PROOF COPY ONLY**DO NOT DISTRIBUTE**

JEM

Reconstruction and exploration of large-scale distributed operations—Multimedia tools for evaluation of emergency management response

Sofie Pilemalm, PhD
Dennis Andersson, MS
Niklas Hallberg

AQ1

ABSTRACT

This study presents an approach for computer-supported reconstruction and exploration (R&E) of distributed tactical operations. The approach involves several steps for constructing a time-synchronized, event-driven multimedia model of the course of events collected from multiple sources in the operational environment and visualizes this model in the F-REX Studio multimedia suite. In this study, the use of R&E and F-REX is explored in large-scale emergency management exercises. The approach's possibilities, limitations, and needs for modification are first outlined followed by a comparison to traditional quantitative and qualitative data collection methods applied in the same context. It is found that the R&E approach in combination with F-REX has several advantages in relation to the other methods, in terms of avoiding problems of retrospection and in being able to provide an overview of the entire operation based on multiple perspectives—addressing the question “why” something happened rather than “what happened.” Correctly used, multimedia-supported R&E can thereby be used for more solid evaluations of large-scale emergency management exercises and operations, thus contributing to more effective handling of future crises.

Keywords: reconstruction & exploration, distributed tactical operations, emergency management, evaluation, computer technology

AQ2

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.¹ To produce effective response to such emergencies, training and evaluation of training activities and of real operations is necessary. To be able to perform such evaluations, data collection and analysis is needed. This may be a rather straightforward process as long as the studied operations or exercises are limited to a reasonable number of individuals and tasks. However, as the context grows, so does the complexity. Several organizations and numerous organizational units and actors may participate displaying multiple perspectives, intrinsic patterns of action, and communication. Here, certain difficulties arise, above all related to capturing the context and integrating the different perspectives.² AQ3

Evaluating large distributed tactical operations (DTOs), such as fire fighting, law enforcement, emergency management, rescue, and military operations, is indeed a complex task. Such operations tend to be large scale with numerous actors and have several units coordinated but are still scattered over large geographical areas and contain multiple sets of parallel activities.¹ This makes them extremely difficult to evaluate, regardless of whether they are training exercises or real operations. If the evaluation relies on quantitative measures, a large group of respondents can be reached. Measures often focus on assessing how the respondents experienced the situation and whether the operation lived up to the predefined goals according to the number-based ratings. However, pure quantitative approaches do not try to explain what actually happened during the

PROOF COPY ONLY**DO NOT DISTRIBUTE**

operation or exercise. In other words, they fail to explore breakdowns and insufficient handling that can be used to determine needs for improvements.

At the other end of the spectrum are the qualitative research methods, such as cognitive analysis approaches.³ These approaches usually involve participatory observation, in situ probing interviews with key actors, and analysis of transcribed after-the-fact interviews. A problem here is that, traditionally, the analysis is seldom based on the reconstruction of a model of the entire work session or context, but rather summarizes particular incidents from the perspective of a few individuals. 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.⁴ Using qualitative methods in large-scale complex DTOs and training sessions is also time consuming and cost some, considering, for instance, the number of observers needed.

AQ4

This study presents an approach for reconstruction-exploration (R&E), which explicitly focuses on rich representations of DTOs with the support of the *F-REX Studio* multimedia tool. The reconstruction phase involves several steps for constructing a time-synchronized, event-driven multimedia model of the course of events collected from multiple sources in the operational environment.⁵ The exploration phase refers to the visualizing of this model in the F-REX Studio. R&A and F-REX have been used successfully to evaluate several DTOs of some scale; however, they have not yet been applied in significantly large and networked contexts. In this study, R&A and F-REX are explored in significantly large DTOs within networked contexts. In the study, their possibilities, limitations, and needs for modification are outlined, followed by a comparison with other quantitative and qualitative data collection methods.

Aim of study

The aim of the study is to explore the use of R&E supported by multimedia tools such as F-REX for evaluations of DTOs in large-scale networked contexts including military, crisis, and emergency management

operations. The study specifically addresses the following question:

- *Is the R&E approach, supported by proper multimedia technology, applicable to large-scale operational contexts and what are its benefits, drawbacks, limitations, and needs for modification to suit these contexts?*

The main question, however, is

- *How can rich reconstructions based on large-operation factual data, and visualized by multimedia tools, complement, improve, or even replace the use of traditional quantitative or qualitative research methods when evaluating large-scale distributed operations?*

The study was performed as a case study at a nation-wide, large-scale military command and as a control exercise. The study results are expected to contribute to future development efforts in the domain of multimedia support for reconstruction and evaluation of both exercised and real operations, both in military settings and in emergency management.

