Evaluation of crisis management operations using Reconstruction and Exploration

Dennis Andersson
Swedish Defense Research Agency, FOI
dennis.andersson@foi.se

Sofie Pilemalm
Swedish Defense Research Agency, FOI
sofie.pilemalm@foi.se

Niklas Hallberg
Swedish Defense Research Agency, FOI
niklas.hallberg@foi.se

ABSTRACT

In this paper we present the Reconstruction and Exploration approach (R&E) and F-REX tool and their applications in a field exercise with the Swedish Rescue Services Agency with the purpose of investigating features needed for a computer supported approach for evaluation of large scale crisis management operations. After the exercise several interviews and one seminar were held to evaluate R&E as a representative for computer supported evaluation approaches for crisis management operations. Initial results indicate that multimedia presentation of key events from an operation can be very valuable not only to stimulate the participants to reflect on their own performance, but also to document and share lessons learned to non-participants.

Keywords

After Action Review, Reconstruction and Exploration, debriefing, computer tools for training and evaluation, crisis management, Distributed Tactical Operations

INTRODUCTION

In our society there is an ever increasing need to effectively respond to threats, emergencies and crises of different kinds, including military and terror attacks and natural disasters (Thomas and Kopczak, 2005). Emergency response is a complex activity requiring people from several organizations to together handle situations that are difficult to fully predict in advance. Hence, it is important that first responders are provided the possibility to practice different situations together. A key part in training and education of first responders to such emergencies is assessment and evaluation of team performance in exercises and live operations. The outcome of the assessment and evaluation is crucial as a basis for further training and development of response strategies. A fundamental step enabling analysis and evaluation of the emergency event is data collection. In the simplest of cases data collection is a rather straightforward process. However, as the exercise or operations grows in size regarding the number of participants, e.g. in Distributed Tactical Operations (DTOs), it is non-trivial to determine what data is needed and how to collect it (Albinsson, Wong, Pilemalm and Morin, 2005). Following the data collection is data analysis which also has different levels of complexity, depending on how well the collected data corresponds to the analysis questions being asked. A common approach to evaluating large and complex exercises is to rely on quantitative measures and self-reported questionnaires. Such an approach will definitely reach a large group of respondents but measures tend to focus on assessing how the respondents experienced the situation and whether the operation managed to complete pre-defined objectives. The drawback of this type of assessment is that it will not allow respondents to comment on what actually happened during the operation or to try to explore breakdowns and determine needs for improvement in the organization.

Qualitative research methods, e.g. cognitive analysis approaches (e.g. Wong and Blandford, 2004) are due to their character comparatively time consuming and resource demanding. Also, they do tend to focus on particular incidents based on the experience of a handful individuals rather than taking the whole operation into account. Therefore traditional qualitative research methods become insufficient in large-scale contexts where the course of events is unfolding and where contextual data and an overall picture based all actors, activities and collaborations involved is crucial (Morin and Albinsson, 2005). In conclusion, both quantitative and qualitative research methods, in their own

respect, seem to capture only fragments of large-scale complex events. They thus need to be combined and efficiently supported by proper computer technology.

Study aim and objectives

The overall aim of this study is to explore features needed in computer supported approaches for evaluation of large scale crisis management operations. The specific objectives of the study are:

- To present F-REX, a set of tools for multimedia supported After Action Reviews (AAR) through Reconstruction and Exploration (R&E) of large crisis management operations.
- Initial evaluation of using the R&E approach and the F-REX tools in a case study from an exercise at the Swedish Rescue Services Agency's (SRSA) training facility in Sandö 2006.

R&E explicitly focuses on rich representations of DTOs such as crisis management operations (Rankin, Gentner & Crissey, 1995; Morrison & Meliza, 1999) and seems highly suitable for analyzing such operations through time-synchronized playback of the complex chain of events in an event-driven multimedia model (Morin, 2002). In this study, R&E and the F-REX tools are chosen as representatives for computer supported evaluation and multimedia tools to be applied in the crisis management context.

