has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University, and by the Department of Industrial and Systems Engineering.

IX.   Subjects Responsibilities

I voluntarily agree to participate in this study. I have the following responsibilities:
1. I should not volunteer for participation, if I now know I will not be able to complete this experiment.
2. After completion of this study, I will not discuss my experiences with any other individual for a period of two months. This will ensure that everyone will begin the study with the same level of knowledge and expectations.
3. I will complete the tasks assigned to me each day of the evaluation period.

X.   Permission

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

Signature	Date

Should I have any questions about this research or its conduct, I may contact:

Jennifer A. Thompson	Investigator	552-3729
Robert C. Williges	Investigator	231-6270
H. T. Hurd	Chair, IRB Research Division	231-5281

C.1.2 Consent Form for Lab/Reporting Treatment Group

Informed Consent for Participants

<u>Title of Project</u>:   Voice Email Usability Study

<u>Investigators</u>: Jennifer A. Thompson and Robert C. Williges

I.   Purpose of the Research

You are invited to participate in a study that involves experimentation for the purpose of evaluating and improving a voice email messaging service (VEMS). VEMS provides voice access to email, thereby allowing its users to access and listen to email from any phone, anywhere in the world. The service uses speaker-independent Automatic Speech Recognition (ASR) technology so that voice commands can be used to navigate through and interact with the system. Text to Speech (TTS) technology allows the system to read out email messages. To use the system, users must dial a toll-free number, and can then access their email by reading out an account number and password. To configure an account, users are required to access a web page account, again by means of an account number and password.

While interacting with VEMS, you will be asked to evaluate it using the critical incident technique. The critical incident technique is a method that allows for the collection of observations pertaining to events, called critical incidents, associated with features that lead to extremely good or extremely poor performance. In order to assist you in learning the critical incident technique, you will be trained using an on-line training tool. The objective of this is to improve your ability to successfully apply the critical incident technique to the evaluation of VEMS.

II.   Procedures

This study is comprised of an introductory session and 4 subsequent mini sessions. Specifically, each day of the evaluation you will be asked to complete a series of tasks consisting of basic emailing functions, such as replying, forwarding, etc. Instructions detailing these tasks can be obtained from the Usability Evaluation Web site, an on-line evaluation tool. You will be required to complete the tasks assigned to you on a particular day on that day, although you may take as much time to do so as you need. Simply perform the tasks to the best of your ability.   While you perform these tasks, you will be asked to evaluate VEMS by applying the critical incident technique. It should be noted that we are not evaluating <u>you</u> or your performance in any ways: you are helping us to evaluate our system. All information that you help us attain will remain anonymous.

During the introductory session, you will be asked to complete a background questionnaire. The experimenter will then provide you with a short introduction to VEMS and to the Usability Evaluation Web Site. You will then be asked to go through the Critical Incident Technique training tool, which involves reading instructional material pertaining to the critical incident technique. The training tool is divided into two main modules, corresponding to the two major components of the technique. You will go through each module in sequence, with the option of reviewing the material or taking a

break at any time. The test session will finish with a questionnaire regarding your interaction with the training tool. The entire training session should last approximately 30 minutes.

During the remainder of the introductory session, you will be asked to carry out your first set of tasks that will involve setting up your VEMS account and learning how to use the VEMS service. While performing these tasks, you will be obligated to apply the critical incident technique to evaluate the system.

You will then be asked to return to the lab each day for the next four days, at times arranged with the experimenter. During these sessions, you will be asked to perform the tasks outlined on the Usability Evaluation Web Site and evaluate the VEMS system by applying the critical incident technique. These mini sessions should last approximately 15-30 minutes.

It will be asked that you only use VEMS in the laboratory and at times arranged with the experimenter. After the evaluation period, you will be asked to complete a satisfaction questionnaire about your usage of the system.

III.   Risks

There are no known risks to the participants of this study. The tasks specified by the experimenter are not tiring, but during the training and test sessions, you are welcome to take rest breaks as needed.

IV.   Benefits of this Research

Your participation in this project will provide information that may be used to improve the usability of VEMS. No guarantee of benefits has been made to encourage you to participate. You may receive a synopsis summarizing this research when completed. Please leave a self-addressed envelope with the experimenter if you wish a copy of the results to be sent to you.

V.   Extent of Anonymity and Confidentiality

The results of this study will be kept strictly confidential. Your written consent is required for the researchers to release any data identified with you as an individual to anyone other than personnel working on the project. The information you provide will have your name removed and only a subject number will identify you during analyses and any written reports of the research.

The screen actions of your usage will be videotaped and your voice will be recorded. These tapes will be stored securely, viewed only by the experimenters (Jennifer Thompson, Dr. Williges, Dr. H. Rex Hartson, and Dr. Kleiner), and erased after 3 months. If the experimenters wish to use a portion of your videotape for any other purpose, they will get your written permission before using it. Your signature on this form does <u>not</u> give them permission to show your videotape to anyone else.

VI.   Compensation

You will receive $10 per hour for participation in the introductory study, which should last approximately 1-1.5 hours. You will receive a flat rate of $5 for each subsequent mini session. Payment will be made immediately after you have completed all evaluation sessions.

VII.   Freedom to Withdraw

You are free to withdraw from this study <u>at any time</u> for any reason. You will be compensated for your participation up to the point of withdrawal.

VIII.   Approval of Research

This research has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University, and by the Department of Industrial and Systems Engineering.

IX.   Subjects Responsibilities

I voluntarily agree to participate in this study. I have the following responsibilities:
1   I should not volunteer for participation, if I now know I will not be able to complete this experiment.
2   After completion of this study, I will not discuss my experiences with any other individual for a period of two months. This will ensure that everyone will begin the study with the same level of knowledge and expectations.

X.   Permission

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

_____   _____

Signature   Date

Should I have any questions about this research or its conduct, I may contact:

Jennifer A. Thompson	Investigator	552-3729
Robert C. Williges	Investigator	231-6270
H. T. Hurd	Chair, IRB Research Division	231-5281

C.1.3   Consent Form for Lab/Non-Reporting Treatment Group

Informed Consent for Participants

Title of Project: Evaluation of a Voice Email Service

Investigators: Jennifer A. Thompson and Robert C. Williges

I.   Purpose of the Research

You are invited to participate in a study that involves experimentation for the purpose of evaluating and improving a voice email service (VEMS). VEMS provides voice access to email, thereby allowing its users to access and listen to email from any phone, anywhere in the world. The service uses speaker-independent Automatic Speech Recognition (ASR) technology so that voice commands can be used to navigate through and interact with the system. Text to Speech (TTS) technology allows the system to read out email messages. To use the system, users must dial a toll-free number, and can then access their email by reading out an account number and password. To configure an account, users are required to access a web page account, again by means of an account number and password.

II.   Procedures

This study is comprised of an introductory session and 4 subsequent mini sessions. During all sessions, you will be asked to interact with VEMS. Specifically, each day of the evaluation you will be asked to complete a series of tasks consisting of basic emailing functions, such as replying, forwarding, etc. Instructions detailing these tasks can be obtained from the Usability Evaluation Web site, an on-line evaluation tool. You will be required to complete the tasks assigned to you on a particular day on that day, although you may take as much time to do so as you need. Simply perform the tasks to the best of your ability. We are not evaluating you or your performance in any ways: you are helping us to evaluate our system. All information that you help us attain will remain anonymous.

During the introductory session, you will be asked to complete a background questionnaire. The experimenter will then provide you with a short introduction to VEMS and to the Usability Evaluation Web Site. You will then be asked to carry out your first set of tasks that will involve setting up your VEMS account and learning how to use VEMS. The entire introductory session should last approximately 30-45 minutes.
You will then be asked to return to the lab each day for the next four days, at times arranged with the experimenter. During these sessions, you will be asked to perform the tasks outlined on the Usability Evaluation Web Site. These mini sessions should last approximately 15-30 minutes.

It will be asked that you only use VEMS in the laboratory and at times arranged with the experimenter. After the evaluation period, you will be asked to complete a satisfaction questionnaire about your usage of the system.

III.   Risks

There are no known risks to the participants of this study. The tasks specified by the experimenter are not tiring, but during the training and test sessions, you are welcome to take rest breaks as needed.

IV.   Benefits of this Research

Your participation in this project will provide information that may be used to improve the usability of VEMS. No guarantee of benefits has been made to encourage you to participate. You may receive a synopsis summarizing this research when completed. Please leave a self-addressed envelope with the experimenter if you wish a copy of the results to be sent to you.

V.   Extent of Anonymity and Confidentiality

The results of this study will be kept strictly confidential. Your written consent is required for the researchers to release any data identified with you as an individual to anyone other than personnel working on the project. The information you provide will have your name removed and only a subject number will identify you during analyses and any written reports of the research.

The screen actions of your usage will be videotaped and your voice will be recorded. These tapes will be stored securely, viewed only by the experimenters (Jennifer Thompson, Dr. Williges, Dr. H. Rex Hartson, and Dr. Kleiner), and erased after 3 months. If the experimenters wish to use a portion of your videotape for any other purpose, they will get your written permission before using it. Your signature on this form does <u>not</u> give them permission to show your videotape to anyone else.

VI.   Compensation

You will receive $10 per hour for participation in the introductory study, which should last approximately 45 minutes. You will receive a flat rate of $5 for each subsequent mini session. Payment will be made immediately after you have completed all evaluation sessions.

VII.   Freedom to Withdraw

You are free to withdraw from this study <u>at any time</u> for any reason. You will be compensated for your participation up to the point of withdrawal.

VIII.   Approval of Research

This research has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University, and by the Department of Industrial and Systems Engineering.

IX.   Subjects Responsibilities

I voluntarily agree to participate in this study. I have the following responsibilities:

1. I should not volunteer for participation, if I now know I will not be able to complete this experiment.
2. After completion of this study, I will not discuss my experiences with any other individual for a period of two months. This will ensure that everyone will begin the study with the same level of knowledge and expectations.

X.   Permission

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

Signature	Date

Should I have any questions about this research or its conduct, I may contact:

Jennifer A. Thompson	Investigator	552-3729
Robert C. Williges	Investigator	231-6270
H. T. Hurd	Chair, IRB Research Division	231-5281

C.1.4 Consent Form for Usability Experts

Informed Consent for Participants

<u>Title of Project</u>:   Voice Email Usability Study

<u>Investigators</u>:   Jennifer A. Thompson and Robert C. Williges

I.   Purpose of the Research

You are invited to participate in a study that involves experimentation for the purpose of evaluating and improving a voice email messaging service (VEMS). VEMS provides voice access to email, thereby allowing its users to access and listen to email from any phone, anywhere in the world. The service uses speaker-independent Automatic Speech Recognition (ASR) technology so that voice commands can be used to navigate through and interact with the system. Text to Speech (TTS) technology allows the system to read out email messages. To use the system, users must dial a toll-free number, and can then access their email by reading out an account number and password. To configure an account, users are required to access a web page account, again by means of an account number and password.

While participating in this study, you will be asked to evaluate it using the critical incident technique. The critical incident technique is a method that allows for the collection of observations pertaining to events, called critical incidents, associated with features that lead to extremely good or extremely poor performance. In order to assist you in learning the critical incident technique, you will be trained using an on-line training tool. The objective of this is to improve your ability to successfully apply the critical incident technique to the evaluation of VEMS.

II.   Procedures

You will be asked to study a series of videotapes of screen usage footage from ten (10) user-subjects interacting with VEMS. These user-subjects were requested to attend test sessions on each of five consecutive days. During each of these test sessions, users were sent a set of 4-7 emails from various fictitious characters and asked to read these emails and use VEMS to perform a series of tasks consisting of basic emailing functions, such as replying, forwarding, etc. Instructions describing these tasks were accessed on-line according to which test session the user was attending (ex. Day 1, Day 2,…,Day 5). Your role in this study is that of an evaluator of user interaction with VEMS. The evaluation methodology you will be asked to use is the critical incident technique. That is, you will be asked to identify and report the critical incidents that occur for each participant during their interactions with the system.

It is important that you undergo critical incident training prior to your observation of user-subjects. This training is available on-line and provides reading instructional material pertaining to the critical incident technique. The training tool is divided into two main modules, corresponding to the two major components of the technique (i.e. identifying and reporting a critical incident). You will go through each module in sequence, with the option of reviewing the material or taking a break at any time. The test session will finish

with a questionnaire regarding your interaction with the training tool. The entire training session should last approximately 30 minutes.

A set of seven videotapes is available immediately for review. All remaining videos will be available no later than Saturday August 7th, 1999. These videotapes are labeled by participant number and are approximately 2 hours in length. You must review these videotapes in Whittemore 530 (the Human Computer Interaction Lab) out of the interest of the confidential nature of the data they contain.   A workstation is available to assist you in the review process. This workstation is comprised of a commercial model VCR, TV monitor and computer with Internet access. You will be given instruction on how to use this equipment. Review can take place at any time during the day, although prior arrangement should be made with the experimenter to guarantee access to the HCI lab (card accessible only for security reasons).

You can take as much time as needed to review the videotapes and generate the critical incident reports, as long as you finish by August 14, 1999. At the completion of your review, you will be asked to complete a questionnaire that relates to your experience analyzing the videotapes.

III.   Risks

There are no known risks to the participants of this study. The tasks specified by the experimenter are not tiring, but during the training and videotape review, you are welcome to take rest breaks as needed.

IV.   Benefits of this Research

Your participation in this project will provide information that may be used to improve the usability of a Voice Email Messaging System (VEMS). No guarantee of benefits has been made to encourage you to participate. You may receive a synopsis summarizing this research when completed. Please leave a self-addressed envelope with the experimenter if you wish a copy of the results to be sent to you.

V.   Extent of Anonymity and Confidentiality

The results of this study will be kept strictly confidential. Your written consent is required for the researchers to release any data identified with you as an individual to anyone other than personnel working on the project. The information you provide will have your name removed and only a subject number will identify you during analyses and any written reports of the research.

VI.   Compensation

You will receive $20 per videotape analyzed. There are 10 videotapes in total, each of which is approximately 2 hours in length. Payment will be made immediately after you have completed all evaluation sessions.

VII.   Freedom to Withdraw

You are free to withdraw from this study <u>at any time</u> for any reason. You will be compensated for your participation up to the point of withdrawal.

VIII.   Approval of Research

This research has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University, and by the Department of Industrial and Systems Engineering.

IX.   Subjects Responsibilities

I voluntarily agree to participate in this study. I have the following responsibilities:
1. I should not volunteer for participation, if I now know I will not be able to complete this experiment.
2. After completion of this study, I will not discuss my experiences with any other individual for a period of two months. This will ensure that everyone will begin the study with the same level of knowledge and expectations.

X.   Permission

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

_____   _____
Participant Signature   Date

_____   _____
Experimenter Signature   Date

Should I have any questions about this research or its conduct, I may contact:

Jennifer A. Thompson	Investigator	552-3729
Robert C. Williges	Investigator	231-6270
H. T. Hurd	Chair, IRB Research Division	231-5281

C.2 PRE-TEST QUESTIONNAIRES

C.2.1 Pre-Test Questionnaire for Remote/Reporting and Lab/Reporting Treatment Groups and Usability Experts

Thank you for participating in the evaluation of the On-line Critical Incident Reporting Tool. Please answer the following questions.

1. Age (years): _____   Sex: ❍ Male   ❍ Female

2. Academic Level: ❍ Freshman ❍ Sophomore   ❍ Junior   ❍ Senior   ❍ Masters   ❍ PhD

3. Major of Study: _____

4. How many university-level courses have you taken that have addressed human factors evaluation methodologies, usability evaluation, or human-computer interaction? *(please check one)*:
   ❍ none
   ❍ 1 - 2
   ❍ 3- 5
   ❍ 5+

5. How many usability evaluations have you conducted in the past (ex. as a participant, for a course project, for research purposes, in industry)? *(please check one)*:
   ❍ none
   ❍ 1 - 2
   ❍ 3- 5
   ❍ 6-10
   ❍ 10+

6. For how long have you been using computers *(please check one)*:
   ❍   less than 6 months
   ❍   between 6 months and a year
   ❍   1 - 3 years
   ❍   3 years or more

7. Please indicate which WWW browsers you have used *(please check all that apply)*:
   ❍ Netscape Navigator
   ❍ Microsoft Explorer
   ❍ AOL
   ❍ Other (*specify):* _____

   Of these browsers, which do you use most frequently: _____

8. Rate your level of expertise using **Microsoft Internet Explorer**:

❍ Very experienced   ❍ Moderate experience   ❍ Minimal experience   ❍No experience

9. Have you used VEMS before?   ❍ Yes   ❍ No

10. If yes, please indicate how long have you used VEMS?   _____ weeks

11. Please indicate any other email software programs or services that you have used *(please check all that apply)*:
    ❍ Eudora Light
    ❍ Hotmail
    ❍ Lotus cc:mail
    ❍ Microsoft Mail
    ❍ Netscape Mail
    ❍ Outlook Express
    ❍ Pine
    ❍ Other *(specify):* _____

12. In a typical day, how many times do you check your email *(please check one)*?
    ❍ 1–2   ❍ 3–5   ❍ 6–10   ❍10+

13. Rate your level of expertise using systems with automatic speech recognition?

❍ Very experienced   ❍ Moderate experience   ❍ Occasional usage   ❍ No experience

14. Rate your level of expertise using systems with synthesized voice?

❍ Very experienced   ❍ Moderate experience   ❍ Occasional usage   ❍ No experience

Rate the following statements by checking the most appropriate response.

15. I am very familiar with the critical incident technique.
    ❍   1 Strongly Agree
    ❍   2 Agree
    ❍   3 Neutral
    ❍   4 Disagree
    ❍   5 Strongly Disagree

16. I am very experienced at applying the critical incident technique.
    ❍   1 Strongly Agree
    ❍   2 Agree
    ❍   3 Neutral
    ❍   4 Disagree
    ❍   5 Strongly Disagree

17. Have you used ever used an on-line instructional tool (ex. training tool, on-line course)?
            ❍ Yes ❍ No

    If yes, please describe this training tool:

    _____

    _____

    _____

    _____

End of Questionnaire. Thank you.

C.2.2   Pre-Test Questionnaire for Lab/Non-Reporting Treatment Group

Thank you for participating in the evaluation of the On-line Critical Incident Reporting Tool. Please answer the following questions.

1. Age (years): _____   Sex: ❍ Male   ❍ Female

2. Academic Level: ❍ Freshman ❍ Sophomore   ❍ Junior   ❍ Senior   ❍ Masters   ❍ PhD

3. Major of Study: _____

4. How many university-level courses have you taken that have addressed human factors evaluation methodologies, usability evaluation, or human-computer interaction? *(please check one)*:
   - ❍   none
   - ❍   1 - 2
   - ❍   3- 5
   - ❍   5+

5. How many usability evaluations have you conducted in the past (ex. as a participant, for a course project, for research purposes, in industry)? *(please check one)*:
   - ❍   none
   - ❍   1 - 2
   - ❍   3- 5
   - ❍   6-10
   - ❍   10+

6. For how long have you been using computers *(please check one)*:
   - ❍   less than 6 months
   - ❍   between 6 months and a year
   - ❍   1 - 3 years
   - ❍   3 years or more

7. Please indicate which WWW browsers you have used *(please check all that apply)*:
   - ❍   Netscape Navigator
   - ❍   Microsoft Explorer
   - ❍   AOL
   - ❍   Other (*specify):* _____

   Of these browsers, which do you use most frequently: _____

8. Rate your level of expertise using **Microsoft Internet Explorer**:

   ❍ Very experienced   ❍ Moderate experience   ❍ Minimal experience   ❍No experience

9. Have you used VEMS before?   ❍ Yes   ❍ No

10. If yes, please indicate how long have you used VEMS*?*   _____ weeks

11. Please indicate any other email software programs or services that you have used *(please check all that apply)*:
   ❍ Eudora Light
   ❍ Hotmail
   ❍ Lotus cc:mail
   ❍ Microsoft Mail
   ❍ Netscape Mail
   ❍ Outlook Express
   ❍ Pine
   ❍ Other (*specify):* _____

12. In a typical day, how many times do you check your email *(please check one)*?
   ❍ 1–2   ❍ 3–5   ❍ 6–10   ❍10+

13. Rate your level of expertise using systems with automatic speech recognition?
   ❍ Very experienced   ❍ Moderate experience   ❍ Occasional usage   ❍ No experience

14. Rate your level of expertise using systems with synthesized voice?
   ❍ Very experienced   ❍ Moderate experience   ❍ Occasional usage   ❍ No experience

15. Have you used ever used an on-line instructional tool (ex. training tool, on-line course)?
   ❍ Yes ❍ No

   If yes, please describe this training tool:

   _____
   _____
   _____
   _____

End of Questionnaire. Thank you.

