The Usefulness of Method-Resources for Evaluating a Collaborative Training Simulator

Ebba Thora Hvannberg, Gyda Halldorsdottir, Jan Rudinsky

University of Iceland, Dunhaga 5, 107 Reykjavík, Iceland
ebba@hi.is, gyda@hi.is, janr@hi.is

Abstract. Voice communication is vital for collaboration between first responders and commanders during crisis management. To decrease cost, training can take place in a virtual environment instead of in a real one. It is non-trivial to build and evaluate a virtual environment for training complex command. To understand the method-resources required for evaluating a training simulator for crisis response, this paper presents a case study of applying several resources. Method-resources were analysed for usability problems and Mechanics of Collaboration (MOC). The results show that the Group Observational Technique and the MOC analysis are appropriate for analysing factors of collaboration and communication. The think-aloud technique, observers, experts in the domain and advanced task scenario were important resources. In only a few cases sound and video were necessary to analyse issues.

Keywords: Virtual Reality; Collaboration; Evaluation; Crisis management; Verbal Communication; Method-Resources; Mechanics of Collaboration.

1 Introduction

Technological developments of collaborative systems have increased demands on resources for evaluating virtual environments that allow multiple modes such as visualization, verbal communication and sound. Evaluating usability of interactive systems designed for Computer Supported Cooperative Work (CSCW) has been found challenging. Because of collaborating users, the evaluation requires more observers than in a single user scenario [1]. Increasingly, sound plays a large role in virtual environments in the form of natural or synthetic sound, music or voice communication [2]. Evaluation of the effect of sound on voice communication has been carried out but mostly in controlled laboratory settings [3].

To gain valid results of a usability evaluation, method selection continues to be a critical activity. The increased complexity of systems has called for ways to choose methods that are most appropriate for usability evaluations. One such framework has been proposed by Antunes et al. [4] who devised a selection strategy for choosing an appropriate method of evaluation. Unfortunately, empirical research on usability eval-

Hvannberg E.T., Halldorsdottir G., Rudinsky J. (2015) The Usefulness of Method-Resources for Evaluating a Collaborative Training Simulator. In: Abascal J., Barbosa S., Fetter M., Gross T., Palanque P., Winckler M. (eds) Human-Computer Interaction – INTERACT 2015. INTERACT 2015. Lecture Notes in Computer Science, vol 9297. Springer, Cham

uation of collaborative virtual environment has not been extensive [4]. While evaluations of collaborative environments have been mostly studied theoretically [5], studies have been reported on the effectiveness of 3D collaborative virtual environments [6]. Recognizing that multiple dimensions of interactive systems call for more than an adoption of one or two evaluation methods, it has been suggested that a collection of resources, that are a part of a whole method, return an efficient evaluation [7]. Examples of resources of methods are participant recruitment, task scenario, reporting format, problem identification, problem classification and a thinking-aloud protocol [7]. This paper uses the term method-resources to describe such resources [8].

Thus motivated, the main objective of this research is to gain empirical knowledge on the suitability of resources for usability evaluation of a collaborative system, set in a virtual environment, where players communicate verbally in a noisy environment. To achieve this objective, we studied a usability evaluation of a prototype virtual environment for training crisis management personnel dealing with mass-casualty accidents.

Crisis management training is organized and developed according to an accurate predefined system for immediate responses to mass-casualty incidents. First responders are personnel from volunteers to professionals with a few to many years of experience. Professional responders come from different organizations, e.g., rescue, police, medical and firefighting. Training these people to obtain efficient skills is crucial.

Figure 1 A snapshot of the virtual environment showing the scene of the accident.