BACKGROUND

An AAR is an assessment conducted after a major activity that allows participants and commanders to discover what happened and why. The aim of the process is to correct deficiencies, sustain strengths and focus on performance of essential task objectives. Key is the spirit in which AARs are given. The environment and climate surrounding an AAR must be one in which the participants and leaders openly and honestly discuss what actually transpired in sufficient detail (Headquarters Department of the Army, 1993).

AARs have existed in less structured forms since time immemorial, but the U.S. Army where the first to formalize the process. Although their version is intended for military use, there are several successful examples of its use on crisis management assessment (e.g. Sexton and McConnan, 2003).

Traditionally AARs have been based on subjective data, coming mostly from individual self-reports and observations. As computation capacity grows the possibilities of using computer technologies to enhance the AAR process has become more apparent (Jenvald, 1999).

Reconstruction and Exploration

The R&E approach used was developed at FOI to enable analysts to objectively explore a course of events anytime after it occurred through visualization of selected parts of the recorded operation (Morin, 2002). The reconstruction phase consists of modeling, instrumentation and data collection. The outcome of the reconstruction phase is a mission history which may incorporate extensive process data including, but not limited to, audio, video, digital photographs, observation protocols, system log files, user-system interaction and position track files from multiple sources in the operational environment. Exploration refers to the work conducted by analysts that involves exploration of the course of events through a multimedia tool. Using this approach the analysts are able to search through the dataset in a synchronized and intuitive way to find answer to their questions without having to rely on subjective and sometimes erroneous observations or reports. The approach also allows for the analysts to feed back their findings to the mission history, thus making it increasingly richer and supporting higher level analysis in the next iterations (Albinsson et al, 2004).

F-REX

F-REX is a toolset developed to support R&E by constructing mission histories and exploring them through multimedia presentation in the main application, the F-REX Studio. The toolset also contains several applications for recording and collecting necessary data including, but not limited to, GPS tracks, computer screens and radio communication. One of the most important features of these data collection tools is time synchronization. All recorded data must be time stamped and tagged by its source in order for the Studio to correctly visualize the contents and enable the analyst to see it in its right context. The F-REX toolset was developed using experiences drawn from several years' usage of an earlier R&E framework, MIND (Jenvald, 1996; Morin, 2002) and consequently uses a similar representation of the mission history, shown in Figure 1.

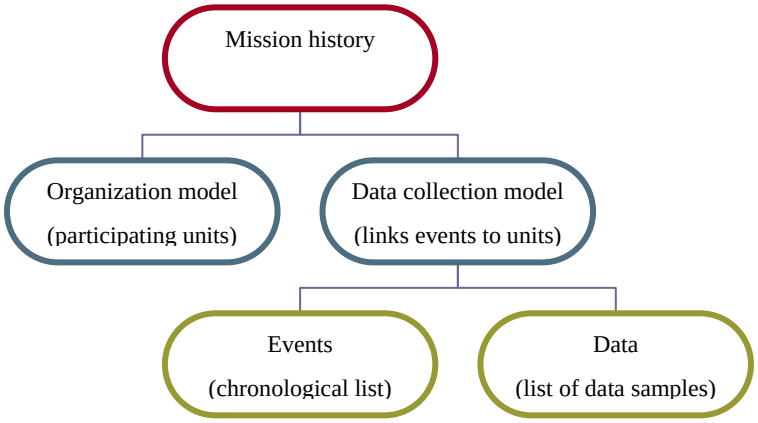

Figure 1. F-REX representation of a mission history. The organization model holds information on the organization, mostly known prior to the operation starts, such as participating units, planned events, incidents and injects, area of operation and more. The data collection model on the other hand holds information about what really happened such as reports, recorded data and observations.