C.3 PARTICIPANT INSTRUCTIONS

C.3.1 Instructions for Remote/Reporting Treatment Group

INTRODUCTION

You are a business professional employed at a computer software company, called TechPointer. Often on the road and away from your computer, you are having trouble keeping up with your email. Your supervisor, Mark Lillehammer, recognized this as a problem and found out about a voice email messaging system (VEMS). This system provides telephone access to email, using a combination of speech recognition input and synthesized speech output.

VEMS Telephone System
You can access your email from any location by dialing a 1-800 number. The system will then request you to say your account number and password. Do so clearly but at a normal speaking pace (that is, do not overemphasize). Upon entering your account, you will be informed of the contents of your mail folder, including the number of new messages and of old messages (if any). You can then use voice commands, or their keypad equivalents, to navigate through these messages and carry out your emailing normal tasks (ex. read, reply, forward, etc.). You can dictate new messages to people listed in your personal directory, which get sent in the form of .wav files. You can also reply to an email by sending a pre-written reply or by dictating a message. While the system will inform you of an attachment and allow you to forward this attachment, you will not be able to read it.   The voice commands are printed on the wallet card for your reference purposes. Read through them now to acquaint yourself with the commands available.

VEMS Web Site
Configuration of your voice mailbox is achieved is done using the VEMS web site.   In order to gain access to this web page, you must enter in your account number and password (same as those used to access your voice mailbox). The web page allows you to update account information, configure email-handling parameters, and set up your personal profile (including your personal directory, reply list, and priority and exclude lists). This web page should be configured prior to using the voice system.

Usability Evaluation Web Site
Your role is to evaluate VEMS using the critical incident technique. You will be trained shortly on how to apply this technique.   In order to facilitate the evaluation, a web site has been created called the Usability Evaluation Web Site (http://hci.ise.vt.edu/~jethomp6). This is the same web site from which you will access the Critical Incident Training Tool. It is best viewed using Internet Microsoft Explorer (version 5 or higher).

Instructions to Start
1. Log in to the Usability Evaluation Web Site (http://hci.ise.vt.edu/~jethomp6). Your password is: _____. Remember this password. You will need it each time you log in to the evaluation web site.
2. You should be taken to the Usability Evaluation home page. Review the contents of the screen.

3. To evaluate VEMS, you will be using the critical incident technique. It is important that you be trained to use this evaluation technique. Press the START TRAINING link to start the Critical Incident Training Session.
4. When you have completed the training session, return to the Usability Evaluation home page. The instructor will then provide you with further instruction

INFORMATION REGARDING THE EVALUATION

Daily Scenario
The evaluation will take place over the remaining four days. Each day you will be asked to perform a specific set of tasks, which you should perform to the best of your abilities. While performing these tasks, you will be required to report all critical incidents that occur. Pressing the DAILY SCENARIOS button (shown below) will give you access to instructions regarding the tasks to perform. It is located in the web site side bar and from the main Usability Evaluation web page.

YOU MUST COMPLETE THE TASKS BETWEEN **8:00 AM AND 10:00 PM** ON THE DAY THAT THEY ARE ASSIGNED. You are to provide confirmation that you have completed the tasks assigned by submitting a confirmation message to the experimenter.

Critical Incident Reporting
Critical incidents can be reported by pressing the REPORT INCIDENT button, shown below. This button is located in the web site side bar and from the main Usability Evaluation web page. Remember you should report both positive and negative critical incidents.

It is important that you report your critical incidents while you are interacting with the VEMS telephone service. It is advised that you keep the Usability Evaluation web site active while you access your email. You can then fill in the Critical Incident Report Form while concurrently interacting with the telephone system. The URL for the Usability Evaluation Web Site and instructions for reporting critical incidents are provided on the back of the voice command card.

Remote Access
You will not be required to return to the lab to perform these tasks. Rather, you can use any computer, provided it has an Internet connection, access to Microsoft Internet Explorer (preferably version 5 or higher), and is located in close proximity to a phone. An ideal place would be at home. If you do not have access to this equipment, please inform the experimenter now.

Email Account
When you register as a VEMS user, you will receive a new email account. You can access this account by dialing a 1-800 number. You are to use **only** this email account for the duration of the study. You will receive email each day and will be required to respond to this email as directed in the Scenario descriptions. Do not forward personal email to this account or ask people to respond to you at this account.

At the conclusion of the study, you will be able to keep your VEMS account and will be free to configure it as you wish.

Evaluation End

On the fifth day of the evaluation, you will be asked to complete a questionnaire to obtain your feedback regarding VEMS and the critical incident technique. You will also be compensated at this time for your participation. The experimenter will arrange a time for you to return to the lab to receive your payment. If you have any questions regarding the instructions provided, please ask the experimenter now.

Instructions to Start

You are now ready to begin your evaluation of VEMS.
1. Go to the Usability Evaluation web site (using Internet Microsoft Explorer).
2. Press the DAILY SCENARIO button to receive your first set of instructions.

C.3.2 Instructions for Lab/Reporting Treatment Group

INTRODUCTION

You are a business professional employed at a computer software company called TechPointer. Often on the road and away from your computer, you are having trouble keeping up with your email. Your supervisor, Mark Lillehammer, recognized this as a problem and found out about a voice email messaging system (VEMS). This system provides telephone access to email, using a combination of speech recognition input and synthesized speech output.

VEMS Telephone System

You can access your email from any location by dialing a 1-800 number. The system will then request you to say your account number and password. Do so clearly but at a normal speaking pace (that is, do not overemphasize). Upon entering your account, you will be informed of the contents of your mail folder, including the number of new messages and of old messages (if any). You can then use voice commands, or their keypad equivalents, to navigate through these messages and carry out your emailing normal tasks (ex. read, reply, forward, etc.). You can dictate new messages to people listed in your personal directory, which get sent in the form of .wav files. You can also reply to an email by sending a pre-written reply or by dictating a message. While the system will inform you of an attachment and allow you to forward this attachment, you will not be able to read it. More information regarding the features of VEMS is provided on the attached sheets. The voice commands are printed on the wallet card for your reference purposes. Read through them now to acquaint yourself with the commands available.

VEMS Web Site

Configuration of your voice mailbox is achieved using the VEMS web site. In order to gain access to this site, you must enter in your account number and password (same as those used to access your voice mailbox). The web page allows you to update account information, configure email-handling parameters, and set up your personal profile (including your personal directory, reply list, and priority and exclude lists). This web page should be configured prior to using the voice system.

Usability Evaluation Web Site

Your role is to evaluate VEMS using the critical incident technique. You will be trained shortly on how to apply this technique.   In order to facilitate the evaluation, a web site has been created called the Usability Evaluation Web Site (http://hci.ise.vt.edu/~jethomp6). This is the same web site from which you will access the Critical Incident Training Tool. It is best viewed using Internet Microsoft Explorer (version 5 or higher).

Instructions to Start
1. Log in to the Usability Evaluation Web Site (http://hci.ise.vt.edu/~jethomp6). Your password is:_____. Remember this password. You will need it each time you log in to the evaluation web site.
2. You should be taken to the Usability Evaluation home page. Review the contents of the screen.
3. To evaluate VEMS, you will be using the critical incident technique. It is important that you be trained to use this evaluation technique. Press the START TRAINING link to start the Critical Incident Training Session.
4. When you have completed the training session, return to the Usability Evaluation home page. The instructor will then provide you with further instruction

INFORMATION REGARDING THE EVALUATION

Daily Scenarios
The evaluation will take place over the remaining four days. Each day you will be asked to perform a specific set of tasks, which you should perform to the best of your abilities. While performing these tasks, you will be required to report all critical incidents that occur. Pressing the DAILY SCENARIOS button (shown below) will give you access to instructions regarding the tasks to perform. It is located in the web site side bar and from the main Usability Evaluation web page.

You must complete the tasks to the best of your ability by the end of the test session. You are to provide confirmation that you have completed the tasks by submitting a confirmation message to the experimenter.

Critical Incident Reporting
Critical incidents can be reported by pressing the REPORT INCIDENT button, shown below. This button is located in the web site side bar and from the main Usability Evaluation web page. Remember you should report both positive and negative critical incidents.

The URL for the Usability Evaluation Web Site and instructions for reporting critical incidents are provided on the back of the voice command card.

Email Account
When you register as a VEMS user, you will receive a new email account. You can access this account on the VEMS system by dialing a 1-800 number. You are to use only this email account for the duration of the study. You will receive email each day and will be required to respond to this email as directed in the Scenario descriptions.   Do not forward personal email to this account

or ask people to respond to you at this account. At the conclusion of the study, you will be able to keep your VEMS account and will be free to configure it as you wish.

Evaluation End
On the fifth day of the evaluation, you will be asked to complete a questionnaire to obtain your feedback regarding VEMS and the critical incident technique. You will also be compensated at this time for your participation. If you have any questions regarding the instructions provided, please ask the experimenter now.

Instructions to Start
You are now ready to begin your evaluation of VEMS. Press the DAILY SCENARIO button to receive your first set of instructions.

C.3.3   Instructions for Lab/Non-Reporting Treatment Group

INTRODUCTION

You are a business professional employed at a computer software company called Techpointer. Often on the road and away from your computer, you are having trouble keeping up with your email. Your supervisor, Mark Lillehammer, recognized this as a problem and found out about a Voice Email Messaging System (VEMS). This system provides telephone access to email, using a combination of speech recognition input and synthesized speech output.

VEMS Telephone System
You can access your email from any location by dialing a 1-800 number. The system will then request you to say your account number and password. Do so clearly but at a normal speaking pace (that is, do not overemphasize). Upon entering your account, you will be informed of the contents of your mail folder, including the number of new messages and of old messages (if any). You can then use voice commands, or their keypad equivalents, to navigate through these messages and carry out your emailing normal tasks (ex. read, reply, forward, etc.). You can dictate new messages to people listed in your personal directory, which get sent in the form of .wav files. You can also reply to an email by sending a pre-written reply or by dictating a message. While the system will inform you of an attachment and allow you to forward this attachment, you will not be able to read it.   More information regarding the features of VEMS is provided on the attached sheets.The voice commands are printed on the wallet card for your reference purposes. Read through them now to acquaint yourself with the commands available.

VEMS Web Site
Configuration of your voice mailbox is achieved using the VEMS web site. In order to gain access to this web site, you must enter in your account number and password (same as those used to access your voice mailbox). The web page allows you to update account information, configure email-handling parameters, and set up your personal profile (including your personal directory, reply list, and priority and exclude lists). This web page should be configured prior to using the voice system.

Usability Evaluation Web Site
You will be required to use the Usability Evaluation Web Site in order to gain access to certain information and on-line tools. This web site is best viewed using Internet Microsoft Explorer (version 5 or higher).

Instructions to Start
1. Log in to the Usability Evaluation Web Site (**http://hci.ise.vt.edu/~jethomp6**). Your password is: _____. Remember this password. You will need it each time you log in to the evaluation web site.
2. You should be taken to the Usability Evaluation home page. Review the contents of the screen.

INFORMATION REGARDING THE EVALUATION

Daily Scenario
The evaluation will take place over the remaining four days. Each day you will be asked to perform a specific set of tasks, which you should perform to the best of your abilities. Pressing the DAILY SCENARIOS button (shown below) will give you access to instructions regarding the tasks to perform. It is located in the web site side bar and from the main Usability Evaluation web page.

You must complete the tasks to the best of your ability by the end of the test session. You are to provide confirmation that you have completed the tasks by submitting a confirmation message to the experimenter. The URL for the Usability Evaluation Web Site and instructions for reporting critical incidents are provided on the back of the voice command card.

Email Account
When you register as a VEMS user, you will receive a new email account. You can access this account on VEMS by dialing a 1-800 number. You are to use **only** this email account for the duration of the study. You will receive email each day and will be required to respond to this email as directed in the Scenario descriptions.   Do not forward personal email to this account or ask people to respond to you at this account.   At the conclusion of the study, you will be able to keep your VEMS account and will be free to configure it as you wish.

Evaluation End
On the fifth day of the evaluation, you will be asked to complete a questionnaire to obtain your feedback regarding VEMS. You will also be compensated at this time for your participation. If you have any questions regarding the instructions provided, please ask the experimenter now.

Start Evaluation
You are now ready to begin using VEMS.

1. Go to the Usability Evaluation web site (using Internet Microsoft Explorer).
2. Press the DAILY SCENARIO button to receive your first set of instructions.

C.3.4   Instructions for Usability Experts

INTRODUCTION

You have been asked to observe user-subjects interacting with a voice email messaging service (VEMS) and to report the critical incidents that occur during this interaction. Provided below is a overview of VEMS service as well as of tools available to assist you in the observation and reporting process.

VEMS Telephone System

[THIS SECTION HAS BEEN DELETED FOR REASONS OF CONFIDENTIALITY]

VEMS Web Site

[THIS SECTION HAS BEEN DELETED FOR REASONS OF CONFIDENTIALITY]

Usability Evaluation Web Site
Your role is to evaluate user interaction with the VEMS telephone system and web site using the critical incident technique. You will be trained shortly on how to apply this technique.   In order to facilitate the evaluation, a web site has been created called the Usability Evaluation Web Site (http://hci.ise.vt.edu/~jethomp6/expert). It is best viewed using Microsoft Internet Explorer (version 5 or higher).

Features of this web site include a training tool for the critical incident, access to critical incident report forms (for reporting critical incidents), and day-specific instructions for tasks that users are asked to complete.   A picture of this web site is provided below.
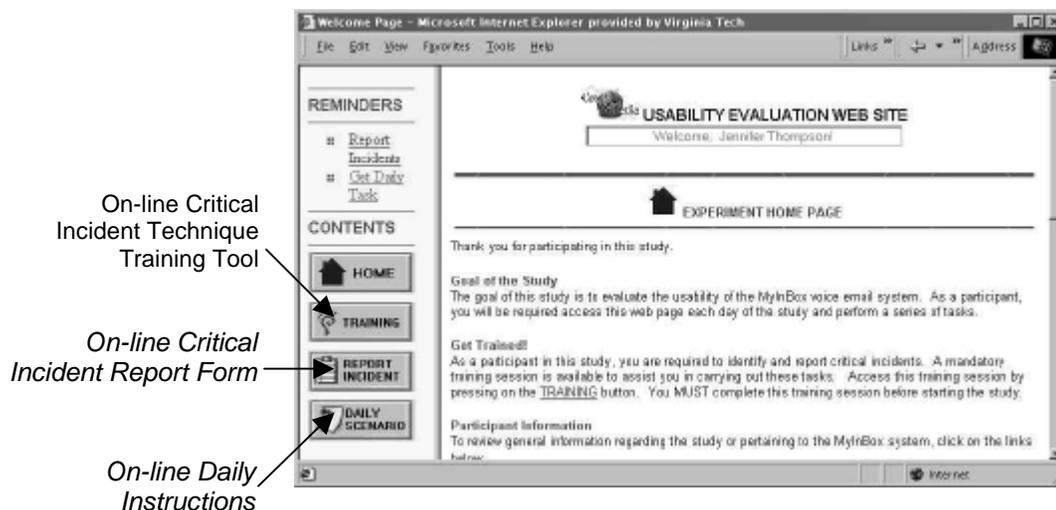

Figure C-1: Usability Evaluation Web Site Home Page

The Usability Evaluation Web Site is password accessible. Your password is **--------.**

220

On-line Critical Incident Training Tool
The training tool was created to provide you with a review of the critical incident technique. It is broken down according to the two major components of the technique: identifying and reporting a critical incident.

On-line Critical Incident Report Forms
Another feature of the Usability Evaluation Web Site is an on-line critical incident report form. There are two versions of the form, corresponding to the two types of critical incidents (positive or negative). Each form requires that you describe the task being formed, rate the frequency at which this task is performed, describe the critical incident, and finally rate the effect of the critical incident on task performance, satisfaction, and error severity. You are required to fill out a report whenever you observe a critical incident to occur. The contents of these reports are submitted via email to the experimenter who can then compile the reports into a summary of negative and positive aspects of the system.

On-line Daily Scenarios
The final component of the web site is access to daily scenarios corresponding to each of the five days a participant was involved in the study. A daily scenario is comprised of tasks, which are in turn broken down into step-by-step instructions. These instructions reflect actions representative of a typical VEMS user, including such actions as reading, replying, and forwarding email messages, excluding or prioritizing senders, and dictating new messages. You are given access to these instructions, as well as to the emails sent to the participants, to facilitate your interpretation of user screen usage footage.

Instructions to Start
1. Log in to the Usability Evaluation Web Site (http://hci.ise.vt.edu/~jethomp6/expert). Your password is: **expert**. Remember this password. You will need it each time you log in to the evaluation web site.
2. You should be taken to the Usability Evaluation home page. Review the contents of the screen.
3. To evaluate VEMS, you will be using the critical incident technique. It is important that you be trained to use this evaluation technique. Press the START TRAINING link to start the Critical Incident Training Session.
4. When you have completed the training session, return to the Usability Evaluation home page. The instructor will then provide you with further instruction.

VIDEOTAPE REVIEW INSTRUCTIONS

Procedure

Whenever you come into the lab to review a videotape, it is advised that you follow the procedure outlined below:

1. Turn on all equipment.
2. Log into the Usability Evaluation web site (http://hci.ise.vt.edu/~jethomp6/expert) and enter your name and password.
3. Determine what "Day" of the evaluation you will be observing.
4. Press the DAILY SCENARIO button (shown below) to gain access to instructions for tasks assigned to the user-subjects for that particular day.

5. Open the Email Binder to the appropriate day to gain access to the contents of all emails sent to the user-subjects on that particular day.
6. Observe the screen usage footage. If a critical incident is identified, report it on-line by pressing the REPORT INCIDENT button, shown below. This button is located in the web site side bar and from the main Usability Evaluation web page. Remember you should report both positive and negative critical incidents.

TIPS:

1. It is important that you report critical incidents while you are observing user interaction with the VEMS interfaces. It is advised that you keep the Usability Evaluation web site active at all times. You can then fill in the Critical Incident Report Form while concurrently interacting with the telephone system.

2. Review the videotapes in order (starting at 5) and keep track of all videotapes by filling out the Expert Review Chart provided.

222

C.5   POST-TEST QUESTIONNAIRES

C.5.1   POST-TEST QUESTIONNAIRE FOR NON-REPORTER USER-SUBJECTS

Please indicate how strongly you disagree or agree to the statements using the number scale provided. Try to respond to all the items. For any item that is not applicable, please use **NA**.

VEMS Web Interface

Respond to the following questions as they pertain to the VEMS web page.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
1. It was simple to use the VEMS web site.	O	O	O	O	O	O
2. I am able to complete my tasks quickly using the VEMS web site.	O	O	O	O	O	O
3. I am able to efficiently complete my tasks using the VEMS web site.	O	O	O	O	O	O
4. It was easy to learn to use the web page	O	O	O	O	O	O
5. The web page gives error messages that clearly tell me how to fix problems.	O	O	O	O	O	O
6. Whenever I make a mistake using the web page, I recover easily and quickly.	O	O	O	O	O	O
7. The information (ex. on-line help, on-screen messages) provided by the VEMS web pages is clear.	O	O	O	O	O	O
8. It is easy to find the information I need.	O	O	O	O	O	O
9. The information provided by the web site is easy to understand.	O	O	O	O	O	O
10. The organization of information on the web pages is clear.	O	O	O	O	O	O
11. I like using the VEMS web site.	O	O	O	O	O	O
12. The VEMS web site has all the functions and capabilities that I expect it to have.	O	O	O	O	O	O

VEMS Voice Interface

Respond to the following questions as they pertain to the VEMS voice system.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
1. It was simple to use the VEMS voice system.	○	○	○	○	○	○
2. I am able to complete my emailing tasks quickly using the VEMS voice system.	○	○	○	○	○	○
3. I am able to efficiently complete my emailing tasks using the VEMS voice system.	○	○	○	○	○	○
4. It was easy to learn to use the voice system.	○	○	○	○	○	○
5. The voice system gives error messages that clearly tell me how to fix problems.	○	○	○	○	○	○
6. Whenever I make a mistake using the voice system, I recover easily and quickly.	○	○	○	○	○	○
7. The information (ex. on-line help, on-screen messages) provided by the VEMS voice system is clear.	○	○	○	○	○	○
8. It is easy to find the information I need.	○	○	○	○	○	○
9. The information provided by the voice system is easy to understand.	○	○	○	○	○	○
10. The organization of information on the web pages is clear.	○	○	○	○	○	○
11. I like using the VEMS voice system.	○	○	○	○	○	○
12. The VEMS voice system has all the functions and capabilities that I expect it to have.	○	○	○	○	○	○

Comments: _____

_____

_____

List the most **negative** aspect(s) of VEMS:

1. _____
2 _____
3. _____

List the most **positive** aspect(s) of VEMS:
1. _____
2 _____
3. _____

End of questionnaire. Thank you for your help.