The application (see Figure 1) that was evaluated is a virtual training simulator for crisis response to a mass-casualty incident at an airport. Trainees are commanders, the On Scene Commander (OSC) managing the responses overall on the scene and the Rescue Coordinator (RC) managing the rescue work on the scene that exercise progress report, rescue resources, information requests and task delegation. They can navigate with an on scene view and a zooming in effect through the virtual environment that has a changing high-fidelity scene. They are wearing earphones with a microphone. A design based on empirical data consisted of three voice communication metaphors reflecting the communication spaces used in a real crisis event and its training, i.e. one way radio, mobile phone and face-to-face (F2F) communication for the two persons to speak to one another [9]. To broadcast a message a GUI button is pressed with a mouse (i.e. Push-to-talk) and in response, a transfer is reversed from receiving to sending with the player microphone no longer muted. The prototype has a high fidelity soundscape with sound from fire trucks, other resources, players talking, fire and wind. Any type of

sound is relayed over communication channels, i.e. players can hear noises at the far end through the channel.

2 Selection of method-resources

2.1 Evaluation methods

To select an evaluation method we applied the CSCW evaluation framework by Antunes et al. [4], where a variety of evaluation methods are presented for each stage of the software development. We selected Groupware Observational User Testing (GOT) as an end-user oriented evaluation method focusing on realism and usability as the main objective [4, 10]. Gutwin and Greenberg [10] proposed GOT, as a cost effective usability method, which was based on a set of fundamental Mechanics Of Collaboration (MOC). The GOT technique is an observational user testing method focusing on usability in a planned situation, collaboration where users perform predefined tasks. The framework of MOC includes seven categories of important collaboration activities: Communication as Explicit communication and Consequential communication (information unintentionally given off by others); Coordination of action; Planning; Monitoring and gathering information in the workspace; Assistance to one another; and Protection of resources in the workspace [10, 11]. The MOC model has been evolving [12], but the original set was more appropriate for our study. Besides MOC, the method-resources needed for GOT are think-aloud, observers, users and tasks. Furthermore, to record observations, screen-captures were used and audio recorded of the communication in the virtual environment and of the think-aloud. The tasks and the users are described in the next section.

2.2 Collaborative scenario and users

Based on extensive observations of crisis management training on-site, interviews and workshops, a collaborative scenario comprising several tasks was written and validated by an experienced crisis management instructor. The scenario aimed to secure a situation at an accident scene and allowing commanders to ask for resources, such as fire fighters, using verbal communication (see Table 1). Each commander was located in a separate room in front of a screen wearing a head-set with a microphone.

Six employees of a rescue and fire organisation at an airport with experience in crisis response were recruited as participants for the study and divided into three pairs of OSC and RC. Participants had all received a one day introduction to training in a virtual environment, but not to the prototype used in this study. The same session was repeated three times, once for each pair of collaborators which were followed by an observer.

2.3 Analysis of usability problems and protocol analysis according to MOC

The data was analysed in two ways, analysing usability issues and collaborations using MOC. A bottom up qualitative analysis was performed identifying usability issues that were consolidated into unique usability problems. Before analysing the data, the second author transcribed the audio data into text while listening to it and observing the video capture. Comments that observers (the first and third authors) had

Table 1. Tasks in a collaborative scenario for OSC and RC.

On Scene Commander tasks	Rescue Coordinator tasks
1. When receiving a mobile call with incident details, please note down the information.	
2. Set up the emergency channel on your radio device by configuring to 116Hz frequency.	**1.** Set up the emergency channel on your radio by configuring it to 116 Hz frequency.
3. You must navigate to the gate. If a person later appears at the gate you will ask him for a name and register the name by writing it down on paper.	
4. You can use radio or mobile phone to contact the RC. Inform RC of the incident details that you received previously and ask him or her to report at the gate.	**2.** Wait for a message from OSC on radio or mobile phone. Once asked by the OSC you will go to the gate.
	3. When you reach the gate you will notify the OSC of your presence and state your name.
	4. Go to the crashed plane and count casualties.
	5. Give the number of casualties to the OSC using radio.
5. You will coordinate the rescue operation with RC. If you receive important information you should write it down. If you receive a call from the RC asking for additional resources, locate the required resources around and send them to the scene.	**6.** Observe the development of fire and listen to messages. If the fire spreads covering most of the plane, contact OSC for additional resources.
6. Once the fire is out, you can go to the scene. If you are contacted by RC asking about casualties' placement, you suggest a location that is near the plane, but safe from fire, smoke and flying debris and not in the line between the scene and gate.	**7.** If you receive a report by the fire fighter's team leader that the fire has been extinguished, you must inform OSC that it is now safe to enter the scene. OSC can be contacted over the radio or F2F if he or she is around.
	8. Discuss casualty placement with OSC over radio or F2F and make an agreement on possible locations.