AQ5

BACKGROUND

In this section, the background of relevance for the study is described, including DTOs, quantitative and qualitative research methods, R&E, current multimedia tools for data analysis and evaluation, and the F-REX Studio.

Distributed tactical operations

DTOs refer to the type of human activities that aims at achieving specific goals with a body of personnel and equipment under unified command and that are geographically distributed.⁶ Typical examples are military operations and emergency management including evacuation, search, and rescue activities. In such activities, the command and control functions are often geographically distributed from the actual operative fieldwork. DTOs may also include distributed and/or interorganizational command and control. That

PROOF COPY ONLY**DO NOT DISTRIBUTE**

PROOF COPY ONLY**DO NOT DISTRIBUTE**

is when two or several organizations jointly perform an operation and their respective command and control functions must coordinate their activities, as is often the case in emergency management. Hence, this could mean several staffs and commanders in the field (eg, policemen, fire fighters, nursing staff). The term *tactical* implies that the operation includes essential amount of considerations and decision-making activities. Important for DTOs is reliable communication so that information can be shared to obtain joint situational awareness. To be able to perform real DTOs, it is vital to exercise in beforehand to overcome potential deficiencies in the accomplishment.

Both real and exercised DTOs are candidates for evaluation. Both cases are also characterized by the multiple sources of complexity involved. As several human activities are performed in parallel, extensive dynamism, significant interdependencies between personnel, multiperspectives, and a need for common situational awareness arise. Further, because conditions can vary dramatically during the course of events in a time-critical situation, the nature of problems also changes over time and the events may take unexpected turns due to the actions taken. The personnel involved have to deal with parallel on-going tasks and problems with multiple possible causes and consequences.²

When it comes to analyzing the work in DTOs, the distribution of cognitive resources presents particular challenges. The analyst, eg, has to consider multiperspectives, seek to avoid fragmentation, and be able to distinguish between intent that is publicly communicated and implicit intent that is assumed from the cultural, organizational, and individual context.⁷ The analysts must also be able to grasp isolated events on a particular level of the operation performed and relate these events to other events (current and past) and to contextual circumstances on various unit levels and geographical locations. This is not without difficulties, using traditional research methods.

Quantitative research methods

Quantitative research methods are based on the belief that it is possible to identify general "truths,"

based on rigorous scientific research and controlled experiments. They thereby rely on, eg, measurements, variables, statistical procedures, testing of hypotheses, and creating validity.⁸ Strict control characterizes the research process. Commonly applied methods are structured questionnaires or interviews where the respondent is supposed to select from a predefined set of alternatives, for example, attitude questionnaires. Often experimental designs are set up, involving an experiment group that is subjected to certain manipulated variables and a control group that is not.⁸ Data analysis usually involves statistics and looking for cause-effect relationships, general tendencies, and confirmation and rejection of hypotheses. Statistics is also important when presenting the results in tables or graphical representations.

When studying DTOs or exercises, quantitative research methods are the least resource demanding means for evaluation. Attitude surveys can be used to evaluate different respondents' experience of the operation and to identify general tendencies and attitudes. Human performance studies can also be carried out.⁹ However, there exist several problems that accompany quantitative evaluations and some of them become intensified in large-scale contexts. Attitude questionnaires indeed provide knowledge on perceived experiences and attitudes; but no explanations of these. In other words, they do not tell what actually happened during the operation and why people have perceived the situation in a certain way. This leads to a risk that problems or difficulties occurring during the operation escapes identification and, subsequently, no actions for future improvements can be taken. Results become decontextualized. Further, quantitative approaches have to deal with general problems of validity and reliability. For instance, error values in questionnaires may occur, due to that the respondent has interpreted a question incorrectly or due to factors other than those put forward by a predefined hypothesis have had an impact of the respondent's ratings. Such problems tend to increase with the growing scale of the evaluation context; it is extremely difficult to control large-scale unpredictable real-life human operations and treat them as experiments.