C.5.2 POST-TEST QUESTIONNAIRE FOR REPORTER USER-SUBJECTS

Please indicate how strongly you disagree or agree to the statements using the number scale provided. Try to respond to all the items. For any item that is not applicable, please use **NA**.

VEMS Web Interface

Respond to the following questions as they pertain to the VEMS web page.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
1. It was simple to use the VEMS web site.	○	○	○	○	○	○
2. I am able to complete my tasks quickly using the VEMS web site.	○	○	○	○	○	○
3. I am able to efficiently complete my tasks using the VEMS web site.	○	○	○	○	○	○
4. It was easy to learn to use the web site.	○	○	○	○	○	○
5. The web page gives error messages that clearly tell me how to fix problems.	○	○	○	○	○	○
6. Whenever I make a mistake using the web site, I recover easily and quickly.	○	○	○	○	○	○
7. The information (ex. on-line help, on-screen messages) provided by the VEMS web pages is clear.	○	○	○	○	○	○
8. It is easy to find the information I need.	○	○	○	○	○	○
9. The information provided by the web site is easy to understand.	○	○	○	○	○	○
10. The organization of information on the web pages is clear.	○	○	○	○	○	○
11. I like using the VEMS web site.	○	○	○	○	○	○
12. The VEMS web site has all the functions and capabilities that I expect it to have.	○	○	○	○	○	○

VEMS Voice Interface

Respond to the following questions as they pertain to the VEMS voice interface.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
1. It was simple to use the VEMS voice system.	O	O	O	O	O	O
2. I am able to complete my emailing tasks quickly using the VEMS voice system.	O	O	O	O	O	O
3. I am able to efficiently complete my emailing tasks using the VEMS voice system.	O	O	O	O	O	O
4. It was easy to learn to use the voice system.	O	O	O	O	O	O
5. The voice system gives error messages that clearly tell me how to fix problems.	O	O	O	O	O	O
6. Whenever I make a mistake using the voice system, I recover easily and quickly.	O	O	O	O	O	O
7. The information (ex. on-line help, on-screen messages) provided by the VEMS voice system is clear.	O	O	O	O	O	O
8. It is easy to find the information I need.	O	O	O	O	O	O
9. The information provided by the voice system is easy to understand.	O	O	O	O	O	O
10. The organization of information on the web pages is clear.	O	O	O	O	O	O
11. I like using the VEMS voice system.	O	O	O	O	O	O
12. The VEMS voice system has all the functions and capabilities that I expect it to have.	O	O	O	O	O	O

Critical Incident Reporting

The following questions pertain to your role as a critical incident identifier and reporter.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
I like the idea of remote-reporting critical incident information to developers.	O	O	O	O	O	O
Reporting critical incidents anonymously is important for me.	O	O	O	O	O	O
The critical incident training provided enough information.	O	O	O	O	O	O
The training was easy to follow.	O	O	O	O	O	O
The training helped me to learn to recognize critical incidents.	O	O	O	O	O	O
The training helped me learn to report critical incidents using the on-line report form.	O	O	O	O	O	O
I found the exercise problems at the end of each training module useful.	O	O	O	O	O	O
I felt better prepared to identify the critical incidents after going through the training tool.	O	O	O	O	O	O
I felt better prepared to report critical incidents after going through the training tool.	O	O	O	O	O	O
It was easy to access the critical incident report form.	O	O	O	O	O	O
It was easy to report critical incidents using the report form.	O	O	O	O	O	O
The questions on the report form were easy to understand.	O	O	O	O	O	O
The questions on the report form covered sufficient detail concerning the critical incident.	O	O	O	O	O	O
I was motivated to report negative critical incidents.	O	O	O	O	O	O
I was motivated to report positive critical incidents.	O	O	O	O	O	O
It was easy to rate the impact of the critical incidents on task performance.	O	O	O	O	O	O
It was easy to rate the impact of the critical incident on satisfaction.	O	O	O	O	O	O
It was easy to rate the severity of errors	O	O	O	O	O	O

I prefer to report critical incidents:
❒   immediately when I encounter them ❒   after completing my task

*Comments:*_____

_____

_____

_____

End of questionnaire. Thank you for your help.

APPENDIX D.   RESULTS

D.1 TRAINING QUESTIONNAIRE DATA

A questionnaire was administered to all participants who underwent the critical incident training to assess their subjective evaluation of the tool. The responses allocated to each question are presented in the table below, along with the average response and standard deviation.

Table D-1. Training Questionnaire Responses

Question	Participant																				Avg.	St. Dev
It was simple to use the training tool.	4	4	4	5	5	4	4	4	4	5	4	5	4	5	4	4	4	5	4	5	4.35	0.48
The training was easy to follow.	4	4	3	5	5	5	5	4	4	5	4	5	4	5	4	2	4	5	4	5	4.30	0.80
The training tool provided sufficient information.	5	5	5	4	5	5	5	4	4	5	4	5	4	5	5	5	4	5	4	5	4.65	0.48
The information was easy to understand.	4	4	5	5	5	4	5	4	4	5	4	5	4	5	5	5	4	5	4	5	4.55	0.49
The organization of information was clear.	4	5	5	5	5	3	5	4	4	5	4	5	3	5	3	5	4	5	4	5	4.40	0.75
I liked interacting with the training tool.	4	5	5	3	5	3	4	3	4	4	4	4	4	3	4	2	3	4	3	4	3.75	0.78
The training helped me learn to **identify positive** critical incidents.	5	5	5	5	5	4	4	4	4	5	4	5	5	4	5	5	4	5	4	4	4.55	0.50
The training helped me learn to **identify negative** critical incidents.	5	5	5	5	5	4	5	4	4	5	4	5	5	4	5	5	4	5	4	4	4.60	0.49
The training helped me learn to **report positive** critical incidents.	5	5	5	4	5	5	4	4	4	5	4	5	5	4	5	5	4	5	4	4	4.55	0.50
The training helped me learn to **report negative** critical incidents.	5	5	5	4	5	5	5	4	4	5	4	5	5	4	5	5	4	5	4	4	4.60	0.49
The material covered by the training tool was sufficient.	5	5	4	5	5	4	5	3	4	5	4	5	4	4	4	5	4	5	4	4	4.40	0.58
I feel better prepared to **identify** critical incidents after going through the training tool.	5	4	5	5	5	4	5	4	4	4	4	5	5	5	5	4	4	5	4	5	4.55	0.50
I feel better prepared to **report** critical incidents after going through the training tool.	5	5	4	4	5	5	5	4	4	4	4	5	5	4	4	5	4	5	4	5	4.50	0.50

D.2 USABILITY PROBLEM AND SUCCESS DESCRIPTIONS

Usability problem and success descriptions were developed based on the contents of each critical incident report form submitted. Table D-2 through Table D-5 present a complete list of usability descriptions, categorized according to critical incident type (problem or success) and interface type (voice interface or web interface).

Table D-2. Usability Problems For Voice Interface - Ordered by Frequency of Occurrence

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Error Severity	Average Overall Criticality
1	Account number/password had to be repeated before successful log-in.	11	4.14	2.77	3.5	3.05	1.1
2	Command confusion - send reply vs. dictate reply	9	3.89	2.56	4.5	3.33	0.33
3	Cannot remove all messages from a particular sender	9	3.56	2.89	3.25	3.11	1.2
4	VR Error - Next message vs. read message	7	5	2.57	3.83	3.71	1.00
5	VR Error - Read message vs. remove message	7	4.29	3.14	3.50	3.29	1.00
6	System does not allow user to list messages from a non-address book member	7	3.71	3.14	4	3.86	1.75
7	Cannot prioritize a particular sender from anywhere in the inbox.	6	3.75	3.83	N/A	3.83	1.833
8	Voice speaks too quickly	6	3	2.67	3.67	3.17	N/A
9	Difficult to keep track of the location of a specific message.	5	4.6	3.8	N/A	3.8	2
10	Voice is difficult to understand	5	3.6	3.2	3.4	3.2	N/A
11	System had difficulty recognizing Next Message command.	4	4.5	2.75	4.5	4	1
12	Cannot find a stop speaking command	4	4.5	2	3.25	3.33	N/A
13	VR Error - Forward message vs. remove message	4	4.25	3.75	3.67	3.75	2
14	Speak slower command has no effect (most likely because used while voice was speaking)	4	3.75	2.5	3.75	3.25	N/A
15	Cannot skip to a particular message	3	4.67	2.67	3	3.33	2
16	Cannot stop voice by voice command (ex. stop reading message command)	3	4.33	2.67	2.5	2.67	0
17	System had difficulty recognizing "Yes" to add a recipient in Forward this Message mode.	3	4.33	2.67	3.5	3.67	2
18	Command to send a dictation once dictation complete is not obvious/clear - user unsure of how to proceed.	3	4.17	4.5	N/A	4.33	2.33
19	How to access list of replies unclear/not obvious once in send reply mode.	3	4.17	3.5	N/A	3.5	0.5

230

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Error Severity	Average Overall Criticality
20	System had difficulty recognizing Dictate New Message command.	3	4	3	5	4	2
21	VR Error - Next message vs. remove message	3	3.83	3	N/A	3.17	1.167
22	Cannot exclude a particular sender from anywhere in the inbox.	3	3.67	3	N/A	3.33	1.33
23	System had difficulty recognizing "Read message" command.	2	5	3	N/A	3	1
24	System had difficulty recognizing Goodbye command.	2	5	2.5	N/A	2.5	0.5
25	VR Error - Read message vs. get 1st new message	2	5	2.5	5	4	1
26	VR Error - Read message vs. next message	2	5	2.5	5	4.00	1
27	VR Error - Previous message vs. remove message	2	4.5	2.5	4	4	N/A
28	Old messages difficult to locate because not where expected (stacked above new messages).	2	4	5	4.5	3.5	N/A
29	VR Error - Dictate new message vs. remove message	2	4	3.5	3	3.5	2
30	Confusion regarding location of old vs. new messages	2	4	3.5	N/A	3.5	1.5
31	System had difficulty recognizing Send Reply command.	2	4	3	N/A	2.5	0
32	Voice reads email addresses as numbers vs. digits - difficult to follow/understand	2	3	2	2.5	2	N/A
33	Commands in general are not understood by system	1	5	4	N/A	5	3
34	System had difficulty recognizing "No" command during recipient name confirmation in Dictate New Message mode.	1	5	4	N/A	4	2
35	VR Error - Dictate new message vs. get first new message	1	5	4	N/A	4	2
36	VR Error - Next message vs. Dictate new message	1	5	4	N/A	4	2
37	Navigation commands limit user to moving down/up one message at a time.	1	5	4	N/A	4	2
38	System does not start listening until a couple seconds after finished speaking	1	5	3	3	3	N/A
39	VR Error - Send reply vs. goodbye	1	5	3	4	2	N/A
40	Too many steps required of user in dictate new message - interferes with ability to remember specific details to be dictated.	1	5	3	N/A	3	-1
41	Cannot specify recipient (in dictate new message) by email address	1	5	3	N/A	3	1
42	Cannot use speak slower/speak faster while voice is talking	1	5	2	4	4	N/A
43	System had difficulty recognizing Send Message command in Dictate Message mode (after dictation complete).	1	5	2	4	4	N/A

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Error Severity	Average Overall Criticality
44	VR Error - Next message vs. get last new message	1	5	2	5	5	N/A
45	System awaits command rather than providing command options (ex. press 1 to do this, press 2 to do this) (user-initiated versus system-initiated dialog) and therefore, do not know what commands are available to use at different points in the interaction.	1	5	2	4	4	N/A
46	No distinction is made between email body and signature - confusing, difficult to interpret	1	5	2	3	3	N/A
47	System had difficulty recognizing Forward this Message command.	1	4.5	4	N/A	4	1.5
48	System had difficulty recognizing Remove Messages More than 1 day old command.	1	4	5	5	5	N/A
49	Help provided in exclude sender mode does not guide the user in what to do (just says Exclude sender means you will exclude this sender from your mail list")	1	4	5	4	4	N/A
50	System had difficulty recognizing "Yes" command to confirm reply in Send Reply mode.	1	4	4	N/A	4	1
51	System had difficulty recognizing the name of recipient in Dictate New Message mode.	1	4	4	N/A	N/A	2
52	VR Error - Remove messages more than 1 day old vs. remove message more than 5/9 days old	1	4	4	N/A	4	2
53	Not obvious what happens to messages once removed	1	4	4	N/A	4	2
54	Command needed to access list of replies in send reply mode not recalled.	1	4	4	N/A	4	2
55	Cannot add a reply while on the phone	1	4	4	N/A	5	2
56	System had difficulty recognizing "Yes" to add a recipient in Dictate New Message mode.	1	4	3	N/A	3	0
57	System had difficulty recognizing "No" to recipient name confirmation in Forward this Message mode.	1	4	3	N/A	4	2
58	System had difficulty recognizing the name of an existing reply.	1	4	3	N/A	3	1
59	VR Error - Forward message vs. read message	1	4	3	4	4	N/A
60	VR Error - Prioritize message vs. read message	1	4	3	N/A	3	1
61	VR Error - Read message vs. restore message	1	4	3	3	3	N/A
62	VR Error - Send reply vs. remove	1	4	3		3	1
63	VR Error - Yes to confirm recipient of a forwarded message vs. help	1	4	3	3	3	N/A

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Error Severity	Average Overall Criticality
64	Too many options given after dictation complete - forget what command to use to send dictation	1	4	3	4	3	N/A
65	Wait time before help prompt in send reply mode (at send reply point) too long	1	4	3	N/A	2	-1
66	Cannot specify recipient by first name	1	4	3	N/A	3	1
67	Having to go through new messages to get to old messages is time-consuming.	1	4	2	5	5	N/A
68	The expression "unread messages" is confusing.	1	4	2	2	2	N/A
69	System does not accept "Send dictated message" as a substitute for "send dictation" in Dictate New Message.	1	4	2	4	4	N/A
70	System had difficulty recognizing "Yes" to add a recipient in Prioritize Sender mode.	1	3.5	3	N/A	3.5	1.5
71	VR Error - Command (not specified) misinterpreted as "remove message"	1	3	5	5	5	N/A
72	Ability to restore message not obvious	1	3	5	5	5	N/A
73	Remove messages more than X days old not apparent/obvious to user.	1	3	4	N/A	4	1
74	System does not listen while voice is speaking	1	3	3	4	3	N/A
75	VR Error - Restore message vs. exclude sender	1	3	3		3	1
76	Wait time before help prompt in prioritize sender mode (at confirm recipient point) too long	1	3	3	N/A	3	1
77	System does not accept "Reply" as a substitute for "Send Reply"	1	3	3	2	2	N/A
78	Cannot review last few seconds of voice text	1	3	3	3	2	N/A
79	VR Error - Commands in general misunderstood by system	1	3	2	3	2	N/A
80	Cannot forward a message to a non-address book member.	1	3	2	4	4	N/A
81	Which passcode to enter in log-in is not specified (artifact of experimental protocol).	1	2	3	N/A	3	-1
82	System does not allow "Repeat message" command to be used instead of "read message".	1	2	3	5	2	N/A

Table D-3. Usability Problems for Web Interface - Ordered by Frequency of Occurrence

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Error Severity	Average Overall Criticality
1	Personal Profile Home link difficult to locate (due to size and location of link - expected to be one of the main icons).	8	2.94	2.94	3.33	3.13	1.2
2	Requirement that reply name must contain at least 2 words is not stated/obvious.	5	3.2	2.8	2	2.9	1.125
3	Hint Option misinterpreted as referring to passcode (hint option must match with passcode).	3	2	2.67			-0.33
4	ExpressLane is difficult to locate	3	1.67	3	3	2.33	N/A
5	Account Number log-in box is mistaken for location in which corporate account number is entered.	2	3.25	3	3	3	1
6	One-word reply names are not accepted by the system.	2	3	2	2.5	1	
7	Protocol for deleting list entries not intuitive.	2	2.75	3	N/A	3	1
8	Polling set-up feature is not easily found.	2	2	4	N/A	3	1
9	Hint Option misinterpreted as referring to passcode (hint option must match with passcode).	2	1.5	3	2.5	2.5	N/A
10	Not obvious that hint & verification text boxes must be filled out (no asterisks denoting required info).	2	1.5	2.5	2	1.5	N/A
11	Error message given when Hint Option not selected on Account Registration Form does not specify what info is wrong/missing.	2	1	3	2.5	2.5	N/A
12	Expected ExpressLane to be configured as a hyperlink embedded within the PPH.	2	1	3	3	2	N/A
13	Log-in information on home page is lost if user accesses Hint page and then returns.	1	5	3	3	3	N/A
14	List entry windows should have a toolbar available (for copy and paste functionality).	1	5	2	5	1	N/A
15	MyAddressBook icon not apparent/easily located.	1	4	4	N/A	4	1
16	MyPriorityList icon not apparent/easily located.	1	4	3	N/A	3	0
17	The number and types of fields that must be specified while creating a new MyPriorityList entry are not clear/obvious.	1	4	3	N/A	3	1
18	The purpose of the Reply Name field is not clear.	1	3.5	3	N/A	1.5	1
19	Concept and implication of "Prioritize Sender" function unclear.	1	3	5	4	4	N/A
20	Changes made to an entry in one list are not reflected in other lists containing that same entry (lists are	1	3	5	2	2	N/A

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Error Severity	Average Overall Criticality
	not linked).						
21	Unable to prioritize someone directly from address book link (lists should be linked)	1	3	4	4	1	N/A
22	If user hits "Back" after submitting Account Registration form or tries to re-enter data after a successful submission and then presses SUBMIT again, a General Database error occurs (which prevents user from submitting).	1	3	3	N/A	3	1
23	Instruction that only one priority field can be specified is not obvious.	1	3	3	3	2	N/A
24	Pressing cancel in list entry (ex. PriorityList entry) window was expected to delete entry.	1	3	3	N/A	2	0
25	Priority list entry in ExpressLane box did not get removed when deleted (by pressing trashcan icon and pressing Yes to confirm).	1	3	3	2	3	N/A
26	"Priority notification is not setup" message is misinterpreted to mean that a priority list entry was not successfully added.	1	3	2	4	4	N/A
27	Layout of MyPriorityList entry form does not make it obvious that only one field should be specified (because similar to other entry forms which do accept multiple fields).	1	3	2	3	2	N/A
28	If user returns to Account Registration form after successful submission, not all information is re-loaded.	1	2	3	N/A	3	1
29	Difference between AutoForwarding and Account Polling not clear.	1	2	3	N/A	3	1
30	Expected a very evident "Sign Up Now" button or link on VEMS home page (similar to Hotmail)	1	2	3	3	3	N/A
31	Lack of specificity with respect to which email address should be specified (ex. VEMS or personal email address) in Email text box on Account Registration form.	1	2	3	3	3	N/A
32	Pressing the return button after completing an input field entry in registration submits the form (instead of moving cursor to next field).	1	2	3	2	2	N/A
33	Description of ExpressLane feature on Personal Profile Home does not give page on which ExpressLane is located.	1	1	3	3	2	N/A
34	Case sensitivity of account log-in password not made clear.	1	1	3	3	2	N/A

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Error Severity	Average Overall Criticality
35	Asterisks appear when typing in Hint text box on Account Registration form - cannot verify that input is correct.	1	1	2	2	2	N/A

Table D-4. Usability Successes for Voice Interface - Ordered by Frequency of Occurrence