The F-REX Studio plays back the entire mission history from the start to the end, with elementary functions such as pause, resume, rewind and fast forward. This together with the filtering functionality enables the analysts to shift between different foci and observe the chain of events in a meaningful context to make extended conclusions from simultaneous incidents occurring at different locations by multiple actors. A screenshot of F-REX Studio in action is given in Figure 2.

The analysis process described above can be more or less time consuming depending on input information known by the analysts and research questions asked. For a traditional AAR though, detailed analysis may not be necessary as the AAR itself can be part of the analysis. In these cases F-REX serves merely as a tool that will help visualize for the participants what actually happened and therefore raise the debriefing discussions from *what happened* to *why it happened*. An able analyst will regard this AAR as another opportunity to collect data about the exercise and feed that back into F-REX and the mission history, extending it with a layer of metadata that can be used as input for further analysis.

METHOD

For this case study F-REX was chosen as a representative system for assessing what features are desired for computerized evaluation of crisis management exercises. F-REX was configured and operated by system experts who also assisted Swedish Rescue Services Agency (SRSA) in conducting the AAR shortly after the exercise. Several informal interviews where then held both with SRSA teachers and with participating students to assess gained and desired advantages of using a computerized evaluation system such as F-REX.

The case study setting

At the SRSA training facility in Sandö several exercises are conducted every year to improve emergency readiness. SRSA are aware of the need for AAR to gain full advantage of exercises. They have traditionally used human observers for data collection and conducted informal AARs in form of interviews.

Some typical exercises at this facility have a rather complex command & control (C2) nature, involving operation upgrading, internal C2, external command post establishing and inter-organizational cooperation. Although most of the incidents usually occur at the same site, there are many trainees working simultaneously training different aspects of emergency response. Ideally every trainee should get the chance to train every exercise in every role at least once, but because of resource limitations that is not an option. Therefore the teachers traditionally host a debriefing after each exercise so that the trainees may exchange experiences from their different roles based on their own memories and on observation reports by subject matter experts (SME) watching and grading performance.