PROOF COPY ONLY**DO NOT DISTRIBUTE***Qualitative research methods*

Qualitative research aims at exploring individual, social, organizational, and cultural phenomena in their real-life context. They provide no rigid set of rules but put an emphasis on subjectivity, sense making, and interpretation. The researcher studies representations of the real world in naturalistic settings, analyzes, and reconstructs these representations.¹⁰ Qualitative research methods span over different interview techniques, questionnaires, various forms of participatory observation, and contextual inquiry.

When it comes to studying contexts involving a number of humans and artifacts, the field of *cognitive analysis* especially provides a multitude of methods. Some of the examples are given below:

The Critical Incident Technique (CIT) relies on the self-report of incidents, those relating to satisfaction and the demonstration of efficiency in the situation, and those that lead to perceived dissatisfaction and inefficiency.^{11,12} The information gained can be used to take precautions to improve the current situation, through the change of work routines.

The Critical Decision Method (CDM) is an offspring of CIT and is a retrospective interview technique for studying decision making in critical, out-of-bound situations.¹³ CDM uses the incident as a timeline, from which to probe the interviewees' memory about information they used, considerations they made, actions they took, and reasons for their behaviors. From these data, conclusions can be systematically drawn about decision strategies, interaction behaviors, problems, and information needs over time.

Contextual inquiry developed in recognition of the role that context played in interpreting data.¹⁴ It generally involves participatory observation of the behaviors of people during routine activities for gaining insights about how work actually is carried out, how people use their tools, and how they coordinate their activities with other team members. One step further, *process tracing* sets out to capture a series of events involving several individuals by collecting data using, eg, interviews, participatory observation, and verbal and behavioral protocols.¹⁵ The data are successively integrated into a time-lined account. Process tracing shows some similarities with the F-REX approach, eg, with regard to the use of phases,

the time-line, and multiple data sources as input to the representation of the results.

Neither general qualitative research methods nor the specific cognitive analysis methods escape problems when moving from small, homogeneous contexts to collecting data from very large and complex events such as DTOs. Here, it is often impossible to conduct probing interviews due to the nonroutine, critical tasks of the participants and due to inaccessible or distributed locations. Similarly, there is no time to fill in questionnaires when activities are still in progress. Hence, contextual inquiry may be ruled out; meanwhile, CIT and CDM have to be carried out in retrospect, which brings problems of subjectivity and reconstruction. The respondents may forget incidents or fragments of incidents and may also intentionally or unintentionally distort the information they provide. CIT and CDM also provide a limited number of individual perspectives on the studied context, and they do not capture the simultaneous multiple perspectives, the dynamic collaboration, the factual course of events, and the wider context underlying the individual reports.

Participant observation/contextual inquiry and process tracing are more oriented toward capturing context. However, to observe large, complex, and geographically dispersed events, an extensive number of observers performing in-depth investigations are required. The approaches are both costly and time consuming and it still does not fully solve the problem of integrating multiple perspectives. It is typical that even process tracing focuses on the decision-making process of an individual or of small teams, and it also focuses on internal cognitive processes of decision making but not on explaining real-life courses of action and what actually happened.

AQ6

Technical support for qualitative data collection and analysis

There have been several attempts to develop multimedia tools to support qualitative data collection and analysis. Examples are MCSHAPA,¹⁶ a tool for analyzing qualitative data from complex tasks using a timeline and DIVA¹⁷ that similarly supports exploratory data analysis using multimedia stream data including input from video and observations in a

PROOF COPY ONLY**DO NOT DISTRIBUTE**

PROOF COPY ONLY**DO NOT DISTRIBUTE**

timeline work frame. Both tools focus on capturing the current work process, and on visualizing and working with the collected data, but they do not have the distributed perspective or capacity that is needed to analyze large DTOs and they do not include facilities for automatic logging of factual data.

The R&E approach and the F-REX Studio

The reconstruction-exploration approach, R&E, was developed as a way to analyze DTOs and to extend the representation of operations above individual self-reports and observations.⁶ It has been successfully applied both in military and emergency management settings. To provide an overview of operations, a rich contextual dataset is required. *Reconstruction* involves several phases for constructing a time-synchronized, event-driven multimedia model of the course of events in a distributed work session. This model—the *mission history*—incorporates extensive process data such as 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 rendering of the mission history in a multimedia suite, the F-REX Studio, which supports explorative analysis of the data. During this exploration, analysis results, or metadata, are incorporated back into the mission history, making it increasingly richer.¹⁸

F-REX uses time as the primary navigation and coordination mechanism. The time stamps assigned to the data in the mission history provide mappings from data to time and from time to data. When the user selects a time point, F-REX reconstructs the state of the mission at that point from the data available in the mission history. Further, it updates all data-presentation views to reflect this state. It is thus possible to see, ie, what was said by different persons and the positions of vehicles at a particular time point. The user can also select a data item, eg, a communication sequence, and let F-REX synchronize the state of the mission to that item. In a replay mode, F-REX uses successive time points to animate the corresponding state changes. An example of what it can look like is provided in Figure 1.