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Overall Criticality
1	Remove messages more than X days old permits quick and easy bulk message removal.	8	3.88	3.5	3.67	1.2
2	Forward message function is quickly and easily executed.	7	4.716	3.29	5	1.17
3	Extra assistance provided during first-time use of a command is helpful and reduces confusion.	7	3.14	3.71	3.86	N/A
4	Send Reply function is quickly and easily executed.	5	4.2	3.6	4	1
5	Exclude Sender function is quickly and easily executed.	5	3.8	3	N/A	1
6	Speak slower command helps to better understand synthesized voice.	3	3	3	N/A	0.33
7	System correctly recognized commands.	2	5	4.5	5	2
8	Entered account number/passcode successfully.	2	4.5	3.5	4	1
9	Restore Message command useful.	2	4	4	4	N/A
10	Prioritize sender function can be is quickly and easily executed.	2	3.5	3	N/A	1
11	Exclude sender function allows junk mail to be quickly eliminated.	2	2.5	4	4	N/A
12	Keypad option useful.	1	5	4	4	N/A
13	Example account number helpful in log-in.	1	5	4	4	N/A
14	Easier to log-in with practice.	1	5	4	4	N/A
15	Dictate Reply function is easily executed.	1	5	3	N/A	1
16	Read message command easy to use.	1	5	3	N/A	0
17	Remove message easy to use.	1	5	3	N/A	1
18	Voice has good inflection.	1	5	2	5	N/A
19	Voice is understandable.	1	5	2	4	N/A
20	Review Dictation command in Dictate New Message mode useful.	1	4	4	5	N/A
21	List messages from X command reduces the amount of search time.	1	4	4	N/A	2
22	System easier to use with practice.	1	4	4	5	N/A
23	Get Last New Message command useful.	1	4	3	3	N/A
24	Ability to hear list of replies helpful.	1	4	3	N/A	1
25	Comeback command useful.	1	3	4	4	N/A
26	Speak faster function easy to use.	1	3	3	N/A	1
27	Voice system is easy to learn.	1	3	3	4	N/A
28	Ability to use help to repeat command options available after dictation complete is helpful.	1	3	3	3	N/A
29	Speak louder command allows user to better hear/understand synthesized voice.	1	2	2	N/A	-1

Table D-5. Usability Successes for Web Interface - Ordered by Frequency of Occurrence

No.	Usability Problem Description	Instances	Average Task Frequency	Average Impact on Task Performance	Average Impact on Satisfaction	Average Overall Criticality
1	Size and position of main icons (ex. for MyReplyList) & location of managing dialog box make them easily recognized and accessible.	12	3.50	3.17	N/A	1.00
2	Size and placement of Add icon (in management dialog box) makes it easily recognized and accessed.	11	4.00	3.00	N/A	1.09
3	ExpressLane allows for quick and easy addition of new entries (provides shortcut).	7	4.00	3.29	3.67	1.00
4	Protocol for deleting list entries intuitive/quick and easy.	6	3.83	3.50	3.60	1.00
5	Protocol for interacting and navigating through web site intuitive/quick and easy.	4	4.00	4.50	5.00	N/A
6	Good feedback for entering an invalid email address in new AddressBook list entry.	3	3.67	3.00	3.00	0.50
7	Protocol for editing list entries intuitive/quick and easy.	3	3.33	3.00	N/A	1.00
8	Setting up account for polling is easily and quickly done.	3	2.67	3.33	3.67	N/A
9	Confirmation of new entry good - indicates add entry action successful + allows user to check for errors.	2	4	3.5	4	2.00
10	Good feedback for specifying too many prioritize fields in new Priority List entry.	2	3.50	3.50	N/A	1.00
11	Protocol for adding new replies intuitive/quick and easy.	2	3.00	3.50	5.00	1.00
12	Size and placement of trash icon(s) make them easily recognized and accessed.	2	3.00	3.00	N/A	1.00
13	Email account polling set-up feature is easily found.	2	2.50	3.00	N/A	1.00
14	Separate window for list entry creation/management good.	1	4.00	3.00	3.00	N/A
15	MyAddressBook icon location makes it easily recognized and draws user's attention.	1	4.00	3.00	N/A	1.00
16	Protocol for adding a new priority list member is intuitive/quick and easy.	1	3.00	4.00	4.00	N/A
17	Ability to add new entries consecutively (I.e. from confirmation screen) good (provides a short-cut).	1	3.00	3.00	4.00	N/A
18	AccountDetails icon size and location make it easily located.	1	3.00	3.00	N/A	1.00
19	The specificity of the error message for invalid account number given after account registration info submission allows for quick recovery.	1	2.00	3.00	N/A	1.00
20	Location of ExpressLane Add icons increases their visibility.	1	2.00	3.00	N/A	0.00
21	Error on registration page causes user to be returned to that page and does not result in deletion of information.	1	1.00	3.00	2.00	

D.3 CHI-SQUARE ANALYSES OF UPI CLASSIFICATION DATA

The Usability Problem Inspector (UPI) was used to classify critical incidents according to the Interaction Activity during which they occurred. There are four Interaction Activities: Planning, Physical Actions, Outcome, and Assessment. Of interest was to determine whether or not the distribution of critical incidents across these Interaction Activity categories was equal or not. A Chi-Square Goodness of Fit test was conducted for each treatment group and interface based on the following test hypotheses and decision rules:

H_o: $O = E$ (observed frequency is equal to the expected frequency)
H_a: $O \neq E$ (observed frequency is not equal to the expected frequency)
Decision Rule: I reject H_o if $\chi^2_{observed} > \chi^2_{tabled}$

A level of significance of 0.05 was used for all tests due to the exploratory nature of usability evaluation research and expected frequencies were based on the expectation that each Interaction Activity category would contain 25 % of the incidents reported by a particular treatment group.

D.3.1 Remote/Reporting Treatment Group Results

Chi-Square Goodness of Fit test results for the web and voice interfaces are presented in Table D-6 and Table D-7 respectively.

Table D-6. Chi-Square Goodness of Fit Test Results for the Remote/Reporting Treatment Group - Web Interface

Interaction Activity	Observed Frequency	Expected Frequency	Chi-Squared
Planning	16	5.25	22.0119
Physical	2	5.25	2.011905
Outcome	1	5.25	3.440476
Assessment	2	5.25	2.011905
Total	21		**29.47619**

Observed Chi-Square	**29.476**
df	3
Tabled Chi-Square	7.814725

238

Table D-7. Chi-Square Goodness of Fit Test Results for the Remote/Reporting Treatment Group - Voice Interface

Interaction Activity	Observed Frequency	Expected Frequency	Chi-Squared
Planning	22	11.25	10.27222
Physical	6	11.25	2.45
Outcome	11	11.25	0.005556
Assessment	6	11.25	2.45
Total	45		**15.17778**

Observed Chi-Square	**15.178**
df	3
Tabled Chi-Square	7.815

Since, $\chi^2_{observed} > \chi^2_{tabled}$, reject the null hypothesis for both the web and voice interfaces. Therefore, it can be concluded that the Remote/Reporting treatment group reported an unequal distribution of critical incidents across the interaction activity types. Inspection reveals that a larger number of planning incidents were reported than expected by the remote users when interacting with the web and voice interface.

D.3.2   Lab/Reporting Treatment Group Results

Chi-Square Goodness of Fit test results for the web and voice interfaces are presented in Table D-8 and Table D-9 respectively.

Table D-8. Chi-Square Goodness of Fit Test Results for the Lab/Reporting Treatment Group - Web Interface

Interaction Activity	Observed Frequency	Expected Frequency	Chi-Squared
Planning	27	8.5	40.26471
Physical	4	8.5	2.382353
Outcome	2	8.5	4.970588
Assessment	1	8.5	6.617647
Total	34		**54.23529**

Observed Chi-Square	**54.235**
df	3
TabledChi-Square	7.814725

Table D-9. Chi-Square Goodness of Fit Test Results for the Lab/Reporting Treatment Group - Voice Interface

Interaction Activity	Observed Frequency	Expected Frequency	Chi-Squared
Planning	34	18.75	12.403
Physical	7	18.75	7.363
Outcome	23	18.75	0.963
Assessment	11	18.75	3.203
Total	75		**23.933**

Observed Chi-Square	**23.933**
df	3
Tabled Chi-Square	7.815

Since, $\chi^2_{observed} > \chi^2_{tabled}$, reject the null hypothesis for both the web and voice interfaces. Therefore, it can be concluded that the number of critical incidents reported by Lab/Reporting treatment were unequally distributed across the interaction activity types. Inspection reveals that a larger number of planning incidents were reported than expected by the laboratory-based users when interacting with the web and voice interfaces. In addition, this treatment group reported far fewer Physical Actions incidents for the voice interface than expected.

D.3.3   Expert 1 Results

Chi-Square Goodness of Fit test results for the web and voice interfaces using Expert 1 data are presented in Table D-10 and Table D-11 respectively.

Table D-10. Chi-Square Goodness of Fit Test Results for the Expert 1 - Web Interface

Interaction Activity	Observed Frequency	Expected Frequency	Chi-Squared
Planning	54	17	80.52941
Physical	7	17	5.882353
Outcome	2	17	13.23529
Assessment	5	17	8.470588
Total	68		**108.1176**

Observed Chi-Square	**108.118**
df	3
Tabled Chi-Square	7.814725

Table D-11. Chi-Square Goodness of Fit Test Results for Expert 1 - Voice Interface

Interaction Activity	Observed Frequency	Expected Frequency	Chi-Squared
Planning	45	28.25	9.931416
Physical	25	28.25	0.373894
Outcome	40	28.25	4.887168
Assessment	3	28.25	22.56858
Total	113		**37.76106**

Observed Chi-Square	**37.761**	
df	3	
Tabled Chi-Square	7.815	

Since, $\chi^2_{observed} > \chi^2_{tabled}$, reject the null hypothesis for both the web and voice interfaces. Therefore, it can be concluded that the number of critical incidents reported by Expert 1 were unequally distributed across the interaction activity types. Inspection reveals that a larger number of planning incidents were reported than expected by Expert 1 when observing interactions with the web interface. In addition, Expert 1 was shown to report far fewer Assessment incidents for the voice interface than expected.

D.3.4 Expert 2 Results

Chi-Square Goodness of Fit test results for the web and voice interfaces using Expert 2 data are presented in Table D-12 and Table D-13 respectively.

Table D-12. Chi-Square Goodness of Fit Test Results for the Expert 2 - Web Interface

Interaction Activity	Observed Frequency	Expected Frequency	Chi-Squared
Planning	7	2.25	10.02778
Physical	0	2.25	2.25
Outcome	0	2.25	2.25
Assessment	2	2.25	0.027778
Total	9		**14.55556**

Observed Chi-Square	**14.556**	
df	3	
Tabled Chi-Square	7.814725	

Table D-13. Chi-Square Goodness of Fit Test Results for Expert 2 - Voice Interface

Interaction Activity	Observed Frequency	Expected Frequency	Chi-Squared
Planning	9	6.25	1.21
Physical	0	6.25	6.25
Outcome	16	6.25	15.21
Assessment	0	6.25	6.25
Total	25		**28.92**

Observed Chi-Square	**28.920**
df	3
Tabled Chi-Square	7.815

Since, $\chi^2_{observed} > \chi^2_{tabled}$, reject the null hypothesis for both the web and voice interfaces. Therefore, it can be concluded that the number of critical incidents reported by Expert 2 were unequally distributed across the interaction activity types. Inspection reveals that a larger number of planning incidents were reported than expected by the Expert 2 when observing interactions with the web interface. In addition, Expert 2 reported far fewer Assessment incidents for the voice interface than expected.

D.4   NEWMAN-KEULS ANALYSES FOR TOTAL NUMBER OF CRITICAL INCIDENTS REPORTED

D.4.1   Main Effect of Treatment

Expert 1 Data

A main effect of Treatment group was found to be significant at $p < 0.0001$. The results of a Newman-Keuls test to isolate significant differences are presented in Table D-14.

Table D-14. Newman-Keuls Results for Main Effect of Treatment on the Total Number of Critical Incidents (Expert 1 Data)

	Increasing Order (n=100; MSE = 1.0522)				
	1	2	3		
Level*	1	2	3		
Mean	0.62	1.01	1.68	r	CDn-k**
Increasing Order 1		**0.39**	**1.06**	3	0.3621
2			**0.67**	2	0.2995
3					

*Level 1 = Remote/Reporting; Level 2= Lab/Reporting; Level 3=Expert 1 (for Lab/Non-Reporting participants)
**The studentized range statistic used to calculate CD_{n-k} was based on df=24

Expert 2 Data

A main effect of Treatment was found to be significant at $p < 0.0001$. Newman-Keuls post hoc analysis was performed to determine which treatment groups generated significantly higher numbers of critical incident reports and the results are shown in Table D-15.

Table D-15. Newman-Keuls Results for Main Effect of Treatment on the Total Number of Critical Incidents (Expert 2 Data)

		Increasing Order (n=100; MSE = 0.9356)				
		1	2	3		
	Level*	3	1	2		
	Mean	0.34	0.62	1.01	r	CDn-k**
Increasing	3	---	**0.28**	**0.67**	3	0.3414
Order	1		---	**0.39**	2	0.2824
	2			---		

*Level 1 = Remote/Reporting; Level 2= Lab/Reporting; Level 3=Expert 1 (for Lab/Non-Reporting participants)
**The studentized range statistic used to calculate CD_{n-k} was based on df=24

D.4.2 Main Effect of Day

Expert 1 Data
A main effect of day was found to be significant at $p < 0.0001$. A Newman-Keuls post hoc test was conducted to determine upon which days a significantly higher numbers of critical incident were reported. Results are shown in Table D-16.

Table D-16. Newman-Keuls Results for Main Effect of Day on the Total Number of Critical Incidents (Expert 1 Data)

		Increasing Order (n=60; MSE=0.7983)					
		1	2	3	4		
	Level	5	3	4	1,2		
	Mean	0.567	0.917	1.267	1.383	r	**CDn-k****
Increasing	a5	---	**0.35**	**0.7**	**0.816**	4	0.4315
Order	a3		---	**0.35**	**0.466**	3	0.3922
	a4			---	0.116	2	0.3265
	a1, a2				---		

**The studentized range statistic used to calculated CD_{n-k} was based on df=60.

Expert 2 Data
The main effect of Day was found to be significant at $p < 0.0001$. A Newman-Keuls post hoc analysis was performed to determine upon which days a significantly higher numbers of critical incident were reported and the results are shown in Table D-17.

Table D-17. Newman-Keuls Results for Main Effect of Day on the Total Number of Critical Incidents (Expert 2 Data)

		Increasing Order (n=60; MSE = 0.4948)						
		1	2	3	4	5		
	Level	5	3	4	1	2		
	Mean	0.367	0.483	0.583	0.900	0.950	r	CDn-k
Increasing	a5		0.117	0.217	**0.533**	**0.583**	5	0.3614
Order	a3			0.100	**0.417**	**0.467**	4	0.3396
	a4				**0.317**	**0.367**	3	0.3088
	a1					0.050	2	0.2570
	a2							

**The studentized range statistic used to calculated CD_{n-k} was based on df=60.

D.4.3 Interaction of Interface and Day

Expert 1 Data

A significant interaction of Interface and Day was found at $p < 0.0001$. A Newman-Keuls post hoc analysis of the comparisons of the interaction between Interface and Day was conducted, the results of which are presented in Table D-18.

Table D-18. Newman-Keuls Results for Interaction of Day and Interface on the Total Number of Critical Incidents (Expert 1 Data)

	Increasing Order (n=30; MSE = 0.7881; I = Interface; D = Day)										
	1	2	3	4	5	6	7	8	9		
Treatment	I1D3	I1D5	I1D2	I2D5	Group E	I2D4	I2D3	I1D1	I2D2		
Mean	0.100	0.300	0.633	0.833	0.867	1.667	1.733	1.900	2.133	r	CDn-k**
I1D3	---	0.200	0.533	**0.733**	**0.767**	**1.567**	**1.633**	**1.800**	**2.033**	9	0.7375
I1D5		---	0.333	0.533	0.567	**1.367**	**1.433**	**1.600**	**1.833**	8	0.7196
I1D2			---	0.200	0.233	**1.033**	**1.100**	**1.267**	**1.500**	7	0.6986
I2D5				---	0.033	**0.833**	**0.900**	**1.067**	**1.300**	6	0.6743
Group E					---	**0.800**	**0.867**	**1.033**	**1.267**	5	0.6451
I2D4						---	0.067	0.233	0.467	4	0.6062
I2D3							---	0.167	0.400	3	0.5511
I1D1								---	0.233	2	0.4587
I2D2									---		

Where:
     Group E = {I1D4, I2D1}
**The studentized range statistic used to calculated CD_{n-k} was based on df=60.

Expert 2 Data

Newman-Keuls test results for the interaction of Interface and Day are presented in Table D-19.

Table D-19. Newman-Keuls Results for Interaction of Day and Interface on the Total Number of Critical Incidents (Expert 2 Data)

	Increasing Order (n=30; MSE = 0.5537; I = Interface; D = Day)											
	1	2	3	4	5	6	7	8	9	10		
Treatment	I1D3	I1D5	I1D2	I1D4	I2D1	I2D5	I2D4	I2D3	I1D1	I2D2		
Mean	0.067	0.1	0.3	0.367	0.6	0.633	0.8	0.9	1.2	1.6	r	CDn-k
I1D3	---	0.033	0.233	0.300	0.533	**0.567**	**0.733**	**0.833**	**1.133**	**1.533**	10	0.6317
I1D5			0.2	0.267	0.5	0.533	**0.7**	**0.8**	**1.1**	**1.5**	9	0.6181
I1D2				0.067	0.300	0.333	0.500	**0.600**	**0.900**	**1.300**	8	0.6032
I1D4					0.233	0.267	0.433	0.533	**0.833**	**1.233**	7	0.5855
I2D1						0.033	0.200	0.300	**0.600**	**1.000**	6	0.5652
I2D5							0.167	0.267	**0.567**	**0.967**	5	0.5407
I2D4								0.1	0.4	**0.8**	4	0.5081
I2D3									0.3	**0.7**	3	0.4619
I1D1										**0.4**	2	0.3845
I2D2										---		

D.5 NEWMAN-KEULS ANALYSES FOR NUMBER OF NEGATIVE CRITICAL INCIDENTS REPORTED

D.5.1 Main Effect of Treatment

Expert 1 Data

A main effect of Treatment group was found to be significant at p = 0.002. The results of a Newman-Keuls test to isolate significant differences are presented in Table D-20.

Table D-20. Newman-Keuls Results for Main Effect of Treatment on the Total Number of Negative Critical Incidents (Expert 1 Data)

		Increasing Order (n=100; MSE = 0.6837)				
		1	2	3		
	Level*	1	2	3		
	Mean	0.47	0.63	0.94	r	CDn-k**
Increasing	1		0.16	**0.47**	3	0.291882
Order	2			**0.31**	2	0.241444
	3					

*Level 1 = Remote/Reporting; Level 2= Lab/Reporting; Level 3=Expert 1 (for Lab/Non-Reporting participants)
**Studentized Range Statistic calculated based on df=24

Expert 2 Data

The main effect of Treatment was found to be significant at p = 0.025. Newman-Keuls test results are presented in Table D-21.

Table D-21. Newman-Keuls Results for Main Effect of Treatment on the Total Number of Negative Critical Incidents (Expert 2 Data)

		Increasing Order (n=100; MSE = 0.6041)				
		1	2	3		
	Level*	3	1	2		
	Mean	0.310	0.47	0.63	r	CDn-k
Increasing	3		0.160	**0.320**	3	0.274365
Order	1			0.16	2	0.226954
	2					

*Level 1 = Remote/Reporting; Level 2= Lab/Reporting; Level 3=Expert 1 (for Lab/Non-Reporting participants)
**Studentized Range Statistic calculated based on df=24

D.5.2 Main Effect of Day

Expert 1 Data

The main effect of Day was found to be significant at p < 0.001. Newman-Keuls test results are presented in Table D-22.

Table D-22. Newman-Keuls Results for Main Effect of Day on the Total Number of Negative Critical Incidents (Expert 1 Data)

		Increasing Order (n=60; MSE = 0.5744)						
		1	2	3	4	5		
	Level	5	3	4	1	2		
	Mean	0.2667	0.5667	0.7333	0.8667	0.9667	r	CDn-k**
Increasing	a5	---	**0.3**	**0.4666**	**0.6**	**0.7**	5	0.3894
Order	a3		---	0.1666	0.3	**0.4**	4	0.3659
	a4			---	0.1334	0.2334	3	0.3327
	a1				---	0.1	2	0.2769
	a2					---		

**Studentized Range Statistic calculated based on df=60

Expert 2 Data

The main effect of Day was found to be significant at p < 0.001. Newman-Keuls test results are presented in Table D-23.