written down during the sessions were integrated to the transcript protocol. The third author analysed the protocol for problems that participants faced that were then verified by the first author. In addition to problems, an observer looked for successful interactions, activities and comments raised by the participants. After analysing the protocol, the observations were categorised into groups emerging from the data. Reviewing the transcribed conversations and the videos of the screen capture, the second author and an engineer analysed the data using MOC. After analysing them independently, they discussed differences and came to a consensus. Additional method-resources are observers with expertise in human-computer interaction and moderate expertise in the domain of crisis management.

3 Results

The experience gained from applying the method-resources will help analyse their usefulness. The sessions lasted 17–20 minutes each, which gave 109 minutes of transcribed audio and video recordings. The number of conversations per pair was 10-12, or a total of 32 for all sessions. Observations, capturing problems or successful interactions, e.g. successful training, were 112.

The data analysis uncovered 13 unique usability problems of 84 problem instances divided into eight categories, communication, information/communication, collaboration, navigation, discrepancies between virtual and real world, sound, wrong or inappropriate tasks and following scripts. The first three problem categories of communication and collaboration are discussed in this paper.

Altogether 71 collaboration instances from the conversations of the three pairs of participants were analysed with respect to the MOC categories (see Figure 2). It is noteworthy that the players seemed to monitor their environment and the situation extensively, learning where people are and what they are doing. That no assistance took place could be attributed to low complexity of the scenario and that the players saw no need to protect their workspaces indicates that there were no threats imposed on them in the environment. To analyse the collaboration protocol according to MOC an exact transcript of the collaboration was required.

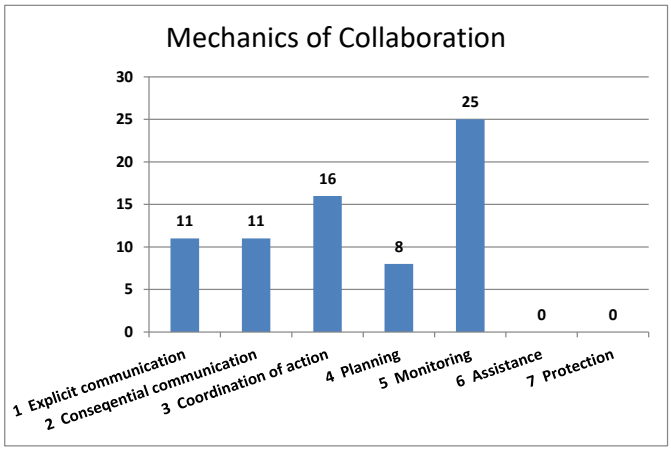

Figure 2 Collaborations analysed using MOC.

The method-resources proved to be helpful for uncovering issues and usability problems of categories addressed here. Table 2 shows examples of how the method-resources were used to uncover particular issues. The issues in the table are labelled with issue IDs (e.g. A01) which are used in the following text and the numbers of the MOC categories are identified according to Figure 2.

To show how the different method-resources have helped, a few examples of the collaboration protocol are given. In several cases, we noticed that the trainees were able to practice skills successfully. One of them is a coordination mechanic (A01):

OSC: [Thinking aloud] I will meet him. The wind actually is like that. I will call. Oops. I will go and meet him like, somewhere here. Not too close to the fire though. I would meet him around here.

The Think-aloud protocol helped the observer to gain insight into what the user is thinking and discovered that he was practicing skills successfully. Such a scenario could not have been practiced without expert users in the domain.