The R&E approach supports a wide system perspective gathering factual, simultaneous, real-time

data on multiple organizational levels on multiple geographical locations and from multiple actors. Using the F-REX Studio, analysts can switch focus between different viewpoints in the substantial dataset. Currently the F-REX lacks well-defined analysis methods and the corresponding technical support for coding of collected data. Moreover, even though F-REX supports a wide system perspective, its scalability up to significantly large-scale DTOs involving several hundred actors, organizational units, and computers have not yet been tested. Because many DTOs in major emergency situations actually are of this size, the R&E approach and the F-REX Studio should adapt to larger-scale contexts than before.

Multimedia tools for evaluation of DTOs

Although the F-REX Studio is unique in certain respects, related efforts to capture DTOs exist. Within the military and civil domain there are numerous systems developed for training and evaluation purposes. Examples are military battle training systems and after-action-review (AAR) systems.¹⁹ Traditional AARs build on observations and are used to identify success and failure aspects of operations to achieve improvements.²⁰ In military settings, there exist AAR systems focusing on evaluating either real exercises or human performance in simulations and virtual environments.^{21,22} Several companies sell AAR systems as “take home packages,” so that the reviews can be used for individual training at home.²³ Several of the AAR systems use maps, GPSs, video films, and statistics and have a timeline as a focal point when evaluating the exercise. AQ7

Although traditional AARs focus both on individual performance and on the whole team/operation, most corresponding AAR systems have the individual user and his or her performance in focus. Only a few systems embrace network aspects and communication (eg, the PowerSTRIPES developed by American AcuSoft, www.acusoft.com). Particularly, the usage of AAR systems for evaluating large-scale DTOs remains unexplored. Therefore, the R&E approach and the F-REX Studio can serve as useful experience paving the way for future developments of similar systems aimed at performing improvement-oriented evaluations.

PROOF COPY ONLY**DO NOT DISTRIBUTE**

Figure 1. Overview of situation at a certain time point as displayed in F-REX providing, from upper left to right, an organization tree, a detailed map with positions of rescue units, a flow of casualties, communication events, and annotated digital photographs; in the middle, the timeline; and from lower left to right, an overview map, a video clip, and communication data in terms of radio traffic.

METHODS

The study was performed as a case study in the Swedish Defence area involving participatory observation and complementary group interviews.

Case studies

Case studies imply that a real phenomenon is studied, eg, an individual, a setting, an incident, or an organization.²⁴ According to Yin,²⁵ they are suitable “when ‘how’ and ‘when’ questions are being posed, when the investigator has little control over events, and when the focus is on a contemporary phenomenon within some real-life context.” Case studies can be of an explorative, descriptive, or explanatory character.²⁵ According to Yin,²⁵ case studies should include the stepwise procedure of designing the study, conducting

it through preparing for data collection and collecting the evidence, analyzing the evidence, and composing the case study report. The current study is explorative in character and follows the recommended procedure.

Participant observation and interviews

Participant observation is the observation of subjects in their real-life settings by their participation in day-to-day activities. It is a direct research method that provides information on how the studied situation actually is, not how it is retold and maybe distorted by the study subjects. Participant observation further supports flexibility and the successive accumulation of knowledge.²⁶ It has been suggested that it can be beneficially used in data triangulation for increased validity of the final results.²⁷

PROOF COPY ONLY**DO NOT DISTRIBUTE**

PROOF COPY ONLY**DO NOT DISTRIBUTE**

Interviewing is a commonly used method to gain information on and understanding of peoples' experiences.²⁸ The concept covers a lot of ground, from completely unstructured to highly formal, controlled interactions with the respondents.²⁶ Interviewing is suitable in cases wherein the researcher is interested in the perspectives of the subjects under study and of how they perceive their reality.²⁸

Focus groups are interviews in which several respondents participate simultaneously. They thus have a collective, multivocal focus and can be used to display inner dynamics of groups and different perspectives.²⁹ Focus groups are often used in triangulation of data, eg, as a complement to participant observation.²⁶

Case study setting and data collection

The case study setting was a nation-wide military exercise taking place in Sweden in the late autumn of 2005. The Swedish Defence has constructed an extensive facility for IT-supported military command and control training. Numerous advanced information and communication systems are available, eg, simulators and common operational pictures providers. Twice every year the Swedish armed forces are gathered in the facility for a large-scale national exercise. At this particular exercise, the total number of participants was approximately 300 using a total of about 200 different workstations. This was the first time the R&E approach and the F-REX Studio were used in such a large-scale environment.