Table D-23. Newman-Keuls Results for Main Effect of Day on the Total Number of Negative Critical Incidents (Expert 2 Data)

		Increasing Order (n=60; MSE = 0.3254)					
		1	2	3	4		
	Level	5	3	4	1 and 2		
	Mean	0.2	0.3333	0.3833	0.7167	r	CDn-k**
Increasing	a5		0.1333	0.1833	**0.5167**	4	0.2754
Order	a3			0.0500	**0.3834**	3	0.2504
	a4				**0.3334**	2	0.2084
	a1 and a2						

**Studentized Range Statistic calculated based on df = 60

D.5.3   Interaction of Interface x Treatment (Expert 1 Data only)

The interaction of Interface and Day was significant at p = 0.036. Newman-Keuls test results are presented in Table D-24.

Table D-24. Newman-Keuls Results for Interaction of Interface and Treatment on the Number of Negative Critical Incidents (Expert 1 Data)

	Increasing Order (n=50; MSE = 0.7163; I=Interface; T=Treatment)							
	1	2	3	4	5	6		
Level	I1T1	I1T2	I1T3	I2T1	I2T2	I2T3		
Mean	0.24	0.34	0.4	0.7	0.92	1.48	r	CDn-k
I1T1	---	0.1	0.16	**0.46**	**0.68**	**1.24**	6	0.49073
I1T2		---	0.06	0.36	**0.58**	**1.14**	5	0.46919
I1T3			---	0.3	**0.52**	**1.08**	4	0.44166
I2T1				---	0.22	**0.78**	3	0.40216
I2T2					---	**0.56**	2	0.33514
I2T3						---		

**Studentized Range Statistic calculated based on df=24

D.5.4 Interaction of Interface x Day

Expert 1 Data

The interaction of Interface and Day was significant at p < 0.0001. Newman-Keuls test results are presented in Table D-25.

Table D-25. Newman-Keuls Results for Interaction of Interface and Day on the Number of Negative Critical Incidents (Expert 1 Data)

Increasing Order (n=30; MSE = 0.5126; I = Interface; D = Day)

	1	2	3	4	5	6	7	8		
Treatment	**I1D3**	**I1D4 & I1D5**	**I1D2**	**I2D5**	**I2D1**	**I1D1 & I2D3**	**I2D4**	**I2D2**	**r**	**CDn-k**
Mean	**0.067**	**0.100**	**0.300**	**0.433**	**0.667**	**1.067**	**1.367**	**1.633**		
I1D3		0.033	0.233	0.367	**0.600**	**1.000**	**1.300**	**1.567**	8	0.5804
I1D4& I1D5			0.200	0.333	**0.567**	**0.967**	**1.267**	**1.533**	7	0.5634
I1D2				0.133	0.367	**0.767**	**1.067**	**1.333**	6	0.5438
I2D5					0.233	**0.633**	**0.933**	**1.200**	5	0.5202
I2D1						**0.400**	**0.700**	**0.967**	4	0.4889
I1D1& I2D3							0.300	**0.567**	3	0.4444
I2D4								0.267	2	0.3699
I2D2										

**Calculated based on df = 60

Expert 2 Data

The interaction of Interface and Day was significant at p < 0.0001. Newman-Keuls test results are presented in Table D-26.

Table D-26. Newman-Keuls Results for Interaction of Day and Interface on the Number of Negative Critical Incidents (Expert 2 Data)

Increasing Order (n=30; MSE = 0.3324; I = Interface; D = Day)

	1	2	3	4	5	6	7	8	9	10		
Treatment	**I1D3**	**I1D5**	**I1D2**	**I1D4**	**I2D1**	**I2D5**	**I2D4**	**I2D3**	**I1D1**	**I2D2**	**r**	**CD****
Mean	**0.067**	**0.1**	**0.3**	**0.367**	**0.6**	**0.633**	**0.8**	**0.9**	**1.133**	**1.567**		
I1D3		0.033	0.233	0.300	**0.533**	**0.567**	**0.733**	**0.833**	**1.067**	**1.500**	10	0.4895
I1D5		---	0.2	0.267	**0.5**	**0.533**	**0.7**	**0.8**	**1.033**	**1.467**	9	0.4789
I1D2			---	0.067	0.300	0.333	**0.500**	**0.600**	**0.833**	**1.267**	8	0.4674
I1D4				---	0.233	0.267	**0.433**	**0.533**	**0.767**	**1.200**	7	0.4537
I2D1					---	0.033	0.200	0.300	**0.533**	**0.967**	6	0.4379
I2D5						---	0.167	0.267	**0.500**	**0.933**	5	0.4189
I2D4							---	0.1	0.333	**0.767**	4	0.3937
I2D3								---	0.233	**0.667**	3	0.3579
I1D1									---	**0.433**	2	0.2979
I2D2										---		

**Studentized Range Statistic calculated based on df = 60

D.6 **NEWMAN-KEULS ANALYSES FOR NUMBER OF NEGATIVE CRITICAL INCIDENTS REPORTED**

D.6.1 Main Effect of Treatment

Expert 1 Data

The main effect of Treatment was significant at $p < 0.0001$. Newman-Keuls test results are presented in Table D-27.

Table D-27. Newman-Keuls Results for Main Effect of Treatment on the Total Number of Negative Critical Incidents (Expert 1 Data)

		Increasing Order (n=100; MSE = 0.3648)				
		1	2	3		
	Level*	1	2	3		
	Mean	0.15	0.38	0.74	r	CDn-k**
Increasing	1		**0.23**	**0.59**	3	0.2132
Order	2			**0.36**	2	0.1764
	3					

*Level 1 = Remote/Reporting; Level 2= Lab/Reporting; Level 3=Expert 1 (for Lab/Non-Reporting participants)
**Studentized Range Statistic calculated based on df=24

Expert 2 Data

The main effect of Treatment was significant at $p < 0.0001$. Newman-Keuls test results are presented in Table D-28.

Table D-28. Newman-Keuls Results for Main Effect of Treatment on the Total Number of Positive Critical Incidents (Expert 2 Data)

		Increasing Order (n=100; MSE = 0.26)				
		1	2	3		
	Level*	3	1	2		
	Mean	0.03	0.15	0.38	r	CDn-k**
Increasing	3	---	0.12	**0.35**	3	0.1800
Order	1		---	**0.23**	2	0.1489
	2			---		

*Level 1 = Remote/Reporting; Level 2= Lab/Reporting; Level 3=Expert 2 (for Lab/Non-Reporting participants)
**Studentized Range Statistic calculated based on df=24

D.6.2 Interaction Between Interface and Treatment (Expert 2 Only)

Newman-Keuls test results for the interaction of Interface and Treatment are presented in Table D-29.

Table D-29. Newman-Keuls Results for Interaction of Interface and Treatment on the Number of Positive Critical Incidents (Expert 2 Data)

Increasing Order (n=50; MSE = 0.18741; I=Interface; T=Treatment)

	1	2	3	4	5	6		
Level	I2T3	I1T3	I2T1	I1T1	I1T2	I2T2		
Mean	0.02	0.04	0.12	0.18	0.26	0.5	r	CDn-k
I2T3	---	0.02	0.1	0.16	**0.24**	**0.48**	6	0.2510
I1T3		---	0.08	0.14	0.22	**0.46**	5	0.2400
I2T1			---	0.06	0.14	**0.38**	4	0.2259
I1T1				---	0.08	**0.32**	3	0.2057
I1T2					---	**0.24**	2	0.1714
I2T2						---		

**Studentized Range Statistic calculated based on df=24

D.6.3   Interaction of Day x Treatment (Expert 1 Only)

Results of a Newman-Keuls test of the interaction of Interface and Treatment are presented in Table D-30.

Table D-30. Newman-Keuls Results for Interaction of Day and Treatment on the Number of Positive Critical Incidents (Expert 1 Data)

Increasing Order (n=20; MSE = 0.2807; D=Day, T=Treatment; based on df=60)

	1	2	3	4	5	6	7	8	9	10		
Level	Group A	Group B	D4T1	D1T2	Group E	Group F	D2T2	Group H	D4T3	D1T3		
Mean	0.1	0.15	0.25	0.3	0.35	0.4	0.5	0.6	1	1.1	r	CDn-k
Group A	---	0.05	0.15	0.2	0.25	0.3	0.4	0.5	**0.9**	**1**	10	0.5509
Group B		---	0.1	0.15	0.2	0.25	0.35	0.45	**0.85**	**0.95**	9	0.5390
D4T1			---	0.05	0.1	0.15	0.25	0.35	**0.75**	**0.85**	8	0.5260
D1T2				---	0.05	0.1	0.2	0.3	**0.7**	**0.8**	7	0.5106
Group E					---	0.05	0.15	0.25	**0.65**	**0.75**	6	0.4928
Group F						---	0.1	0.2	**0.6**	**0.7**	5	0.4715
D2T2							---	0.1	**0.5**	**0.6**	4	0.4431
Group H								---	**0.4**	**0.5**	3	0.4028
D4T3									---	0.1	2	0.3353
D1T3										---		

Where:
    Group A = {D3T1, D5T1}
    Group B = {D1T1, D2T1}
    Group E = {D3T2, D4T2}
    Group F = {D5T2, D5T3}
    Group G = {D2T3, D3T3}
*Level 1 = Remote/Reporting; Level 2= Lab/Reporting; Level 3=Expert 1 (for Lab/Non-Reporting participants)
**Studentized Range Statistic calculated based on df=60

D.6.4 Interaction of Interface x Day

Expert 1 Data

Newman-Keuls test results for the interaction between Interface and Day are presented in Table D-31.

Table D-31. Newman-Keuls Results for Interaction of Day and Interface on the Number of Positive Critical Incidents (Expert 1 Data)

	Increasing Order (n=30; MSE = 0.2630; I = Interface; D = Day)										
	1	2	3	4	5	6	7	8	9		
Treatment	**I1D3**	**Group B**	**I2D4**	**I1D2**	**I2D5**	**I2D2**	**I2D3**	**I1D4**	**I1D1**		
Mean	**0.033**	**0.200**	**0.300**	**0.333**	**0.400**	**0.500**	**0.667**	**0.767**	**0.833**	**r**	**CDn-k****
I1D3	---	0.167	0.267	0.300	0.367	**0.467**	**0.633**	**0.733**	**0.800**	9	0.4260
Group B		---	0.100	0.133	0.200	0.300	**0.467**	**0.567**	**0.633**	8	0.4157
I2D4			---	0.033	0.100	0.200	0.367	**0.467**	**0.533**	7	0.4035
I1D2				---	0.067	0.167	0.333	**0.433**	**0.500**	6	0.3895
I2D5					---	0.100	0.267	**0.367**	**0.433**	5	0.3726
I2D2						---	0.167	0.267	0.333	4	0.3502
I2D3							---	0.100	0.167	3	0.3183
I1D4								---	0.067	2	0.2650
I1D1									---		

Where:
  Group B = {I1D5, I2D1}

**Calculated based on df = 60

Expert 2 Data

Newman-Keuls test results for the two-way interaction of Interface and Day are presented in Table D-32.

Table D-32. Newman-Keuls Results for Interface x Day Interaction on the Number of Positive Critical Incidents (Expert 2 Data)

	Increasing Order (n=30; MSE=0.1337; I=Interface; D=Day)									
	1	2	3	4	5	6	7	8		
Treatment	I1D3	I2D4	Group C	I2D5	I1D1	I2D3	I1D4	I2D2		
Mean	0.000	0.067	0.100	0.233	0.267	0.300	0.333	0.367	r	CDn-k
I1D3		0.067	0.100	0.233	**0.267**	**0.300**	**0.333**	**0.367**	8	0.2964
I2D4			0.033	0.167	0.200	0.233	0.267	**0.300**	7	0.2877
Group C				0.133	0.167	0.200	0.233	0.267	6	0.2777
I2D5					0.033	0.067	0.100	0.133	5	0.2657
I1D1						0.033	0.067	0.100	4	0.2497
I2D3							0.033	0.067	3	0.2270
I1D4								0.033	2	0.1889
I2D2										

Where:
  Group C = {I1D2, I1D5, I2D}

**Studentized Range Statistic calculated based on df=60

D.6.5 Interaction of Interface x Treatment x Day (Expert 1 Only)

A Newman-Keuls test results for the three-way interaction of Interface, Treatment, and Day are presented Table D-33.

Table D-33. Newman-Keuls Results for Interface x Day x Treatment Interaction on the Number of Positive Critical Incidents (Expert 1 Data)

Increasing Order (n=10; MSE = 0.2807; I=Interface, D=Day, T=Treatment, based on df=60)

Level	1 Group A	2 Group B	3 Group C	4 Group D	5 I1D1T2	6 Group F	7 I2D5T2	8 Group H	9 I2D2T2	10 I2D3T3	11 I1D4T3	12 I1D1T3	r	CDn-k
Total	0	1	2	3	4	5	6	7	8	11	13	19		
Group A		1	2	3	4	5	6	7	**8**	**11**	**13**	**19**	12	8.06
Group B			1	2	3	4	5	6	7	**10**	**12**	**18**	11	7.92
Group C				1	2	3	4	5	6	**9**	**11**	**17**	10	7.79
Group D					1	2	3	4	5	**8**	**10**	**16**	9	7.62
I1D1T2						1	2	3	4	**7**	**9**	**15**	8	7.44
Group F							1	2	3	6	**8**	**14**	7	7.22
I2D5T2								1	2	5	**7**	**13**	6	6.97
Group H									1	4	**6**	**12**	5	6.67
I2D2T2										3	5	**11**	4	6.27
I2D3T3											2	**8**	3	5.70
I1D4T3												**6**	2	4.74
I1D1T3														

Where:
Group A = {I1D3T1, I1D3T2, I2D4T1}
Group B = {I1D2T1, I1D3T3, I1D5T1, I2D1T1, I2D5T1}
Group C = {I1D1T1, I1D2T2, I1D5T2, I2D1T2, I2D2T1, I2D3T1, I2D4T2}
Group D = {I1D5T3, I2D1T3}
Group F = {I1D4T1, I1D4T2, I2D2T3, I2D5T3}
Group H = {I1D2T3, I2D3T2, I2D4T3}

D.7   POST-TEST QUESTIONNAIRE DATA

Post-test questionnaires were administered to all participants to solicit subjective feedback about VEMS and the usability evaluation tools (ex. critical incident report form, Usability Evaluation web site).   Participants were asked to specify the positive and negative aspects of each interface as well as provide general comments. Table D-34 to Table D-41 present the responses given.

Table D-34. Positive Aspects of the Web Interface

POSITIVE ASPECT – WEB INTERFACE
Small pop-up windows for the forms
All tasks I performed were easy to figure out on the system.
Functions are mostly well organized and categorized
Distinct colors and simple layout make it visually easy to see/find what you need
Quick loading
Organized and provide useful descriptions of what certain functions are for.
It gave examples to functions that quickly illustrated to me what to do and how the function could be used.
Provides lots of useful features
Automatic reply messages
All the critical "help" was on the page with the commands that you are using; making it easy to remember what info it was looking for.
Ease of use
All the functions were easily accessible from the top of the page.
Set up like most web sites with iconic buttons
Most things have shortcuts and are easy to do quickly
Easy to use interface with easily understood buttons
Easy to use and comprehend
Easy to navigate.
It's overall very straightforward
This web site was extremely clear and easy to use, almost a zero learning curve.
Quick buttons for frequently used commands
Eye Pleasing site.
Express lane
It was easy to get around and find things within the web site. Except it took some time to configure the polling option
Innovative technology makes it fun
Consistent from anywhere.
Visual icons
Icon for each function
Help is actually helpful
Very Organized
Customized reply messages
Quick adds and deletions were very efficient.
Really easy to use
Simply organized, not too hard to eventually find something.
Trash Can
Easy to understand, graphic icons are easily identifiable.
I like the FYI statements and links offered to help you learn more about the system.
Efficient
Voice e-mail capability

POSITIVE ASPECT – WEB INTERFACE
Good amount of icons, so I didn't have to scroll down a huge list to find what I wanted
There are step by step instructions for the different functions
Icons on the top of each page
Icons made it easier to identify which page I needed to go to [in terms of reconfigurement].
The links to functions are clearly labeled and easy to see.
Well structured and laid out, which enables easy and efficient navigation.
Express function buttons
Easy to manage email.
Fast
Very easy to set up and change my preferences
The layout is easy to follow and understand
Instructions were clear
It is easy to find what you are looking for.

Table D-35. Negative Aspects of the Web Interface

NEGATIVE ASPECT – WEB INTERFACE
Not very useful.
The hint section usually is a hint to your real password and not another password
The page should allow for more information to be stored in the address books. For instance phone numbers and addresses. It would be very nice to be able to access this information over the phone.
The personal Profile home link should be a large icon not a small link in the upper right hand corner of the screen.
Requires high resolution
You can't control the performance of the phone interface significantly (see negative aspects of the phone interface, those are the things I'd like to be able to control).
Needs better organization
Useful advanced commands (delete older than, list from sender, etc.) are buried in non-obvious places
Slow in loading
Have error messages link to the appropriate area in the help page
I don't like the icons on the top of the screen. I prefer them along the side
Challenging to find some functions, but with some practice is fine
Some of the terms were vague; mini-descriptions under options like "handling" would make it easier to know you were choosing the right option
Does not provide clear messages of what the user should do to fix a problem if it occurs or if the user is experiencing difficulty handling a certain function/feature of the web site.
Every time I went to the next "layer" of the website I had to scroll down to where I wanted to be.
Not as much hints and help as needed
In my address book I would like to add people who have sent me a message without typing there address in
It would be nice to have the option to retrieve email via the website.
There are confirmation messages for some field but not others
A few more links could be added to text in page
The boxes where the user has to type in the account number and pass code could be bigger and more noticeable, perhaps moved to a better location because for first time users they are hard to notice/locate.
I did not notice a quick reference guide on the opening page. One would have to search for that through the page.
A little difficult to understand for someone not quite used to computers and e-mail

NEGATIVE ASPECT – WEB INTERFACE
It was a hassle to have to search for certain functions or menus that was not made apparent.
Menus and Icon locations could be improved upon.
Not so user friendly
The trashcan icon, sometimes you don't know if it is going to clear completely out of it or just from a list.
No outbox, so you don't know if you sent something or not.
Profile page was difficult to get to the first time around
You cannot check your e-mail there.
Login boxes should be more visible [such as being either in the center or at the bottom of the page to make it more visible for the user]
When you want to add a name or address, it asks you twice if you are done. (It asks if you want to add another... it is a little annoying if you only have one name to add, but then again kind of convenient if you want to add more than one name.)

Table D-36. Positive Aspects of the Voice Interface

POSITIVE ASPECT – VOICE INTERFACE
Easy to reply
The voice recognition works surprising well.
Efficient if I were to travel a lot
It was very clear at each point where you were in the system, and what the system needed.
Very convenient to be able to check email wherever there is a phone
The VEMS telephone system is a good idea
The forwarding and reply features are nice. I can see where a businessperson would become reliant on these features.
Provides help for first time use
Good selection of commands
The option to use the touch-tone commands will allow for faster overall interaction, such as the interrupt option.
I think the menu system is well organized. I like how it tells you from the start how many new and saved messages you have.
Good voice recognition
Ease of use
It has a very organized structure that facilitates easy use.
Comprehensive
If you can't get to your email, it is really convenient to be able to check them from any phone
The system, when reading the messages is clear.
I love the dictation command. I will be interested to send a real dictation and see how easy it is on the receiving end.
Surprisingly easy to navigate through
You can access your email without a computer
Reliable for getting email without computer
Fairly easy to understand voice
Most everything is by voice and not by numbers on the keypad
The system was polite
Cool to hear your emails
Can check from any phone
The electronic speech handles a lot of common phrases well.
Access from anywhere
Every I say Goodbye, the system responds quickly.

POSITIVE ASPECT – VOICE INTERFACE
Can get email anywhere no problem
Does not require a computer to check email
Toll free number
Sending replies and forwarding messages was fairly simple.
Lots of Fun
Voice Recognition is actually pretty good for the most part.
Could have some advantages for visually impaired people
It offers many useful built in commands.
I like being able to forward and dictate replies to e-mail.
Simple to use.
Different
Easy accessibility
Easy to restore messages that you deleted
The commands are straightforward so it is relatively easy to use
Extra Options (Send Reply, List Messages From.., Remove Messages More than....)
Easy to get used to after a few days
It is easy to recover from mistakes that I made.
With practice the system is easily understood and expectations of system are clearer.
I like the convenience of not being at a computer (or on the internet) to check my e-mail
Easy access to emails.
Effective
Respond both voice and touch tone commands
The convenience of its accessibility is its major advantage
Easy to reply to messages
It is extremely convenient [better than toting a laptop and trying to connect it to a nearby phone line]
It is convenient to call in and check e-mail.