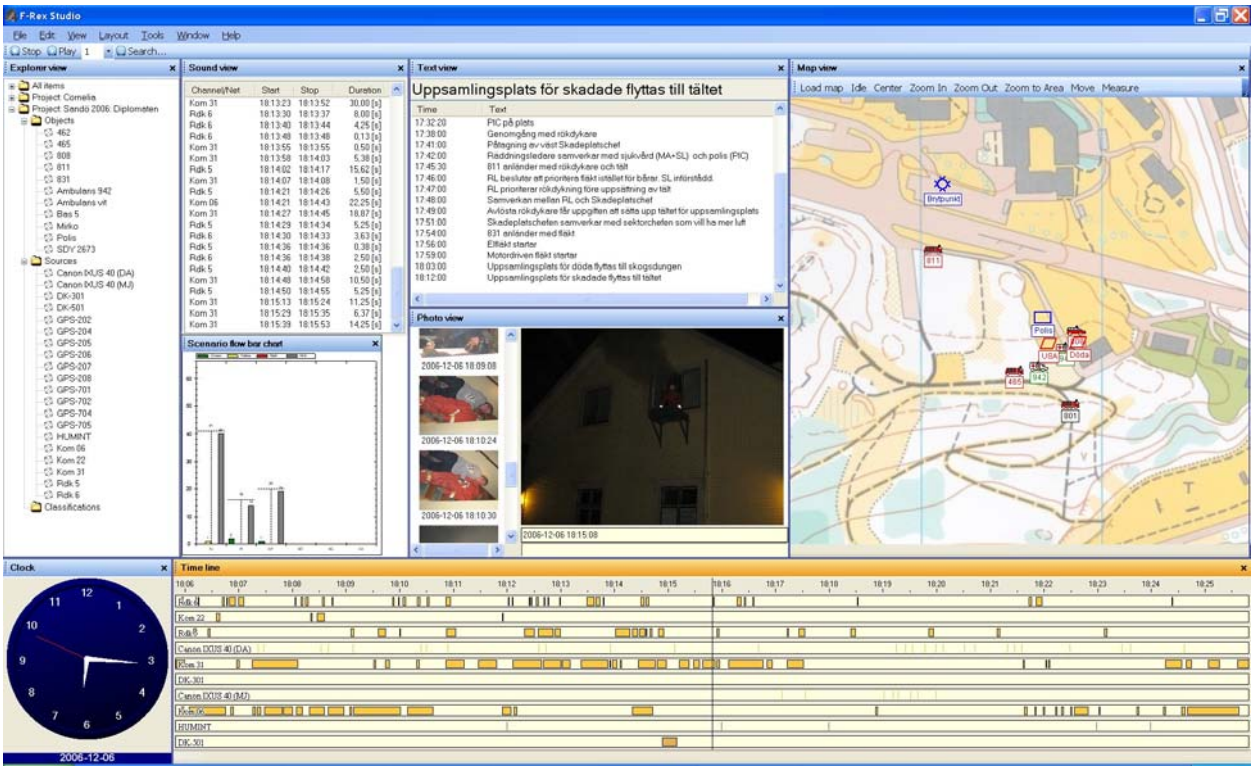

Figure 2. Evaluation of an SRSA exercise at Sandö training facility in 2006, visualized in F-REX Studio.

The applying of R&E and F-REX

In December 2006 a field study was performed in cooperation with SRSA to investigate the usefulness of R&E at the facility and try to identify additional features that would further improve both R&E in general and F-REX in particular when it comes to capability of supporting emergency operation AARs. The contract for this field study was to analyze a set of exercises during three days with focus on a burning house exercise the second day. Extensive data was collected and an AAR within 36 hours from exercise completion, assisted by FOI researchers that were using F-REX to demonstrate R&E capabilities and identify benefits of and additional requirements for computer supported AARs. (Thorstensson, Johansson, Andersson and Albinsson, 2007)

It was never a goal of this study to reconstruct an extensive model of the exercise but instead to keep it simple and focus on demonstrating the method using some of the most significant data sources for analyzing the C2 aspects of this large scale emergency response exercise only. For this reason it was natural to focus data collection on the trainee in the role of commanders, their communication and to some degree the events that directly influences their perception of the current situation.

Interviews

Immediately after the AAR all participating trainees were assembled to discuss the added effect of using F-REX and R&E. No formal interview protocol was used for this discussion; instead the trainees were given the chance to openly share their reflections on this debriefing method.

In August 2007 a seminar was held at the same facility with most of the SRSA teachers and some of the trainees who participated in the December exercise. Add to this a few more attendees from different parts of the crisis management organizations to get a total of around 30 persons attending the seminar. During this four hour seminar nine questions were asked (see Table 1) and the attendees were grouped into groups of 3-4 persons to discuss each question internally before sharing to the rest of the group. The seminar was led by FOI researchers and the SRSA teacher who organized the exercise.

1. What do you consider to be the main benefits of using similar approaches?
2. What do you consider to be the main limitations of using similar approaches?
3. For what purposes can this type of methods be used?
4. When and for whom will this type of methods be most useful?
5. Do you find any ethical issues of detailed recording activities for playback?
6. Order each data source according to usefulness with motivations
7. What data sources are you missing?
8. What data sources need to be enhanced or otherwise further developed?
9. How do you think the collected data should be presented in a better way?

Table 1. Seminar questions

RESULT

This section presents how the R&E approach and the F-REX Studio were applied in detail in the current case study context. This is followed by initial evaluation of the approach and tool according to interviews with participating teachers and trainees.

The R&E approach and F-REX application

The R&E approach was used to collect necessary data to reconstruct a useable model of the exercise for use as discussion material in the AAR. R&E consists of two main phases; *Reconstruction* and *Exploration* respectively. The Reconstruction phase in turn consists of *Domain analysis*, *Modeling*, *Instrumentation* and *Data Collection* while the Exploration phase consists of one or many *Presentations* (Morin, 2002). How each of these steps was conducted is described below.