The exercise consisted of four linked scenarios played over a period of 2 weeks. Each scenario was followed by short periods of evaluation and debriefing. During the entire exercise, R&E was applied to capture the course of events and reconstruct it in F-REX for further analysis and evaluation. For this purpose, three system experts and one project leader were involved, the *F-REX operators*. Following each scenario an AAR was scheduled with participating exercise actors, where selected parts of the scenario were replayed to the actors, allowing them to reflect over and provide feedback on anything related to the course of events. Concurrently, the exercise was evaluated by traditional qualitative and quantitative research

methods applied by different groups of researchers and by the military exercise managers.

The *F-REX evaluator* performing this study participated as an observer of the exercise for the entire 2 weeks. Exercise actors in the field were being closely monitored as were the system operators keeping track of data sources and recording to ensure that sufficient data were collected for F-REX Studio. The observation, thus, had its focus on the R&E/F-REX applicability in this new environment, perceived benefits, and experienced problems, and how it complemented traditional research methods for evaluation. Notes were collected in a digital log, sometimes complemented with video documentation, eg, of the AAR sessions. In addition, a focus group interview was held with five military observers who had been participating in the field during the exercise. The interview was regarding their experience of certain aspects of F-REX.

In the data analysis, the notes from the diary and the video documentation were scrutinized by the evaluator. Similar data were initially clustered and successively formed into categories displaying certain aspects of the use of R&E and F-REX. The categories were discussed with two of the F-REX operators and some modifications were made before arriving at the final results. All three operators and the evaluator discussed the final results at a half-day workshop before documenting the final results.

RESULTS

The results of the study are presented followed by an analysis. The findings are presented in four different main categories: technical issues, methodological issues, integration with qualitative research methods, and integration with quantitative research methods.

Technical issues

At the technical level, the large-scale environment clearly posed new challenges for applying the F-REX Studio. It took the first day of play before things began to work somewhat smoothly. The most urgent problem concerned audio because audio transmission did not work at all from some workstations and where it worked it was of insufficient quality for playback to make any sense. At the end of the day, a decision was

PROOF COPY ONLY**DO NOT DISTRIBUTE**

made to focus on the smaller number of the workstations, ie, those 50 computers that were regarded the most relevant to capture the exercise in its essence. During the second day, audio transmission was established from selected workstations, still with some quality problems.

The rest of the data capture performed well with timed screen capture of selected workstations, video feeds, observer reports, and system logs. Compilation of data was initiated on the second day of play, followed by initial editing of the video clips and classification of the available communication data. The process was performed rather smoothly. A problem, however, was that it took a long time to handle data input, which was much more voluminous than has been the case in previous F-REX applications. While playing back it was also sometimes difficult to interpret the incoming stream of communication because many different actors spoke simultaneously. The problems remained through all the four scenarios.

Methodological issues

The large-scale context also brought difficulties of a methodological nature. At a general level, it soon became clear that dealing with the technical issues took too much of the time that was supposed to be devoted to analyzing data and constructing the mission history; technology directed methodology instead of the other way round. Further, several specific methodological problems were noted. At the most concrete level, with a growing number of exercise participants, more and more of them were ignoring their instructions to wear a headset (for recording purposes), thereby introducing a certain level of data loss.

It was also difficult to navigate and identify important events in data, much due to the large dataset. The difficulties delayed data analysis, construction of mission histories, and the subsequent AARs. In other words, the usual R&E methodology loop where it should be possible to hold an AAR 3 hours following the completion of an exercise was challenged. With a more limited number of workstations to focus on, things improved from the second scenario. Another contribution to the improvement was a change in the R&E methodology: a prioritization was now performed initially in the data analysis phase, deciding in what

order to analyze the different units' actions and starting out with those most central for directing the course of events in the current exercise. The AAR could not be held within 4 hours for any of the remaining scenarios as scheduled but at least they could be conducted during the same day as each scenario was completed, which was considered a partial success.