Table D-37. Negative Aspects of the Voice Interface

NEGATIVE ASPECT – VOICE INTERFACE
Some confusion over commands
It should recognize the "speak" commands while it is talking.
Some functions are similar sounding (ex. Remove messages, Restore messages, Prioritize and Previous)
It takes much longer to listen to messages than to read them, so the system felt slow and like a waste of time
Takes a few days to get used to saying commands right in order for system to understand you
The system has hard time understanding certain commands, causing the user to repeat the command or the system misunderstands and performs another different command.
I think it is too easy to delete messages by mistake.
Too many options provided sometimes
Not being able to set up separate folders to organize my e-mails
The menu system is organized well, although sometimes I still get confused as to where I am within it. The best example is the critical incident that I reported today, where the account thought I said help after I listed the name of a person to forward the message to. So I was in the help menu, which then meant I was unsure as to whether the names I had already listed as people to forward the message were still there, or whether I had to start over.
There is no voice-activated "stop reading"
It does not have a function to remove all messages from a particular sender.

NEGATIVE ASPECT – VOICE INTERFACE
If you get lost it is hard to recover at this point
Some of the commands were not as natural; like "send reply" instead of just reply
Navigating through messages not easy, despite a few useful commands (get first save message, etc); should be more commands such as listing messages from someone NOT listed in address book
Should have more voice commands such as "stop reading message"
Speech recognition problems
Not being able to go directly to a message
The primary reason of difficulty, when I had it, was in the system not properly recognizing the commands that I said. For example, confusion occurred between "read" and "remove", and "yes" and "help".
Sometimes skips words and it sometimes doesn't understand what you say
Can't interrupt the computer with instructions mid reading
Takes longer than reading e-mail on a computer
Cannot state my account number like forty-three, ten, fifteen etc.
At times I was confused as to exactly where I was.
It is hard to use because it isn't easy to hear him
Can be slow
System sometimes doesn't even listen to you until a second or so after it stops talking.
Synthesized voice
No different Voice between different functions or messages
Time consuming
Can't send a dictated reply in writing (like a regular email)
It always misunderstands what I say and I have to repeat the commands numerous times.
Slow to navigate through messages
I was not able to access my old messages.
It is frustrating when he doesn't understand you
Voice activation gets confused sometimes
System doesn't' t listen to you while it's speaking.
Navigating message list
Can not recognize my voice sometimes
Voice recognition problems yield frustration and it is tough to find old messages quickly.
It can be frustrating to locate particular e-mails when not familiar with all of the commands
The automatic computer voice was annoying and frustrating to listen to at times.
There is no way to configure the system using the names of the sender only
Very difficult to recover from a mistake
At first I think the person should say a couple of words and numbers so the machine would get used to it. It would make it a lot easier
The voice system is not sensitive enough for 'speech imperfections,' different pronunciations from the standard for instance especially for someone with an accent.
It is difficult to access names that are not in the Address Book (web page)
System kept misinterpreting commands [notably "_____" message commands and the remove messages more than (x) days old]
I often did not speak clear enough and this resulted in the wrong function being performed.
Difficult to pick up system the first few times because it is different from my usual email system.
it is difficult to visualize e-mail organization at times
Login on was a problem, especially if you do not want anyone else around to hear your password and account information if you were to ask to say it.
Once again, no outbox
I do not like the fact that the messages are queued and you have to follow a set pattern to get to

NEGATIVE ASPECT – VOICE INTERFACE
your messages. There should be a way to skip messages and attend to the ones you want to hear without having to 'wait in line' so to speak (this has to do with previously saved messages)
Computerized voice
Voice used by the telephone system made it difficult to understand what he was saying [often had to repeat messages to understand content]
I did not know how to locate a specific e-mail that I was looking for, so I had to go through all 13 messages (with subject lines) to find the one message that I knew I wanted.

Table D-38. Positive Aspects of the Online Critical Incident Report Form

POSITIVE ASPECT – CRITICAL INCIDENT REPORT FORM
Simple
The technique did effectively prompt me to give a complete account of the incident.
It was fun!!!!! I enjoyed the possibilities.
Clear, outlined way to report, so you didn't worry about if you were missing telling something
Helped to make clearer in my mind how I felt about what I was reporting
The examples help when the user has no idea what to write
It made me more aware of my own frustration threshold and dealing with computers and voice systems (telephone).
Easy to use
Better hard data than surveys
The way it was divided into reporting our actions and then the incident made it easier to be clear as to what exactly happened.
It made me more aware that things aren't working the way should.
Provides useful information
Since there are so many fields you can pinpoint the problem
Even though sometimes it was awkward to stop what we were doing, having the c.i. technique button reminded us to remember the incidents and then to report them before we forgot the circumstances.
Gave the information the developers wanted
It lets you say what you feel about the system
I feel like I'm having a say in an important product's development
Didn't take too much time
Very detailed.
It was there to report things that either went wrong or right
Allows you to register frustrations
Easy to access from anywhere.
Clear way of understanding one's interaction with computer software
Every time I need help, just click help then get a good example
Helpful to experimenter
Not hard to figure out, ultimately.

Table D-39. Negative Aspects of the Online Critical Incident Report Form

NEGATIVE ASPECT – CRITICAL INCIDENT REPORT FORM
I had trouble describing my feelings
Some of the prompts seemed repetitive or superfluous.
It was not accessible while on the phone
Sometimes I wasn't sure if my satisfaction was an important factor, or just ease of system (i liked and didn't like things that I didn't know if they fit into the incident technique
Sometimes I wished there were more gradients to the ratings used
Perhaps have a suggestion box?
I found myself looking for positive responses. Even though there was plenty of positive incidents it was harder for me to report since having the program work would seem natural.
A little lengthy.
Switching back and for between the web sites, but you can't really do it any other way
A little easier to report something negative, such as when you are frustrated with what you are doing. Maybe a little negative-biased, although you did a good job reminding us to report positive experiences as well.
I couldn't offer suggestions as to possible fixes
Some of the questions didn't fit every situation, so I would repeat information
Sometimes it was not easy to quantify the incident
I found that going to the CI report immediately after an incident was distracting and a little repetitive. I often wrote what happened and reported at the end of the task.
Sometimes it is not so easy to get motivated to report an incident
It is often hard to separate the problem into so many different fields, and it takes a fair bit of time
I found it both a positive and a negative that we had to stop what we were doing to fill out the incident report. It sometimes broke the flow of the task. (I think this is more positive than negative, because it's so easy to forget to report it if you don't do it right away)
Sometimes skips words and it sometimes doesn't understand what you say
Only do one if you are impressed or depressed with the way the system worked
It's easier to spot negative incidents than positive
The report does not reflect the fact that most of the time its a lot of little, low rated, details added up that make an impact
I felt somewhat discouraged to do the report because some problems were not too extreme, one-way or the other.
Some of the question weren't easily understood
Tedious; time-consuming.
The questions seem a little ill-suited to the process of naturally describing an "incident", although they basically work.
It was hard getting motivated after a while, especially after my ability to use the VEMS system improved
Would rather just be able to comment on the use
I might be more critical of the system if I'm having a bad day or something like that
Not all questions needed to be answered for all incidents
I was more prone to report negative incidents.
Positive critical incidents often don't come to mind.

Table D-40. Comments pertaining to the Statement "I will continue to use VEMS"

GENERALS COMMENT - VEMS
I think it is a useful tool. However, I won't be using it in the immediate future.
The system is great. I will continue to use the account. My use of the system while in Blacksburg will likely be limited because of easy internet access, however I will use it while traveling.
yes I will if I travel. It will serve to get occasional messages since I do not have email at my house.
It is cool; unfortunately so many emails are really long that I wouldn't want to listen to them all; if I was expecting important information via email, this would be a nice way to always have access to it, but for the general emails from friends, or emails with a lot of information, like the news, and not just weather, it takes too long to listen to. I will probable use it for fun, but not for all of my email
I think I will use VEMS, but maybe just as a stand-alone email address first. For example, I may give out my VEMS email address to friends to use for social messages, but I don't think I will link my Virginia Tech account to VEMS because I prefer to read important messages that may contain a lot of information and details. VEMS works well with brief messages, but when the messages are verbose, I think details can be forgotten more easily.
I will continue to experiment with the system. I can see how valuable it could become when traveling and away from email access. When listening to the computer voice, I find it hard sometimes to comprehend what is being read. I would not want this to replace written email.
I might continue to use it but it is not practical for professional people because they in general receive a lot of e-mails sometimes in the hundreds, it is OK for some home users new to e-mail but I guess they would probably not want to get use to a whole new system just as they got used to one
When I don't have access to the Internet for a couple days, but do have access to a phone.
I think it would be nicer/easier to have all statements actually human read, such as "Would you like to forward this message to ____", and then fill in the blank with the computerized voice. It involves less concentration overall, although this may be something we get used to over time. I will definitely continue to use this account and think that I will find it very useful.
I might continue to use the account, but most likely will just continue to use standard email format.
I will probably keep the account for a little while, to see if I do keep using it. I set it to retrieve email form my VT account, so I'll see if I ever use it.
I prefer to read my email in person, and have time to think about replies rather than leaving an answering machine type message.
Yes I will continue to use VEMS. It would be very handy to have when I go on trips.
It is a good idea and is very nice to be able to check email from anywhere. I would certainly like to continue using it. I would imagine that with practice and experience with the phone system, I would become more efficient and enjoy using it.
I don't think that I will be using it.
I removed my account information from the system; I don't imagine I'll do much more than maybe try it with a few different types of text after this. It was just too frustrating that the system wouldn't listen to you while it was speaking, you should at least be able to configure it to shut up when you start talking. In the same vein, the ability to at least say "pause" or "stop" would make the system truly hands-free. As it is, it feels like you're being assaulted with the text of your mail, it doesn't ever pause for breaths; it's the ultimate one-sided conversation when it should be like an assistant that responds to your cues. The web site was much more palatable, in fact I ended up using it while I was listening to the phone messages, which is probably not going to happen much. They should really add the ability to actually read your messages from the web site, that way it really would be a single source for all email needs and conditions, instead of something you'd only use in the unlikely event that you had a phone but not a computer. Even a palm pilot can be used to read email using a phone line, so the usefulness of the voice reading service doesn't really stand up well by itself, especially since it's got a lot of disadvantages vs. reading your email as text using conventional means.
Not really, too slow and a little annoying. Suggestion, use different voice for Command and Message Reading

Table D-41. General Comments Pertaining to the Usability Evaluation Web Site

GENERAL COMMENTS – USABILITY EVALUATION WEB SITE
Better than most...but I am not a big fan of tables
It was clear and concise.
It is often redundant, but that is necessary to cover all possible reactions I may have had.
i know you couldn't help this but the size of the font and pages made it very awkward to navigate; really inconvenient b/c you couldn't see all the information available
Nicely designed with forms and new browsers launching when reporting a Critical Incident, so place is not lost on page you are on.
I can tell that a lot of work went into making the site. I believe that much thought was given to how people would understand and navigate through the site.
It was fine; it worked well.
Sometimes I know that I left out some info you were looking for in the sense that you had a few things listed in the help link, but I forgot to check and just answered the question in an intuitive way. I liked it when you put a reminder of where to find the info that we needed to complete a certain task, so that if we were not paying attention closely enough to the content of the emails the day before we could still perform the tasks of the day. (ex. - that info was given to you in the email sent to you yesterday from Graham Roeburg). Not really a comment on the web site, but more the study in general.
Sometimes I could not remember all the details of the "story line". So when it would ask me details about the "meeting" coming up or what hotel I was at I did not remember it.
I felt lost in the immense amount of links. Information and examples for the incident reports could have been incorporated into the text further, rather than having to link to another area. It took me some time to remember where I had been and where I was going.
It was easy to get to and it was helpful.
The only minor drawbacks of the Usability web site were its dependence on using Internet Explorer, and the problems with the JavaScript that undermined the feedback loop where the system would ask for different information, you'd hit the back button, change the information it requested, then be forced to reselect every radio button before it would let you continue. I could explain my theory on why the JavaScript problems happen if you'd like.
The tasks were presented in a clear fashion but could use a bit more detail in explaining what exactly to do especially the first couple of times because of unfamiliarity of the system to the user.
Very well organized and easy to navigate.
The instructions given were lucid and clear
Maybe you should tell the subject to read all e-mail messages, than reply and finally do the tasks desired. Sometimes I replied someone and later on, I found out that I should give another reply. But I fixed that in the next session. Overall it was very easy to use and realistic. You gave both scenarios of "junk mail" and important situations. I think maybe an interaction between VEMS and the "usual" e-mail system would be a very good experience, since I believe most of us going to use both systems together
I only used it for straightforward tasks so it was not hard to use or navigate at all. I confess that I really didn't pay attention to much else apart from my daily tasks.
The site was very usable.

D.8 ANOVA RESULTS FOR POST-TEST QUESTIONNAIRE DATA

D.8.1 Comparisons Amongst All Treatment Groups

Analyses of variance were conducted for each post-test questionnaire rating question to assess the effects of Treatment (between-subject factor; 3 levels) on the mean ratings allocated. Ratings were assigned based on a five-point Likert-scale where 1=strongly disagree and 5=strongly agree.

D.8.2 It was simple to use the VEMS web site

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It was simple to use the VEMS web site". Insufficient evidence of a main effect of treatment was found.

Table D-42. Analysis of Variance for Question 1: It was simple to use the VEMS web site.

Source	DF	SS	MS	F	P
Treatment	2	0.867	0.433	0.68	0.517
Error	27	17.300	0.641		
Total	29	18.167			

D.8.3 I was able to complete my tasks quickly using the VEMS web site

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "I am able to complete my tasks quickly using the VEMS web site". Insufficient evidence of a main effect of treatment was found.

Table D-43. Analysis of Variance for Question 2: I was able to complete my tasks quickly using the VEMS web site.

Source	DF	SS	MS	F	P
Treatment	2	4.067	2.033	2.73	0.083
Error	27	20.100	0.744		
Total	29	24.167			

D.8.4 I am able to efficiently complete my tasks using the VEMS web site.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "I am able to efficiently complete my tasks using the VEMS web site". Insufficient evidence of a main effect of treatment was found.

Table D-44. Analysis of Variance for Question 3: I am able to efficiently complete my tasks using the VEMS web site.

Source	DF	SS	MS	F	P
Treatment	2	2.600	1.300	1.59	0.223
Error	27	22.100	0.819		
Total	29	24.700			

D.8.5 It was easy to learn to use the VEMS web site

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It was easy to learn to use the VEMS web site". A significant main effect of treatment at p=0.02 was found.

Table D-45. Analysis of Variance for Question 4: It was easy to learn to use the VEMS web site.

Source	DF	SS	MS	F	P
Treatment	2	2.400	1.200	4.56	0.020
Error	27	7.100	0.263		
Total	29	9.500			

A Newman-Keuls post hoc analysis was conducted to isolate the significant differences, the results of which are presented below.

Table D-46. Newman-Keuls Results for "It was Easy to Learn the VEMS web site"

	Increasing Order (n=10; MSE=0.263)			
	1	2		
Treatment	1	2 & 3		
Mean	4.1	4.7	r	CDn-k
1	---	**0.6**	2	0.474
2 & 3		---		

The results indicate that ratings allocated by the laboratory-based users (reporting and non-reporting) were significantly higher than those allocated by the remote users.

D.8.6 The VEMS web page gives error messages that clearly tell me how to fix problems.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The VEMS web page gives error messages that clearly tell me how to fix problems". Insufficient evidence of a main effect of treatment was found.

Table D-47. Analysis of Variance for Question 5: The VEMS web page gives error messages that clearly tell me how to fix problems.

Source	DF	SS	MS	F	P
Treatment	2	0.785	0.393	0.40	0.673
Error	22	21.375	0.972		
Total	24	22.160			

D.8.7 Whenever I make a mistake using the VEMS web site, I recover easily and quickly.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "Whenever I make a mistake using the VEMS web site, I recover easily and quickly". Insufficient evidence of a main effect of treatment was found.

Table D-48. Analysis of Variance for Question 6: Whenever I make a mistake using the VEMS web site, I recover easily and quickly.

Source	DF	SS	MS	F	P
Treatment	2	0.20	0.10	0.07	0.934
Error	26	38.56	1.48		
Total	28	38.76			

D.8.8 The information (ex. on-line help, on-screen messages) provided by the VEMS web pages is clear.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The information (ex. on-line help, on-screen messages) provided by the VEMS web pages is clear". Insufficient evidence of a main effect of treatment was found.

Table D-49. Analysis of Variance for Question 7: The information (ex. on-line help, on-screen messages) provided by the VEMS web pages is clear.

Source	DF	SS	MS	F	P
Treatment	2	1.867	0.933	1.92	0.166
Error	27	13.100	0.485		
Total	29	14.967			

D.8.9 It is easy to find the information I need using the VEMS web pages

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It is easy to find the information I need using the VEMS web pages". Insufficient evidence of a main effect of treatment was found.

Table D-50. Analysis of Variance for Question 8: It is easy to find the information I need using the VEMS web pages.

Source	DF	SS	MS	F	P
Treatment	2	2.067	1.033	1.60	0.220
Error	27	17.400	0.644		
Total	29	19.467			

D.8.10 The information provided by the VEMS web site is easy to understand.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The information provided by the VEMS web site is easy to understand". A significant main effect of Treatment was found at p=0.028.

Table D-51. Analysis of Variance for Question 9: The information provided by the VEMS web site is easy to understand.

Source	DF	SS	MS	F	P
Treatment	2	2.400	1.200	4.10	0.028
Error	27	7.900	0.293		
Total	29	10.300			

A Newman-Keuls post hoc analysis was conducted to isolate the significant differences, the results of which are presented below.

Table D-52. Newman-Keuls Results for " The information provided by the VEMS web site is easy to understand."

Increasing Order (n=10; MSE=0.293)

Treatmen	1	2		
t	1	2 & 3		
Mean	3.9	4.5	r	CDn-k
1	---	**0.6**	2	0.500
2 & 3		---		

The results indicate that ratings allocated by the laboratory-based users (reporting and non-reporting) were significantly higher than those allocated by the remote users.

D.8.11 The organization of information on the VEMS web pages is clear

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The organization of information on the VEMS web pages is clear". Insufficient evidence of a main effect of treatment was found.

Table D-53. Analysis of Variance for Question 10: The organization of information on the VEMS web pages is clear.

Source	DF	SS	MS	F	P
Treatment	2	2.600	1.300	1.46	0.251
Error	27	24.100	0.893		
Total	29	26.700			

D.8.12 I like using the VEMS web site

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "I like using the VEMS web site". Insufficient evidence of a main effect of treatment was found.

Table D-54. Analysis of Variance for Question 11: I like using the VEMS web site.

Source	DF	SS	MS	F	P
Treatment	2	2.067	1.033	1.77	0.190
Error	27	15.800	0.585		
Total	29	17.867			

D.8.13 The VEMS web site has all the functions and capabilities that I expect it to have

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The VEMS web site has all the functions and capabilities that I expect it to have". Insufficient evidence of a main effect of treatment was found.

Table D-55. Analysis of Variance for Question 12: The VEMS web site has all the functions and capabilities that I expect it to have.

Source	DF	SS	MS	F	P
Treatment	2	3.200	1.600	2.47	0.104
Error	27	17.500	0.648		
Total	29	20.700			

D.8.14 I am able to complete my emailing tasks quickly using the VEMS voice system.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "I am able to complete my emailing tasks quickly using the VEMS voice system". Insufficient evidence of a main effect of treatment was found.

Table D-56. Analysis of Variance for Question 16: I am able to complete my emailing tasks quickly using the VEMS voice system.

Source	DF	SS	MS	F	P
Treatment	2	7.47	3.73	2.93	0.071
Error	27	34.40	1.27		
Total	29	41.87			

D.8.15 I am able to efficiently complete my emailing tasks using the VEMS voice system

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "I am able to efficiently complete my emailing tasks using the VEMS voice system". Insufficient evidence of a main effect of treatment was found.

Table D-57. Analysis of Variance for Question 18: I am able to efficiently complete my emailing tasks using the VEMS voice system.

Source	DF	SS	MS	F	P
Treatment	2	4.867	2.433	2.86	0.075
Error	27	23.000	0.852		
Total	29	27.867			

D.8.16 It was easy to learn to use the voice system

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It was easy to learn to use the voice system". A significant main effect of Treatment was found at p=0.010.