Domain analysis

The domain analysis aims to find crucial topics in the work situation, particular problems of interest and some sort of prioritization among them (Morin, 2002). In this case study SMEs conducted several interviews with SRSA teachers at Sandö several weeks in advance to find a general priority of what to assess.

Modeling

The modeling step serves to produce an object-oriented conceptual model that defines the main actors and activities in the domain and an instrumentation plan outlining the needs for data collection and data presentation, based on the output from the domain analysis (Morin, 2002). While it would certainly be possible to model each unit, vehicle and even person to a very detailed level, the outcome of the domain analysis showed that the most important participants in this study were the commanders therefore only the commanders and those that they interact were modeled in F-REX for this particular exercise.

Instrumentation

Instrumentation is the step of turning models and plans into procedures, equipment and software components for data collection and presentation (Morin, 2002). For this experiment six different data types were collected, all using commercial off-the-shelf devices (COTS) and in-house developed software: GPS tracks, audio, video, photos, human observations and mock patient cards. It was clear at an early stage that even though it would be desired to have all data captured and collected live via streaming techniques, some level of post-processing would be required for all data to be fitted into the F-REX Studio for presentation.

Data collection

Data collection is the only step of R&E that is conducted during the operation. In this step procedures and tools devised in the instrumentation step are used to capture the activities in the operational environment. The data collected are converted and combined to produce the mission history. (Morin, 2002)

Handheld GPS devices were used to capture the locations of each response vehicle until arrival at the scene (fire engines, police patrol cars, ambulances, mobile command and control posts). In this scenario GPS tracks were secondary to other data types since most of the interesting action occurred on the same site.

Each radio channel was captured using a laptop connected to a radio unit monitoring the channel. For direct (internal) communication, such as order sessions, offline sound recording units were attached to several significant trainees. Finally a laptop was used to tap into the telephone system for the dispatch center to monitor incoming reports via cell phones.

Observers using handheld digital camcorders enabled capturing of interesting sequences on video. This was used sparsely because of problems with darkness and risk of interfering with the scenario.

Several observers with digital photo cameras were taking pictures of significant events in the exercise. Automatic time stamping of these enabled F-REX Studio to visualize important incidents in the chaotic environment and thereby helping the trainees to remember what actually happened. (Thorstensson et al, 2007)

Subject Matter Experts (SMEs) were continuously monitoring the exercise and taking time stamped notes on performance and significant events. These reports function as an index into the large set of data for F-REX Studio, enabling analysts to quickly get some idea of what happened and were to begin their detailed studies of the exercise.

A large number of younger students were volunteering as extras acting as casualties in the exercise. Each student was given a synchronized watch and instructions to create time stamped checkpoints (Thorstensson, Morin and Jenvald, 1999) on when they were discovered by the trainees, when they were carried out from the burning house and finally when they were attended to by medical staff. Collected data was then compiled into a detailed chart for F-REX visualization. (Thorstensson et al, 2007)

When using COTS devices one has to expect and plan for system errors and keep a descent level of redundancy to ensure successful data collection, or accept the consequences of loosing a few data sources due to errors. During this exercise a small level of redundancy was kept, enough to ensure that as little as possible of the most critical data was lost, but still there was some data loss because of failing GPS loggers and drifting system clocks.

After the exercise all devices were connected to the F-REX database and aligned with the conceptual model to complete the mission history and make it accessible for the teachers, analysts and trainees.

Presentation

Presentation is the step of turning the data included in the mission history into a form that supports exploratory analysis of the operation (Morin, 2002). In this study, the exploratory analysis consisted of an AAR. The AAR required some post-processing of data, and some manual work to connect all data to the model. An F-REX assisted AAR was then led by SRSA assisted by R&E experts. The principles of the AAR closely followed the layout suggested by the U.S. Army (Headquarters Department of Army, 1993), but with the added benefits of R&E support as described by Morin (Morin, 2002).