Also, the AARs, when held, illustrated several interesting points. The reviews provided explanations for several mistakes and confusion that had arisen during the scenarios. For instance, one AAR showed that a certain task of crucial importance for the assignment indeed was handled correctly by the military management during the game, but nevertheless the wrong task was solved at the given time due to the fact that different actors had different interpretations of the phase of the assignment they were currently in. This and similar findings led to a new way of applying R&E where several AAR participants wanted to proceed with the analysis by digging through the F-REX communication data themselves. With the support of the F-REX operators they were able to perform quick analyses by going through the data at an early stage, replaying communications, clarifying obscurities, and finding explanations to the courses of events. Several participants from the exercise management acknowledged that the F-REX Studio may be used to identify information structures and gaps in information during large-scale operations. During the exercise, the multiple simultaneous chains of information had caused many problems and misunderstandings. Applying F-REX was considered a smooth way of identifying and providing explanations to the problems.

Integration with qualitative research methods

Whereas the R&E approach was applied to capture and F-REX to reconstruct the large exercise, the exercise management used participant observation as the exercise main evaluation method. Approximately 20 observers, all with military background, were sent out to different units, locations, and workstations in the facility. Briefly defined guidelines in a paper-based protocol were supposed to guide their observations. Following each scenario, the observers were requested to fill in their findings in a web-based questionnaire.

PROOF COPY ONLY**DO NOT DISTRIBUTE**

PROOF COPY ONLY**DO NOT DISTRIBUTE**

To make use of data from participant observation within F-REX, a computer-based tool for the reporting of observation had been developed and integrated with the overall F-REX framework. This tool is called NBOT.³⁰ By means of NBOT the observers were able to report events in real time, informing the F-REX operators about important incidents at different locations. The observers carried personal digital assistants (PDAs) with NBOT installed and observations were reported in textual format, but it was also possible to send in audio clips and attachments containing drawings and digital pictures. The observers were also able to categorize their observations in real time, using predefined categories from the NBOT tool or creating their own.

In the study interviews, those five observers who had used NBOT expressed great enthusiasm with and clearly saw benefits of using the tool. They enjoyed the real-time aspect and considered the tool easy to use, time saving, and enabling them to make more notes than before. However, they also pointed out a risk with reporting "on the fly": that the observations provided in real time may be less reflected upon and less processed, as compared with traditional participatory observation. When comparing the observation material from the first scenario (where NBOT was not used) to the material from the others, the observation notes were indeed more substantial and numerous in the latter. Several observers even chose to abandon the protocols provided by the exercise management totally in favor of the new tool, reporting everything in real time.

The PDAs were perceived as small but still clumsy to handle by some observers. They asked for Tablet PCs that could be jacked to PC keyboards when possible, but at the same time also noted that PDAs are handier, if one has to move between different places during the observation. Someone also noted that it would be difficult to use the PDAs in outdoor exercises if the weather was rough and rainy.

The NBOT functionality, on the other hand, was highly appreciated and the observers requested additional functionalities. They wanted to be able to review and complement reports they had sent in earlier, a feature that was added after the second scenario. Other features recommended concerned the ability to report comprised messages of findings via voice/speech data, access to templates for reporting, the inclusion of

two-way communication between the observers and the F-REX operators, and facilities such as a war diary and a chat where observers can notify each other about important incidents that occur at different places and that may be missed by those who are not there.

A problem related to the observation data was that the observers were quite unsure of what to report and asked the F-REX operators "what kind of observations they wanted." Even though the operators provided some guidance, eg, that the observers might report both when things went smooth, as well as during sudden breakdowns and exceptions, this was found difficult to achieve in practice. Observations were still described in very general terms without illustrative comments that could actually aid the operators in the analysis process. Two observers complained that the predefined categories, ie, the scheme with categories and clickable sub-categories, in the NBOT did not make sense. The problem was accentuated by the fact that both the observers and the operators lacked sufficient domain knowledge to quickly identify illustrative events.

Integration with quantitative research methods

Another group of researchers evaluated the scenarios from a human-computer-interaction (HCI) perspective using a quantitative approach. Within the framework of this study, the HCI researchers, the F-REX system operators, and the evaluator collaborated to investigate whether R&E and the F-REX Studio could be integrated with and support the quantitative analysis.