Table 2 Examples of use of method-resources.

Issue ID	Description	Expert in domain	Task scenario	Thinking aloud protocol	Voice	Sound	Video	Observation	Notes by observers	Transcripts of collaboration	Problem identification	Problem classific	MOC
A01	Thinking aloud; Checking conditions	X	X	X	X					X	X	X	3
A02	Thinking aloud; Checking ones role in the experiment;	X	X	X	X			X	X	X	X	X	3
A03	Acknowledging information wrongly.	X	X		X	X	X	X	X	X	X	X	1
A04	OSC is not aware of that RC can hear him. Radio is up on the screen, but it is not lighted as put on for talking			X	X		X			X	X	X	2
A05	Needing some guidance to navigate			X	X			X		X	X	X	3
A06	Does not realize that he is located at wrong place			X	X			X	X	X	X	X	5
A07	Mismatch in prototype vs. real operation	X	X	X	X		X	X	X	X	X	X	5
A08	RC can hear everything OSC is saying, altough OSC is not aware of			X	X			X	X	X	X	X	5
A09	Coordinating own action. Needing guidance to call.	X	X	X	X			X		X	X	X	5

In two of the three experiments, it took participants some time to realise that they should collaborate and take turns, but after a short while they got used to it (A02). Four usability problem instances were observed affecting two participants:

OSC: [Thinking-aloud] So, Do I then… – do I wait for him to call a backup team or help. Or do I? Is that my decision, or does he...

The current method-resources make it difficult to conclude the nature of this problem and the cause of the confusion. It could be that the commander is unsure of the collaborative scenario, that the virtual environment does not provide adequate affordance, or even that he has not been trained adequately in his role. An additional method-resource would be needed to inquire about certain critical points.

Some problems were about providing information verbally. Users relayed wrong information to their partners, or missed to respond when being addressed, either altogether or not responding accurately (A03):

OSC: "RC we have a plane crash on runway 19 the intersection. Plane is on fire. We have 35 people on board. It is a mini-jumbo jet. You go on scene with your team."

RC: "RC got that. A plane crashed down on fire, an intersection 11. No 19 and 01. And 38 people on board" "What type of aircraft is that?"

The participant acknowledged inaccurately and wrongly, or he did not hear correctly the number of people on board, i.e., 38 instead of 35. The video was checked to confirm conditions that showed good sound and no noise disturbing. Five participants had such a problem in a total of 15 instances. It is not analysed as a usability problem and it may

even reflect an accurate picture of a normal training scenario. Another thing we noticed is that the expert users in the domain were able to play their role and act out from the given scenario (*"What type of aircraft is that?")*, thus indicating that the virtual environment is a useful training tool. The MOC analysis proved useful in separating the explicit communication issues from others. The advanced collaborative scenario and expert users in the domain are vital method-resources to create a dialogue and reveal such a scenario.

It was observed that OSC talked while navigating to the gate in the virtual environment and kept the radio channel open, allowing RC to hear what he said (A04, A05, A06, and A07). RC told the observer that he could hear everything OSC said (A08). Later, OSC resolved how to use the radio and called his partner successfully by identifying himself and the receiver (A09). We noticed that participants used the phone much less but encountered similar problems. Two of the three pairs tried to talk F2F, one of them took a few minutes to make it work smoothly but a second pair used it without problems. Expertise of the users in using communication devices in crisis management, e.g. radios, was crucial to understand how they used the communication metaphors in the virtual environment. Observers are expensive resources for evaluating how much certain features are used and could be replaced with automatic monitors. Finally, we see in this example, when noting that the OSC talked on his way to the gate, that it is essential to have a screen-capture of the experiment.

The GOT method and its resources worked well for focusing on usability evaluation in a planned situation, especially when focusing on collaboration and users performing particular predefined tasks. The factors of collaboration and communication were the focus of this study and using the MOC analysis of collaborations fits well for them. The examples show that observation, transcripts and the think-aloud technique are fundamental in researching verbal communication in a collaborative environment. Experts in the domain and the task scenarios are used to uncover fewer issues, but are nonetheless essential. The captured voice is used to produce exact transcript of the collaboration, but as Table 2 shows the sound and the video is used less to understand the issues.