By reliving the exercise through this visualized exploration in F-REX, the trainees were given the opportunity to perceive their own actions and decisions, enabling them to reflect and comment on their own leadership, so that they, as well as the other trainees, could benefit from experiences gained in this exercise. (Thorstensson et al, 2007)

Case study evaluation

Informal interviews with AAR participants and teachers show that all AAR participants found the presentation to be a very useful support for the evaluation. Partly because it helped remembering many details of the chain of events that sometimes is very hard in a chaotic environment such as an emergency response site, but also to clarify what orders were given and what decisions the commanders themselves made. Multimedia data such as still images, video and recorded sound of the surrounding environment helped visualizing the current situation for those trainees not in the role of commander and who were not themselves at the current site at that specific time. After the AAR all collected data was handed out to participating students who were then able to explore it themselves. Several students took advantage of this opportunity to single-handedly focus more on evaluating their own performances. (Thorstensson et al, 2007)

The most appreciated feature of this computerized exercise evaluation was without doubt considered to be the time synchronized playback of radio communication (Thorstensson et al, 2007). Being able to see the chronology of

important events in a non-disputable list was also very much appreciated by the trainees to help them organize their memories. The other data sources: GPS tracks, observation notes and patient data cards were regarded as secondary information, yet very useful such as a means of navigation through the dataset to find interesting events.

Most participants had never seen a computerized evaluation tool before and where so overwhelmed with the fact that they could see and hear themselves in action that they had a hard time answering what they could possibly need more, but a few features where commonly missed; most notably the ability to realize what was being presented to them by the toolset. While the current version of F-REX does support shifting of focus between different trainees the functionality requires much manual labor with the dataset prior to the AAR and was therefore not activated for this exercise. Another much missed feature was the direct communication between the commander and sub-commanders. Although this was recorded during the exercise using offline voice recorders, the post-processing took too much time to incorporate in the mission history before the AAR. (Thorstensson et al, 2007)

CONCLUSION

Our case study indicates that computer supported approaches for evaluation of large scale crisis management operations can be very useful to help participants getting a clear picture of what happened during the operation. Time synchronized data sources enables recreation of a chronological model which is essential in order to enable participants to sort their memories. Documenting and replaying multimedia data from the exercise also proved very valuable to non-participants who found this approach very useful for sharing lessons learned.

In the special case of assessing the performance of commanders communication is essential. Therefore, another important feature of computer supported evaluations is the ability to reproduce the commanders' communication in a time-synchronized way. This refers to all forms of communication, i.e. written or spoken; inbound or outbound.

Photographs and videos help visualizing the chaotic environment that often occurs at a crisis site. These are most useful when trying to realize the stress under which the personnel are working, something that may explain certain behaviors and thus should not be neglected in performance evaluation.

Secondary data, such as GPS tracks and observer notes are excellent to be used as a rough estimation of the situation in a chronology and should thus be used as navigational support when exploring the dataset.

For a computerized AAR data presentation is a key feature of an evaluation system. In an AAR it is clear that data must be easily accessible to the participants, for the AAR to gain maximum advantage from the computer model. How to best overcome this problem needs to be investigated further.

In conclusion we have presented an alternative approach to operation evaluation that has several benefits over traditional quantitative and qualitative methods. In a larger context we have identified some important system features in a rapidly growing research area where methods and systems are being developed all over the world, both within the military and the civil domain for training and evaluation purposes. Many of these systems are focusing on AARs, using maps, GPS tracks, video clips, statistical data and use the timeline to navigate when evaluating the exercise. There exist several AAR systems focusing either on evaluating operations or on evaluating human performance in simulations and virtual environments (Allen and Smith, 1994; Vasend, 1995; Gratch and Mao, 2003; Ford, Gallagher, Shockley and Wadsworth, 2005). The results of this study can provide knowledge and guidelines for future approaches such as these.