One aim of the HCI-oriented study was to investigate how the actors at a specific unit, known as *Location C*, made sense of different operational pictures, ie, how different visualizations and formulations affected the exercise actors' situational awareness. The underlying motivation was to be able to develop proper communication methods and systems support for usage at Location C. Several hypotheses were formulated. Following each game, all participants somehow affected by Location C filled in a predefined questionnaire where they were to choose from different operational pictures and tell which ones they thought correctly displayed a certain situation during the game. They were also asked to rank how sure they were that their choice was correct on a 1-7 scale. Similarly, the

PROOF COPY ONLY**DO NOT DISTRIBUTE**

respondents were asked to place different situations that had occurred during the game in a correct order and to estimate the correction of their choice.

As for the integration, the F-REX Studio was extended with a tool for displaying different hypotheses. The hypotheses should be connectable to the dataset in F-REX, above all to the communication data. In this way, it will be possible to identify events that confirm or contradict the hypotheses and to provide possible explanations for this. In other words, the hypothesis tool in F-REX will constitute an initial value that is subsequently dressed with data. Such an integration with the NBOT observation tool will further ensure that the field participant observers report from different events that are of relevance to the hypotheses. During the exercise, the component was restricted to a visualization mechanism displaying hypotheses in the F-REX hypothesis tool. The component will be subsequently integrated with the tools that are presently provided by F-REX for linking hypotheses to data and explanations, and using statistics to display general results and trends (Figure 2).

Figure 2. The hypothesis view in F-REX displaying, in order, the current hypothesis (H0), variables, data sources, results and notes, preliminary and local interpretations, and comments and local comments.

The HCI researchers experienced various problems when performing their study. Several of the hypotheses they had proposed were not confirmed as expected and they perceived some of their findings as strange. In the quantitative questionnaires, the respondents claimed to have a sufficient situational awareness during the first scenario, but indeed performed poorly. After the second scenario, they claimed their situational awareness had improved, whereas, in reality, they performed even worse. Following the third scenario, they claimed that their situational awareness had further improved as compared with the second, whereas, in reality, they performed the worst. Also, the HCI researchers experienced difficulties as to identify which respondents among the multiple actors had the right competence and background to properly fill in the questionnaires, ie, they did not achieve a representative set of respondents. Yet another problem was that the exercise management continually changed the injections that obliged the HCI researchers to do continuous corresponding changes in their questionnaires. The problem regarding confusing questionnaire results was partly solved because it showed that they were in part due to miscalculations on the part of one of the researchers, who had compared the wrong operational pictures. Still, there remained hypotheses that were not confirmed without any explanation as to why.

As for exploring the initial integration of the hypotheses with the F-REX Studio, one of the HCI researchers scrutinized the observation data provided by the NBOT tool, identified, and listed those issues that were relevant for Location C. The chosen observations were used as a basis for reconstructing a mission history in F-REX to connect the hypotheses to the mission history and to analyze them in light of the mission context as provided. Even the mere investigation of the listed observations revealed some explanations as to why some hypotheses were not confirmed. A contributing factor was that several of the operational pictures that were applied during the exercise contained serious errors such as the wrong markings. The operational pictures had thus confused the respondents. Also, in certain phases of the scenarios, problems with different technology and systems were noted, probably affecting the HCI survey results.

PROOF COPY ONLY**DO NOT DISTRIBUTE**

PROOF COPY ONLY**DO NOT DISTRIBUTE****ANALYSIS**

The study demonstrates that the overall R&E approach and that of the F-REX Studio indeed are usable when applied in large-scale contexts involving numerous actors, computers, and masses of data. The strengths of the approach may even be accentuated as the growing size and complexity of operations and exercises increases the difficulties involved in, not computer supported, traditional quantitative and qualitative research methods such as lack of control, inadequate representation, and insufficient resources. However, the large-scale context also poses certain challenges to F-REX both at a technical and a methodological level, which need to be dealt with if the approach is to be successfully applied.