4 Conclusion

The contribution of the paper is twofold. First, we have shown a method for analysing the usefulness of method-resources. Such a method can be useful for other researchers analysing method-resources. Second, its application shows that the GOT method and the MOC analysis are appropriate for analysing factors of collaboration and communication. An essential part of that is to include the think-aloud technique and observers. In only a few cases sound and video were necessary to analyse issues. Other resources that were especially important were expert users in the domain and advanced task scenario.

A few ideas emerged for decreasing the cost of method-resources and raising their effectiveness. Developing software tools for monitoring the scenario, e.g. the frequency of use of features could help decrease the expenses of observation in the collaborative scenario and data analysis. In an environment where the dialogue is rich, the domain is

complex and user domain expertise is high, it may be more difficult than ever to understand the causes of users' actions. A method-resource to analyse critical points of understanding causes may be needed.

References

1. Grudin, J.: Why CSCW applications fail: problems in the design and evaluation of organizational interfaces. Proceedings of the 1988 ACM conference on Computer-supported cooperative work, pp. 85-93. ACM, Portland, Oregon, United States (1988)
2. Divjak, M., Kore, D.: Visual and audio communication between visitors of virtual worlds. Advances in Neural Networks and Applications 41-46 (2001)
3. MacDonald, J.A., Balakrishnan, J., Orosz, M.D., Karplus, W.J.: Intelligibility of speech in a virtual 3-D environment. Human Factors: The Journal of the Human Factors and Ergonomics Society 44, 272-286 (2002)
4. Antunes, P., Herskovic, V., Ochoa, S.F., Pino, J.A.: Structuring dimensions for collaborative systems evaluation. ACM Comput. Surv. 44, 1-28 (2012)
5. de Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G., Poulovassilis, A.: Developing an Evaluation Methodology for Immersive Learning Experiences in a Virtual World. In: Games and Virtual Worlds for Serious Applications, 2009. VS-GAMES '09. Conference in, pp. 43-50. (Year)
6. Montoya, M.M., Massey, A.P., Lockwood, N.S.: 3D Collaborative Virtual Environments: Exploring the Link between Collaborative Behaviors and Team Performance. Decision Sciences 42, 451-476 (2011)
7. Woolrych, A., Hornbæk, K., Frøkjær, E., Cockton, G.: Ingredients and meals rather than recipes: a proposal for research that does not treat usability evaluation methods as indivisible wholes. International Journal of Human-Computer Interaction 27, 940-970 (2011)
8. Law, E.L.-C., Hvannberg, E.T., Vermeeren, A.P., Cockton, G., Jokela, T.: Made for sharing: HCI stories of transfer, triumph and tragedy. In: CHI'13 Extended Abstracts on Human Factors in Computing Systems, pp. 3235-3238. ACM, (Year)
9. Rudinsky, J., Hvannberg, E.T., Helgason, A.A., Petursson, P.B.: Designing soundscapes of virtual environments for crisis management training. Proceedings of the Designing Interactive Systems Conference, pp. 689-692. ACM, Newcastle Upon Tyne, United Kingdom (2012)
10. Gutwin, C., Greenberg, S.: The mechanics of collaboration: developing low cost usability evaluation methods for shared workspaces. In: Enabling Technologies: Infrastructure for Collaborative Enterprises, 2000. (WET ICE 2000). Proeedings. IEEE 9th International Workshops on, pp. 98-103. (Year)
11. Pinelle, D., Gutwin, C.: Evaluating teamwork support in tabletop groupware applications using collaboration usability analysis. Personal and Ubiquitous Computing 12, 237-254 (2008)
12. Pinelle, D., Gutwin, C., Greenberg, S.: Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the mechanics of collaboration. ACM Trans. Comput.-Hum. Interact. 10, 281-311 (2003)