ACKNOWLEDGMENTS

This study was funded by the Swedish Rescue Services Agency and the authors would like to thank teachers and trainees at the Sandö training facility for their cooperation and invaluable contribution to this paper and their interest in our research.

REFERENCES

1. Albinsson, P-A., Wong, W., Pilemalm, S. and Morin, M. (2005) Using a Reconstruction-Exploration Approach to Empower Cognitive Analysis Methods: Possibilities and Challenges, *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL, USA
2. Allen, G. and Smith, R. (1994) After Action Review in Military Training Simulations, *Proceedings of the 26th conference on Winter simulation*, Orlando, FL, USA.

3. Ford, R., Gallagher, V., Giuli, R., Oberg, S., Shockley, J. and Wadsworth, I. (2004) The JTEP Takehome Package: An AAR Tool for Distributed Training, Paper 04F-SIW-112, *Proceedings of IEEE 2004 Fall Simulation Interoperability Workshop*, September 2004.
4. Headquarters Department of the Army (1993) A Leader's Guide to After-Action Reviews (TC 25-20), Washington, DC, 30 September 1993.
5. Gratch, J. and Mao, W. (2003) Automating After-Action-Review: Attributing Blame or Credit in Team Training, *Proceedings of the 12th Conference on Behavior Representation in Modeling and Simulation*, Scottsdale, AZ, USA.
6. Jenvald, J. (1999) Methods and Tools in Computer-Supported Taskforce Training, *Linköping Studies in Science and Technology, Dissertation No. 598*, Linköping University, Linköping, Sweden.
7. Morin, M. (2002) Multimedia Representation of Distributed Tactical Operations, *Linköping Studies in Science and Technology, Dissertation No. 771*, Linköping University, Linköping, Sweden.
8. Morin, M. and Albinsson, P-A. (2005) Exploration and context in communication analysis. In C. Bowers, E. Salas and F. Jentsch (eds.), *Creating High-Tech Teams: Practical Guidance on Work Performance and Technology*. Washington DC: APA Press.
9. Morrison, J.E. and Meliza, L. L. (1999) Foundations of the After Action Review Process, *IDA Document 2332*, Institute for Defense Analyses, Alexandria, VA, USA, DTIC/NTIS AD-A368 651.
10. Rankin, W. J., Gentner, F.C. and Crissey, M. J. (1995) After action review and debriefing methods: technique and technology, *Proceedings of the 17th Interservice/Industry Training Systems and Education Conference*, Albuquerque, NM, USA.
11. Sexton, R. and McConnan, I. (2003) A Comparative Study of After Action Review (AAR) in the Context of the Southern Africa Crisis.
12. Thomas A. and Kopczak, L. (2005) From Logistics to Supply Chain Management: The Path Forward in the Humanitarian Sector, Fritz Institute.
13. Thorstensson, M., Morin, M. and Jenvald, J. (1999) Monitoring and Visualisation Support for Management of Medical Resources in Mass-Casualty Incidents, *Proceedings of the International Emergency Management Society's Sixth Annual Conference*, pp. 179-188, Delft, The Netherlands.
14. Thorstensson, M., Johansson, M., Andersson, D., Albinsson, P-A. (2007) Improved outcome of exercises – Methods and tools for training and evaluation at the Swedish Rescue Services school at Sandö, User report FOI-R--2305--SE, Totalförsvarets forskningsinstitut, Linköping, Sweden *(in Swedish)*.
15. Vasend, G. (1995) After Action Review System Development Trends, *Proceedings of the 27th conference on Winter simulation*, Arlington, VA, USA.
16. Wong, W. B. L., And Blandford, A. (2004) Situation awareness in emergency medical dispatch, *International Journal on Human-Computer Studies 61*, pp. 421-452.