Table D-58. Analysis of Variance for Question 19: It was easy to learn to use the voice system.

Source	DF	SS	MS	F	P
Treatment	2	9.800	4.900	5.47	0.010
Error	27	24.200	0.896		
Total	29	34.000			

A Newman-Keuls post hoc analysis was conducted to isolate the significant differences, the results of which are presented below.

Table D-59. Newman-Keuls Results for "It was easy to learn to use the voice system"

	Increasing Order (n=10; MSE=0.896)				
	1	2	3		
Treatment	1	3	2		
Mean	3.2	4.3	4.5	r	CDn-k
1	---	**1.1**	**1.3**	3	1.057
3		---	0.2	2	0.874
2			---		

The results indicate that ratings allocated by the laboratory-based users (reporting and non-reporting) were significantly higher than those allocated by the remote users.

D.8.17   The voice system gives error messages that clearly tell me how to fix problems.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The voice system gives error messages that clearly tell me how to fix problems". Insufficient evidence of a main effect of treatment was found.

Table D-60. Analysis of Variance for Question 20: The voice system gives error messages that clearly tell me how to fix problems.

Source	DF	SS	MS	F	P
Treatment	2	2.57	1.28	1.03	0.371
Error	26	32.40	1.25		
Total	28	34.97			

D.8.18 The information (ex. help or prompts) provided by the VEMS voice system is clear

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The information (ex. help or prompts) provided by the VEMS voice system is clear". A significant main effect of Treatment was found at p=0.050.

Table D-61. Analysis of Variance for Question 21: The information (ex. help or prompts) provided by the VEMS voice system is clear.

Source	DF	SS	MS	F	P
Treatment	2	5.928	2.964	3.37	0.050
Error	26	22.900	0.881		
Total	28	28.828			

A Newman-Keuls post hoc analysis was conducted to isolate the significant differences, the results of which are presented below.

Table D-62. Newman-Keuls Results for "It was easy to learn to use the voice system"

	Increasing Order (n=10; MSE=0.896)				
	1	2	3		
Treatment	1	3	2		
Mean	3	3.9	4	r	CDn-k
1	---	**0.9**	**1**	3	1.048
3		---	0.1	2	0.867
2			---		

The results indicate that ratings allocated by the laboratory-based users (reporting and non-reporting) were significantly higher than those allocated by the remote users.

D.8.19   Whenever I make a mistake using the voice system, I recover easily and quickly.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "Whenever I make a mistake using the voice system, I recover easily and quickly". Insufficient evidence of a main effect of treatment was found.

Table D-63. Analysis of Variance for Question 22: Whenever I make a mistake using the voice system, I recover easily and quickly.

Source	DF	SS	MS	F	P
Treatment	2	1.40	0.70	0.46	0.638
Error	27	41.30	1.53		
Total	29	42.70			

D.8.20   It was easy to navigate within the voice email system.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It was easy to navigate within the voice email system". Insufficient evidence of a main effect of treatment was found.

Table D-64. Analysis of Variance for Question 23: It was easy to navigate within the voice email system.

Source	DF	SS	MS	F	P
Treatment	2	3.80	1.90	1.59	0.222
Error	27	32.20	1.19		
Total	29	36.00			

D.8.21   It is easy to find the information I need using the VEMS voice system.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It is easy to find the information I need using the VEMS voice system". Insufficient evidence of a main effect of treatment was found.

Table D-65. Analysis of Variance for Question 24: It is easy to find the information I need using the VEMS voice system.

Source	DF	SS	MS	F	P
Treatment	2	1.867	0.933	1.29	0.291
Error	27	19.500	0.722		
Total	29	21.367			

D.8.22 The information provided by the voice system is easy to understand.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The information provided by the voice system is easy to understand". Insufficient evidence of a main effect of treatment was found.

Table D-66. Analysis of Variance for Question 25: The information provided by the voice system is easy to understand.

Source	DF	SS	MS	F	P
Treatment	2	0.801	0.400	0.52	0.599
Error	26	19.889	0.765		
Total	28	20.690			

D.8.23 The organization of email messages is clear.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The organization of email messages is clear". Insufficient evidence of a main effect of treatment was found.

Table D-67. Analysis of Variance for Question 26: The organization of email messages is clear.

Source	DF	SS	MS	F	P
Treatment	2	4.200	2.100	2.43	0.107
Error	27	23.300	0.863		
Total	29	27.500			

D.8.24 I like using the VEMS voice system.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "I like using the VEMS voice system". Insufficient evidence of a main effect of treatment was found.

Table D-68. Analysis of Variance for Question 27: I like using the VEMS voice system.

Source	DF	SS	MS	F	P
Treatment	2	6.87	3.43	2.10	0.142
Error	27	44.10	1.63		
Total	29	50.97			

D.8.25 The VEMS voice system has all the functions and capabilities that I expect it to have.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The VEMS voice system has all the functions and capabilities that I expect it to have". Insufficient evidence of a main effect of treatment was found.

Table D-69. Analysis of Variance for Question 28: The VEMS voice system has all the functions and capabilities that I expect it to have.

Source	DF	SS	MS	F	P
Treatment	2	0.267	0.133	0.13	0.874
Error	27	26.700	0.989		
Total	29	26.967			

D.8.26 It was easy to navigate through the Usability Evaluation web site

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It was easy to navigate through the Usability Evaluation web site". A significant main effect of Treatment was found at p=0.035.

Table D-70. Analysis of Variance for Question 49: It was easy to navigate through the Usability Evaluation web site.

Source	DF	SS	MS	F	P
Treatment	2	5.000	2.500	3.81	0.035
Error	27	17.700	0.656		
Total	29	22.700			

A Newman-Keuls post hoc analysis was conducted to isolate the significant differences, the results of which are presented below.

Table D-71. Newman-Keuls Results for "It was easy to navigate through the Usability Evaluation web site"

	Increasing Order (n=10; MSE=0.656)				
	1	2	3		
Treatment	1	2	3		
Mean	3.6	4.1	4.6	r	CDn-k
1	---	0.5	**1**	3	0.9041
2		---	0.5	2	0.7479
3			---		

The results indicate that ratings allocated by the non-reporting laboratory-based users were significantly higher than those allocated by the remote users.

D.8.27 It was easy to learn to use the Usability Evaluation web site

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It was easy to learn to use the Usability Evaluation web site". A significant main effect of Treatment was found at p=0.007.

Table D-72. Analysis of Variance for Question 50: It was easy to learn to use the Usability Evaluation web site.

Source	DF	SS	MS	F	P
Treatment	2	3.800	1.900	6.04	0.007
Error	27	8.500	0.315		
Total	29	12.300			

A Newman-Keuls post hoc analysis was conducted to isolate the significant differences, the results of which are presented below.

Table D-73. Newman-Keuls Results for "It was easy to learn to use the Usability Evaluation web site"

	Increasing Order (n=10; MSE=0.315)				
	1	2	3		
Treatment	1	2	3		
Mean	3.8	4.5	4.6	r	CDn-k
1	---	**0.7**	**0.8**	3	0.6265
2		---	0.1	2	0.5182
3			---		

The results indicate that ratings allocated by the laboratory-based users (reporting and non-reporting) were significantly higher than those allocated by the remote users.

D.8.28   The information provided by the Usability Evaluation web pages is clear.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The information provided by the Usability Evaluation web pages is clear". Insufficient evidence of a main effect of treatment was found.

Table D-74. Analysis of Variance for Question 51: The information provided by the Usability Evaluation web pages is clear.

Source	DF	SS	MS	F	P
Treatment	2	1.400	0.700	2.74	0.083
Error	27	6.900	0.256		
Total	29	8.300			

D.8.29   It is easy to find the information I need.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "It is easy to find the information I need". Insufficient evidence of a main effect of treatment was found.

Table D-75. Analysis of Variance for Question 52: It is easy to find the information I need.

Source	DF	SS	MS	F	P
Treatment	2	2.400	1.200	1.99	0.157
Error	27	16.300	0.604		
Total	29	18.700			

D.8.30 The information provided by the Usability Evaluation web site is easy to understand.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The information provided by the Usability Evaluation web site is easy to understand". Insufficient evidence of a main effect of treatment was found.

Table D-76. Analysis of Variance for Question 53: The information provided by the Usability Evaluation web site is easy to understand.

Source	DF	SS	MS	F	P
Treatment	2	0.467	0.233	0.97	0.392
Error	27	6.500	0.241		
Total	29	6.967			

D.8.31   The organization of information on the Usability Evaluation web pages is clear.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The organization of information on the Usability Evaluation web pages is clear". Insufficient evidence of a main effect of treatment was found.

Table D-77. Analysis of Variance for Question 54: The organization of information on the Usability Evaluation web pages is clear.

Source	DF	SS	MS	F	P
Treatment	2	1.867	0.933	2.05	0.148
Error	27	12.300	0.456		
Total	29	14.167			

D.8.32   I liked using the Usability Evaluation web site.

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "I liked using the Usability Evaluation web site". Insufficient evidence of a main effect of treatment was found.

Table D-78. Analysis of Variance for Question 55: I liked using the Usability Evaluation web site.

Source	DF	SS	MS	F	P
Treatment	2	1.667	0.833	1.99	0.156
Error	27	11.300	0.419		
Total	29	12.967			

D.8.33 The emailing tasks that I was required to perform were realistic

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "The emailing tasks that I was required to perform were realistic". A significant main effect of Treatment was found at p=0.041.

Table D-79. Analysis of Variance for Question 56: The emailing tasks that I was required to perform were realistic.

Source	DF	SS	MS	F	P
Treatment	2	3.163	1.581	3.61	0.041
Error	26	11.389	0.438		
Total	28	14.552			

A Newman-Keuls post hoc analysis was conducted to isolate the significant differences, the results of which are presented below.

Table D-80. Newman-Keuls Results for "The emailing tasks that I was required to perform were realistic"

	Increasing Order (n=10; MSE=0.438)				
	1	2	3		
Treatment	1	2	3		
Mean	3.9	4.4	4.7	r	CDn-k
1	---	0.5	**0.8**	3	0.7388
2		---	0.3	2	0.6111
3			---		

The results indicate that ratings allocated by the non-reporting laboratory-based users were significantly higher than those allocated by the remote users.

D.8.34 I will continue to use VEMS now that I have completed the study

A one-way ANOVA was carried out to evaluate the effect of treatment group on ratings allocated to the statement "I will continue to use VEMS now that I have completed the study". A significant main effect of Treatment was found at $p < 0.0001$.

Table D-81. Analysis of Variance for Question 57: I will continue to use VEMS now that I have completed the study.

Source	DF	SS	MS	F	P
Treatment	2	15.793	7.896	11.63	0.000
Error	26	17.656	0.679		
Total	28	33.448			

A Newman-Keuls post hoc analysis was conducted to isolate the significant differences, the results of which are presented below.

Table D-82. Newman-Keuls Results for "I will continue to use VEMS now that I have completed the study"

	Increasing Order (n=10; MSE=0.679)				
	1	2	3		
Treatment	1	2	3		
Mean	2.8	4.2	4.5	r	CDn-k
1	---	**1.4**	**1.7**	3	0.9198
2		---	0.3	2	0.7609
3			---		

The results indicate that ratings allocated by the laboratory-based users (reporting and non-reporting) were significantly higher than those allocated by the remote users.

D.9 COMPARISONS BETWEEN LAB/REPORTING AND REMOTE/REPORTING TREATMENT GROUPS ONLY

Analyses of variance were conducted for post-test questionnaire rating questions related to the Critical Incident Report Form to assess the effects of Treatment (between-subject factor; 2 levels) on the mean ratings allocated. In this case, only comparisons amongst the remote/reporting and lab/reporting treatment groups were feasible. Ratings were assigned based on a five-point Likert-scale where 1=strongly disagree and 5=strongly agree.

D.9.1 I like the idea of reporting critical incident information to developers.

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "I like the idea of reporting critical incident information to developers". Insufficient evidence of a main effect of Treatment was found.

Table D-83. Analysis of Variance for Question 32: I like the idea of reporting critical incident information to developers.

Source	DF	SS	MS	F	P
Treatment	1	0.025	0.025	0.05	0.821
Error	12	5.689	0.474		
Total	13	5.714			

D.9.2 It was easy to access the critical incident report form

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "It was easy to access the critical incident report form". Insufficient evidence of a main effect of Treatment was found.

Table D-84. Analysis of Variance for Question 33: It was easy to access the critical incident report form.

Source	DF	SS	MS	F	P
Treatment	1	0.800	0.800	2.32	0.145
Error	18	6.200	0.344		
Total	19	7.000			

D.9.3 It was easy to report critical incidents using the report form

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "It was easy to report critical incidents using the report form". Insufficient evidence of a main effect of Treatment was found.

Table D-85. Analysis of Variance for Question 34: It was easy to report critical incidents using the report form.

Source	DF	SS	MS	F	P
Treatment	1	0.450	0.450	1.00	0.331
Error	18	8.100	0.450		
Total	19	8.550			

D.9.4 The questions on the report form were easy to understand

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "The questions on the report form were easy to understand". Insufficient evidence of a main effect of Treatment was found.

Table D-86. Analysis of Variance for Question 35: The questions on the report form were easy to understand.

Source	DF	SS	MS	F	P
Treatment	1	0.450	0.450	0.95	0.342
Error	18	8.500	0.472		
Total	19	8.950			

D.9.5 The questions on the report form covered sufficient detail concerning the critical incident

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "The questions on the report form covered sufficient detail concerning the critical incident". Insufficient evidence of a main effect of Treatment was found.

Table D-87. Analysis of Variance for Question 36: The questions on the report form covered sufficient detail concerning the critical incident.

Source	DF	SS	MS	F	P
Treatment	1	1.800	1.800	2.70	0.118
Error	18	12.000	0.667		
Total	19	13.800			

D.9.6 I was motivated to report negative critical incidents

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "I was motivated to report negative critical incidents". Insufficient evidence of a main effect of Treatment was found.

Table D-88. Analysis of Variance for Question 37: I was motivated to report negative critical incidents.

Source	DF	SS	MS	F	P
Treatment	1	1.800	1.800	4.05	0.059
Error	18	8.000	0.444		
Total	19	9.800			

D.9.7 I was motivated to report positive critical incidents

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "I was motivated to report positive critical incidents". A significant main effect of Treatment was found at p=0.039.

Table D-89. Analysis of Variance for Question 38: I was motivated to report positive critical incidents.

Source	DF	SS	MS	F	P
Treatment	1	3.200	3.200	4.97	0.039
Error	18	11.600	0.644		
Total	19	14.800			

Table D-90 presents the mean responses per treatment group. Since the Treatment factor has only two levels in this case, comparison of the mean responses is sufficient to determine the significant difference. Therefore, it can be concluded that the lab/reporting treatment group gave a significantly higher rating (indicated stronger agreement) to the statement "I was motivated to report positive critical incidents".

Table D-90. Comparison of Mean Responses Allocated to the Statement "I was motivated to report positive critical incidents" Per Treatment Group

Treatment Group	Mean Response
Remote/Reporting	3.0
Lab/Reporting	3.8

D.9.8 It was easy to describe the task I was performing using the Critical Incident Report Form

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "It was easy to describe the task I was performing using the Critical Incident Report Form". Insufficient evidence of a main effect of Treatment was found.

Table D-91. Analysis of Variance for Question 39: It was easy to describe the task I was performing using the Critical Incident Report Form.

Source	DF	SS	MS	F	P
Treatment	1	0.000	0.000	0.00	1.000
Error	18	11.200	0.622		
Total	19	11.200			

D.9.9   It was easy to describe the critical incident using the Critical Incident Report Form

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "It was easy to describe the critical incident using the Critical Incident Report Form". Insufficient evidence of a main effect of Treatment was found.

Table D-92. Analysis of Variance for Question 40: It was easy to describe the critical incident using the Critical Incident Report Form.

Source	DF	SS	MS	F	P
Treatment	1	0.000	0.000	0.00	1.000
Error	18	6.000	0.333		
Total	19	6.000			

D.9.10 It was easy to rate the impact of the critical incident on task performance.

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "It was easy to rate the impact of the critical incident on task performance". Insufficient evidence of a main effect of Treatment was found.

Table D-93. Analysis of Variance for Question 41: It was easy to rate the impact of the critical incident on task performance.

Source	DF	SS	MS	F	P
Treatment	1	0.200	0.200	0.26	0.618
Error	18	14.000	0.778		
Total	19	14.200			

D.9.11 It was easy to rate the impact of the critical incident on satisfaction.

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "It was easy to rate the impact of the critical incident on satisfaction". Insufficient evidence of a main effect of Treatment was found.

Table D-94. Analysis of Variance for Question 42: It was easy to rate the impact of the critical incident on satisfaction.

Source	DF	SS	MS	F	P
Treatment	1	0.426	0.426	1.19	0.291
Error	17	6.100	0.359		
Total	18	6.526			

D.9.12 It was easy to rate the severity of errors (negative critical incidents only).

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "It was easy to rate the severity of errors (negative critical incidents only)". Insufficient evidence of a main effect of Treatment was found.

Table D-95. Analysis of Variance for Question 43: It was easy to rate the severity of errors (negative critical incidents only).

Source	DF	SS	MS	F	P
Treatment	1	0.000	0.000	0.00	1.000
Error	18	13.200	0.733		
Total	19	13.200			

D.9.13 The rating scales used were appropriate

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "The rating scales used were appropriate". Insufficient evidence of a main effect of Treatment was found.

Table D-96. Analysis of Variance for Question 44: The rating scales used were appropriate.

Source	DF	SS	MS	F	P
Treatment	1	0.200	0.200	0.47	0.500
Error	18	7.600	0.422		
Total	19	7.800			

D.9.14 I consider the critical incident technique an effective way of evaluating an interface

A one-way ANOVA was carried out to evaluate the effect of treatment group (including only the remote/reporting and lab/reporting conditions) on ratings allocated to the statement "I consider the critical incident technique an effective way of evaluating an interface". Insufficient evidence of a main effect of Treatment was found.

Table D-97. Analysis of Variance for Question 45: I consider the critical incident technique an effective way of evaluating an interface.

Source	DF	SS	MS	F	P
Treatment	1	0.001	0.001	0.00	0.975
Error	17	9.789	0.576		
Total	18	9.789			

D.10 COMPARISONS BETWEEN WEB AND VOICE INTERFACES

Analyses of variance were conducted for post-test questionnaire rating questions asked of both the VEMS web interface and voice interface to assess the effects of Interface (within-subject factor; 2 levels) on the mean ratings allocated. Again, ratings were assigned based on a five-point Likert-scale where 1=strongly disagree and 5=strongly agree.

D.10.1 I was able to complete my tasks quickly using the [interface type]

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "I was able to complete my tasks quickly using the [interface type]". A main effect of interface was found at $p<0.0001$.

Table D-98. ANOVA Results for "I was able to complete my tasks quickly using the [interface type]"

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	18.1500	18.1500	18.1500	22.54	0.000
Subject	29	42.6833	42.6833	1.4718	1.83	0.055
Error	29	23.3500	23.3500	0.8052		
Total	59	84.1833				

Table D-99 presents the mean responses per Interface type. Since the Interface factor has only two levels, comparison of the mean responses is sufficient to determine the significant difference. Therefore, it can be concluded that a significantly higher rating to the statement "I was able to complete my tasks quickly using the [interface type]" was allocated for the web interface versus the voice interface.

Table D-99. Comparison of Mean Responses Allocated to the Statement " I was able to complete my tasks quickly using the VEMS web site " Per Interface Type

Mean Responses	
Web Interface	**Voice Interface**
4.2	3.1

D.10.2 I was able to complete my tasks efficiently using [interface type]

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "I was able to complete my tasks efficiently using [interface type]". A main effect of interface was found at $p<0.0001$.

Table D-100. ANOVA Results for "I was able to complete my tasks efficiently using [interface type]"

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	16.0167	16.0167	16.0167	25.13	0.000
Subject	29	34.0833	34.0833	1.1753	1.84	0.052
Error	29	18.4833	18.4833	0.6374		
Total	59	68.5833				

Table D-101 presents the mean responses per Interface type. It can be concluded from this data that a significantly higher rating to the statement "I was able to complete my tasks efficiently using the [interface type]" was allocated for the web interface versus the voice interface.

Table D-101. Comparison of Mean Responses Allocated to the Statement "I was able to complete my tasks efficiently using [interface type]" Per Interface Type

Mean Responses	
Web Interface	**Voice Interface**
4.1	3.1

D.10.3 It was easy to learn how to use the [interface type]

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "It was easy to learn how to use the [interface type]". A main effect of interface was found at $p=0.007$.

Table D-102. ANOVA Results for "It was easy to learn how to use the [interface type]"

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	3.7500	3.7500	3.7500	8.53	0.007
Subject	29	30.7500	30.7500	1.0603	2.41	0.010
Error	29	12.7500	12.7500	0.4397		
Total	59	47.2500				

Table D-103 presents the mean responses per Interface type. It can be concluded from this data that a significantly higher rating to the statement "It was easy to learn how to use the [interface type]" was allocated for the web interface versus the voice interface.

Table D-103. Comparison of Mean Responses Allocated to the Statement "It was easy to learn how to use the [interface type]" Per Interface Type

Mean Responses	
Web Interface	**Voice Interface**
4.5	4.0

D.10.4 The [interface type] gives error messages that clearly tell me how to fix problems

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "The [interface type] gives error messages that clearly tell me how to fix problems". Insufficient evidence of a main effect of Interface was found.

Table D-104. ANOVA Results for "The [interface type] gives error messages that clearly tell me how to fix problems "

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	3.7078	2.0000	2.0000	2.53	0.125
Subject	28	38.1255	38.1255	1.3616	1.72	0.090
Error	24	19.0000	19.0000	0.7917		
Total	53	60.8333				

D.10.5 Whenever I make a mistake using the [interface type], I recover easily and quickly

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "Whenever I make a mistake using the [interface type], I recover easily and quickly". A main effect of interface was found at p=0.005.

Table D-105. ANOVA Results for "Whenever I make a mistake using the [interface type], I recover easily and quickly"

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	7.0838	8.3448	8.3448	9.11	0.005
Subject	29	55.8034	55.8034	1.9243	2.10	0.026
Error	28	25.6552	25.6552	0.9163		
Total	58	88.5424				

Table D-106 presents the mean responses per Interface type. It can be concluded from this data that a significantly higher rating to the statement "Whenever I make a mistake using the [interface type], I recover easily and quickly" was allocated for the web interface versus the voice interface.

Table D-106. Comparison of Mean Responses Allocated to the Statement "Whenever I make a mistake using the [interface type], I recover easily and quickly" Per Interface Type

Mean Responses	
Web Interface	Voice Interface
3.8	3.1

D.10.6 The information provided by the [interface type] is clear

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "The information provided by the [interface type] is clear". Insufficient evidence of a main effect of Interface was found.

Table D-107. ANOVA Results for " The information provided by the [interface type] is clear "

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	1.7651	1.7241	1.7241	2.79	0.106
Subject	29	26.5184	26.5184	0.9144	1.48	0.151
Error	28	17.2759	17.2759	0.6170		
Total	58	45.5593				

D.10.7 It is easy to find information I need using the [interface type]

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "It is easy to find information I need using the [interface type]". A main effect of interface was found at p=0.001.

Table D-108. ANOVA Results for " It is easy to find information I need using the [interface type]"

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	6.0167	6.0167	6.0167	15.19	0.001
Subject	29	29.3500	29.3500	1.0121	2.56	0.007
Error	29	11.4833	11.4833	0.3960		
Total	59	46.8500				

Table D-109 presents the mean responses per Interface type. It can be concluded from this data that a significantly higher rating to the statement "It is easy to find information I need using the [interface type]" was allocated for the web interface versus the voice interface.

Table D-109. Comparison of Mean Responses Allocated to the Statement "It is easy to find information I need using the [interface type]" Per Interface Type

Mean Responses	
Web Interface	Voice Interface
3.9	3.2

D.10.8 The information provided by the [interface type] is easy to understand

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "The information provided by the [interface type] is easy to understand". Insufficient evidence of a main effect of Interface was found.

Table D-110. ANOVA Results for "The information provided by the [interface type] is easy to understand "

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	2.4002	2.0862	2.0862	4.05	0.054
Subject	29	16.5759	16.5759	0.5716	1.11	0.392
Error	28	14.4138	14.4138	0.5148		
Total	58	33.3898				

D.10.9 I like using the [interface type]

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "I like using the [interface type]". A main effect of interface was found at p=0.004.

Table D-111. ANOVA Results for "I like using the [interface type]"

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	7.3500	7.3500	7.3500	9.62	0.004
Subject	29	46.6833	46.6833	1.6098	2.11	0.025
Error	29	22.1500	22.1500	0.7638		
Total	59	76.1833				

Table D-112 presents the mean responses per Interface type. It can be concluded from this data that a significantly higher rating to the statement "I like using the [interface type]" was allocated for the web interface versus the voice interface.

Table D-112. Comparison of Mean Responses Allocated to the Statement "I like using the [interface type]" Per Interface Type

Mean Responses	
Web Interface	Voice Interface
4.1	3.4

D.10.10 The [interface type] has all the functions and capabilities that I expect it to have

A one-way within-subject ANOVA was carried out to evaluate the effect of Interface on ratings allocated to the statement "The [interface type] has all the functions and capabilities that I expect it to have". A main effect of interface was found at p=0.016.

Table D-113. ANOVA Results for "The [interface type] has all the functions and capabilities that I expect it to have"

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Interface	1	4.2667	4.2667	4.2667	6.60	0.016
Subject	29	28.9333	28.9333	0.9977	1.54	0.124
Error	29	18.7333	18.7333	0.6460		
Total	59	51.9333				

Table D-114 presents the mean responses per Interface type. It can be concluded from this data that a significantly higher rating to the statement "The [interface type] has all the functions and capabilities that I expect it to have" was allocated for the web interface versus the voice interface.

Table D-114. Comparison of Mean Responses Allocated to the Statement "The [interface type] has all the functions and capabilities that I expect it to have" Per Interface Type

Mean Responses	
Web Interface	Voice Interface
3.9	3.4

APPENDIX E.   PILOT STUDY TEST FORMS

E.1 INFORMED CONSENT FORMS

E.1.1 Instruction With Exercises Training Group

Informed Consent for Participants

<u>Title of Project</u>: Evaluation of an On-line Critical Incident Technique Training Tool

<u>Investigators</u>:   Jennifer A. Thompson and Robert C. Williges

Purpose of the Research
You are invited to participate in a study that involves experimentation for the purpose of evaluating and improving an on-line training tool for the critical incident technique. The critical incident technique is a method that allows for the collection of observations pertaining to events, called critical incidents, associated with features that lead to extremely good or extremely poor performance. The training tool is comprised of two modules, each of which pertains to a different component of the critical incident technique.   The objective of the training tool is to improve your ability to successfully apply the critical incident technique to the evaluation of an interface.

Procedures
In this study, you will be asked to go through the training tool, which involves reading instructional material pertaining to the critical incident technique and performing a series of hands-on exercises. The training tool is divided into two modules, corresponding to the two major components of the technique. Each module has its own set of instructions and exercises. You will go through each module in sequence, with the option of reviewing the material or taking a break at any time. At the end of the training session, you will be given a short break, after which time you will be asked to evaluate an interface using the critical incident technique. The test session will finish with a questionnaire regarding your interaction with the training tool. The entire study will last approximately two hours.

Risks
There are no known risks to the participants of this study. The tasks specified by the experimenter are not tiring, but during the training and test sessions, you are welcome to take rest breaks as needed.

Benefits of this Research
Your participation in this project will provide information that may be used to improve the on-line critical incident training tool. No guarantee of benefits has been made to encourage you to participate. You may receive a synopsis summarizing this research when completed. Please leave a self-addressed envelope with the experimenter if you wish a copy of the results to be sent to you.

Extent of Anonymity and Confidentiality
The results of this study will be kept strictly confidential. Your written consent is required for the researchers to release any data identified with you as an individual to anyone other than personnel working on the project. The information you provide will have your name removed and only a subject number will identify you during analyses and any written reports of the research.

The screen actions of your usage will be videotaped. These tapes will be stored securely, viewed only by the experimenters (Jennifer Thompson, Dr. Williges, Dr. H. Rex Hartson, and Dr. Kleiner), and erased after 3 months. If the experimenters wish to use a portion of your videotape for any

other purpose, they will get your written permission before using it. Your signature on this form does not give them permission to show your videotape to anyone else.

Compensation
You will receive $10 per hour that you participate in the study. The test session will last for approximately 2 hours. Payment will be made immediately after you have finished your participation.

Freedom to Withdraw
You are free to withdraw from this study at any time for any reason. You will be compensated for your participation up to the point of withdrawal.

Approval of Research
This research has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University, and by the Department of Industrial and Systems Engineering.

Subjects Responsibilities
I voluntarily agree to participate in this study. I have the following responsibilities:
4. I should not volunteer for participation if I now know I will not be able to complete this experiment.
5. After completion of this study, I will not discuss my experiences with any other individual for a period of two months. This will ensure that everyone will begin the study with the same level of knowledge and expectations.

Permission
I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

_____   _____

Signature   Date

Should I have any questions about this research or its conduct, I may contact:

Jennifer A. Thompson	Investigator	552-3729
Robert C. Williges	Investigator	231-6270
H. T. Hurd	Chair, IRB Research Division	231-5281

Informed Consent for Participants

Title of Project: Evaluation of an On-line Critical Incident Technique Training Tool

Investigators: Jennifer A. Thompson and Robert C. Williges

Purpose of the Research
You are invited to participate in a study that involves experimentation for the purpose of evaluating and improving an on-line training tool for the critical incident technique. The critical incident technique is a method that allows for the collection of observations pertaining to events, called critical incidents, associated with features that lead to extremely good or extremely poor performance. The training tool is comprised of two modules, each of which pertains to a different component of the critical incident technique.   The objective of the training tool is to improve your ability to successfully apply the critical incident technique to the evaluation of an interface.

Procedures
In this study, you will be asked to go through the training tool, which involves reading instructional material pertaining to the critical incident technique. The training tool is divided into two modules, corresponding to the two major components of the technique. You will go through each module in sequence, with the option of reviewing the material or taking a break at any time. At the end of the training session, you will be given a short break, after which time you will be asked to evaluate an interface using the critical incident technique. The test session will finish with a questionnaire regarding your interaction with the training tool. The entire study will last approximately two hours.

Risks
There are no known risks to the participants of this study. The tasks specified by the experimenter are not tiring, but during the training and test sessions, you are welcome to take rest breaks as needed.

Benefits of this Research
Your participation in this project will provide information that may be used to improve the on-line critical incident training tool. No guarantee of benefits has been made to encourage you to participate. You may receive a synopsis summarizing this research when completed. Please leave a self-addressed envelope with the experimenter if you wish a copy of the results to be sent to you.

Extent of Anonymity and Confidentiality
The results of this study will be kept strictly confidential. Your written consent is required for the researchers to release any data identified with you as an individual to anyone other than personnel working on the project. The information you provide will have your name removed and only a subject number will identify you during analyses and any written reports of the research.

The screen actions of your usage will be videotaped. These tapes will be stored securely, viewed only by the experimenters (Jennifer Thompson, Dr. Williges, Dr. H. Rex Hartson, and Dr. Kleiner), and erased after 3 months. If the experimenters wish to use a portion of your videotape for any other purpose, they will get your written permission before using it. Your signature on this form does <u>not</u> give them permission to show your videotape to anyone else.

Compensation
You will receive $10 per hour that you participate in the study. The test session will last for approximately 2 hours. Payment will be made immediately after you have finished your participation.

Freedom to Withdraw
You are free to withdraw from this study <u>at any time</u> for any reason. You will be compensated for your participation up to the point of withdrawal.

Approval of Research
This research has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University, and by the Department of Industrial and Systems Engineering.

Subjects Responsibilities
I voluntarily agree to participate in this study. I have the following responsibilities:
6. I should not volunteer for participation; if I now know I will not be able to complete this experiment.
7. After completion of this study, I will not discuss my experiences with any other individual for a period of two months. This will ensure that everyone will begin the study with the same level of knowledge and expectations.

Permission

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project.

_____   _____
Signature   Date

Should I have any questions about this research or its conduct, I may contact:

Jennifer A. Thompson	Investigator	552-3729
Robert C. Williges	Investigator	231-6270
H. T. Hurd	Chair, IRB Research Division	231-5281

E.3   PRE-TEST QUESTIONNAIRE FORM

Thank you for participating in the evaluation of the On-line Critical Incident Reporting Tool.
Please answer the following questions.

1. Age (years): _____   Sex: ❍ Male   ❍ Female

2. Academic Level: ❍ Freshman   ❍ Sophomore   ❍ Junior   ❍
   Senior   ❍ Masters   ❍ PhD

3. Major of Study: _____

4. How many university-level courses have you taken that have addressed human factors
   evaluation methodologies, usability evaluation, or human-computer interaction? *(please
   check one)*:
   ❍ none
   ❍ 1 - 2
   ❍ 3- 5
   ❍ 5+

5. How many usability evaluations have you conducted in the past (ex. as a participant, for
   a course project, for research purposes, in industry)? *(please check one)*:
   ❍ none
   ❍ 1 - 2
   ❍ 3- 5
   ❍ 6-10
   ❍ 10+

6. For how long have you been using computers *(please check one)*:
   ❍ less than 6 months
   ❍ between 6 months and a year
   ❍ 1 - 3 years
   ❍ 3 years or more

7. Please indicate which WWW browsers you have used *(please check all that apply)*:
   ❍   Netscape Navigator
   ❍   Microsoft Explorer
   ❍   AOL
   ❍   Other (*specify*): _____

   Of these browsers, which do you use most frequently: _____

8. Rate your level of expertise regarding the use of WWW browsers:

❍ Very experienced   ❍ Moderate experience   ❍ Minimal experience   ❍No experience

Rate the following statements by checking the most appropriate response.

9. I am very familiar with the critical incident technique.
   - ❍ 1 Strongly Agree
   - ❍ 2 Agree
   - ❍ 3 Neutral
   - ❍ 4 Disagree
   - ❍ 5 Strongly Disagree

10. I am very experienced at applying the critical incident technique.
    - ❍ 1 Strongly Agree
    - ❍ 2 Agree
    - ❍ 3 Neutral
    - ❍ 4 Disagree
    - ❍ 5 Strongly Disagree

11. Have you used ever used an on-line instructional tool (ex. training tool, on-line course)?
    - ❍ Yes ❍ No

If yes, please describe this training tool:

_____

_____

_____

_____

End of Questionnaire. Thank you.

E.4 POST-TEST QUESTIONNAIRE FORMS

E.4.1 Instruction with Exercises Training Condition

Please indicate how strongly you disagree or agree to the statements using the number scale provided. Try to respond to all the items. If an item is not applicable, use **NA**.

Critical Incident Training
The following questions pertain to your use of the on-line training tool.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
It was simple to use the training tool.	O	O	O	O	O	O
The training was easy to follow.	O	O	O	O	O	O
The training tool provided sufficient information.	O	O	O	O	O	O
The information was easy to understand.	O	O	O	O	O	O
The organization of information was clear.	O	O	O	O	O	O
I liked interacting with the training tool.	O	O	O	O	O	O
The training helped me learn to **identify positive** critical incidents.	O	O	O	O	O	O
The training helped me learn to **identify negative** critical incidents.	O	O	O	O	O	O
The training helped me learn to **report positive** critical incidents.	O	O	O	O	O	O
The training helped me learn to **report negative** critical incidents.	O	O	O	O	O	O
The exercise problems were useful.	O	O	O	O	O	O
The exercise problems were at an appropriate level of difficulty.	O	O	O	O	O	O
The exercise problems helped clarify the written material.	O	O	O	O	O	O
The number of exercise problems was sufficient.	O	O	O	O	O	O
The material covered by the exercise problems was sufficient.	O	O	O	O	O	O
I feel better prepared to **identify** critical incidents after going through the training tool.	O	O	O	O	O	O
I feel better prepared to **report** critical incidents after going through the training tool.	O	O	O	O	O	O

List the two most **negative** aspect(s) of the On-line Critical Incident Training Tool:

1. _____

2. _____

List the two most **positive** aspect(s) of the On-line Critical Incident Training Tool:

1. _____

2. _____

Critical Incident Reporting
The following questions pertain to your role as a critical incident identifier and reporter.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
I like the idea of reporting critical incident information to developers.	O	O	O	O	O	O
It was easy to access the critical incident report form.	O	O	O	O	O	O
It was easy to report critical incidents using the report form.	O	O	O	O	O	O
The questions on the report form were easy to understand.	O	O	O	O	O	O
The questions on the report form covered sufficient detail concerning the critical incident.	O	O	O	O	O	O
I was motivated to report negative critical incidents.	O	O	O	O	O	O
I was motivated to report positive critical incidents.	O	O	O	O	O	O
It was easy to describe the task I was performing using the Critical Incident Report Form.	O	O	O	O	O	O
It was easy to describe the critical incident using the Critical Incident Report Form.	O	O	O	O	O	O
It was easy to rate the impact of the critical incidents on task performance.	O	O	O	O	O	O
It was easy to rate the impact of the critical incident on satisfaction.	O	O	O	O	O	O
It was easy to rate the severity of errors (negative critical incidents only).	O	O	O	O	O	O
The rating scales used were appropriate.	O	O	O	O	O	O

I prefer to report critical incidents:
❐   immediately when I encounter them   ❐   after completing my task

Comments: _____
_____
_____
_____
_____

End of questionnaire. Thank you for your help.

E.4.2 Instruction Only Training Condition

Please indicate how strongly you disagree or agree to the statements using the number scale provided. Try to respond to all the items. If an item is not applicable, use **NA**.

Critical Incident Training
The following questions pertain to your use of the on-line training tool.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
It was simple to use the training tool.	○	○	○	○	○	○
The training was easy to follow.	○	○	○	○	○	○
The training tool provided sufficient information.	○	○	○	○	○	○
The information was easy to understand.	○	○	○	○	○	○
The organization of information was clear.	○	○	○	○	○	○
I liked interacting with the training tool.	○	○	○	○	○	○
The training helped me learn to **identify positive** critical incidents.	○	○	○	○	○	○
The training helped me learn to **identify negative** critical incidents.	○	○	○	○	○	○
The training helped me learn to **report positive** critical incidents.	○	○	○	○	○	○
The training helped me learn to **report negative** critical incidents.	○	○	○	○	○	○
I feel better prepared to **identify** critical incidents after going through the training tool.	○	○	○	○	○	○
I feel better prepared to **report** critical incidents after going through the training tool.	○	○	○	○	○	○

List the two most **negative** aspect(s) of the On-line Critical Incident Training Tool:

1. _____
_____

2 _____
_____

List the two most **positive** aspect(s) of the On-line Critical Incident Training Tool:

1 _____
_____

2 _____
_____

Critical Incident Reporting

The following questions pertain to your role as a critical incident identifier and reporter.

	strongly disagree	disagree	neutral	agree	strongly agree	NA
I like the idea of reporting critical incident information to developers.	O	O	O	O	O	O
It was easy to access the critical incident report form.	O	O	O	O	O	O
It was easy to report critical incidents using the report form.	O	O	O	O	O	O
The questions on the report form were easy to understand.	O	O	O	O	O	O
The questions on the report form covered sufficient detail concerning the critical incident.	O	O	O	O	O	O
I was motivated to report negative critical incidents.	O	O	O	O	O	O
I was motivated to report positive critical incidents.	O	O	O	O	O	O
It was easy to describe the task I was performing using the Critical Incident Report Form.	O	O	O	O	O	O
It was easy to describe the critical incident using the Critical Incident Report Form.	O	O	O	O	O	O
It was easy to rate the impact of the critical incidents on task performance.	O	O	O	O	O	O
It was easy to rate the impact of the critical incident on satisfaction.	O	O	O	O	O	O
It was easy to rate the severity of errors (negative critical incidents only).	O	O	O	O	O	O
The rating scales used were appropriate.	O	O	O	O	O	O

I prefer to report critical incidents:

❒   immediately when I encounter them   ❒   after completing my task

Comments: _____
_____
_____
_____
_____

End of questionnaire. Thank you for your help.

E.5 DATA COLLECTION OBSERVATION SHEET

TIME LOG – TRAINING + EVALUATION

	Time (s)
Training Intro Start	
Module 1 Start	
Exercise 1 Start	
Module 2 Start	
Exercise 2 Start	
Conclusion Start	
Conclusion End	
Break Start	
Evaluation Start	
Evaluation End	

CRITICAL INCIDENT REPORT SUMMARY

Critical Incident	Type (P or N)	Start Time	End Time	Description	Help Accessed?
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					