Technical challenges

It is clear that the large-scale context demonstrated some shortcomings of the current F-REX Studio capacity above all relating to scalability. This included insufficient and nonefficient data storage capabilities, which made it impossible to log all workstations. As for communication technology, it was not possible to verify that all the exercise participants used their microphones all the time. The experienced problems can be solved both technically and methodologically. Clearly, the F-REX data logging, collection, and storage capacity need improvements to handle the massive amount of data generated in this kind of operations. However, with fully functioning data collection routines there may be other scalability problems when it comes to performance of the visualization tools. A trade-off between the benefits of larger datasets versus performance must thus be considered. As for overcoming the problems of getting subjects to wear microphones, an alternative is to use a set of microphones fixed to the infrastructure in each cell in the exercise subsuming all occurring sounds. Of course, this solution applies only to indoor exercises and not to outdoor training and real operations. For the analysis of the large amount of data, automatic pre-data analysis methods are needed to be incorporated with the R&E studio. Such methods have been suggested in several contexts, for instance, video analysis by Courtney³¹ and team communication and cognition (PRONET and CHUMS) by Kiekel et al.³²

Methodological challenges

However, most challenges concern methodology issues. An alternative when approaching extremely large-scale contexts is to focus on a limited predefined set of actors and units; equip these with high-quality microphones and focus observation and logging of computers on these individuals. Even if an entire large context may not be reached in this way, when logging 50 or 100 actors, it is still possible to capture entire chains of events and to produce multiperspective, time-rich context data. Also, all aspects of a large DTO and exercise are not always relevant. Here, a trade-off between technology capacity and the specific aim of reconstructing and exploring the current operation must be made.

The full-scale perspective seems unrealistic also when reconstructing mission histories from the available data. It is better to begin with a limited set of data and then successively add to the history. Again, all data are not always necessarily needed, and the F-REX Studio still manages larger data sets than can be handled with traditional research methods. A further experience is that in large-scale contexts it is necessary to start compiling data while the operation/exercise and data logging is still in progress. This is because much more data is present than in smaller operations/exercises, and more time is needed for the analysis and the reconstruction of the mission history. It is also possible that in large-scale contexts it is not feasible to hold AARs based on analysis only a few hours following completion of the mission. Rather, it is desirable to hold "open AARs" presenting material from the mission (based on the predefined initial values on what data to collect) and letting the operation or exercise participants analyze the course of events in the presentation.

Formalization. To predefine a set of actors and units, beforehand knowledge of the context and which actors/units are most relevant for the themes that will be in focus during the reconstruction and analysis of the operation is required. In other words, the need for formalization becomes much more important than when applying the F-REX approach in smaller settings. Previously, it has been possible for the F-REX operators to participate in the field and acquire an overview of events during the logging of data. This has guided their analysis and the reconstruction of the operation.

PROOF COPY ONLY**DO NOT DISTRIBUTE**

Therefore, it has not been necessary to establish clear goals of recording the exercise beforehand even though R&E in theory states that goals should be identified in advance. In the large-scale context, however, the lack of distinct goals as to what to use R&E for will, even more than in smaller contexts, reduce the benefits of the approach and delay the analysis and the following AAR. To have a study design is, thus, even more crucial when going from more limited to larger contexts where huge amounts of data are to be sought through. Predefined goals, for instance, support the identification of important actors and units, which in turn support a structured approach to handle communication data, starting out with the most important voices rather than “digging” into them all at the same time.

Observers/experts. As the context expands and the identification of critical events and illustrative data becomes more complex, the importance of domain knowledge and observers/domain experts similarly becomes crucial. In previous smaller exercises, the person in charge of the exercise has also been in charge of the AAR and supported the F-REX operators in the data analysis to quickly identify critical events. In the current case, there existed no single person responsible for the exercise and no specific domain expert was available to the F-REX operators.

In the future, several domain experts/observers should be continuously available to the F-REX operators to work together with them in interpreting data, but also to observe in the field at different locations. The observers can then provide the F-REX operators with the information the operators previously had to acquire themselves and support the analysis by identifying where to look for and how to interpret data. The predefined goals for recording the operation/exercise must be present to the observers/domain experts. The study design should therefore be worked out by the operators and domain experts together.

In the study, the integration of the observation tool NBOT with the F-REX suite was a clear success and should be developed further. In the current case, NBOT was used by general observers with multiple observation tasks in the exercise. In their observations, they were colored by their simultaneous use of a web-based questionnaire where they were supposed

to respond to a number of predefined questions provided by the game management. When equipped with NBOT they indeed reported more observations than before but these observations were still restricted in content. For the future, the users of NBOT should be the observers/domain experts working closely with the F-REX operators, as outlined above.

Domain experts are also needed to enable close to real-time data analysis if R&E is to support a sequence of scenarios or exercises where findings from one AAR can provide input to improve the next exercise. To render the analysis phase even more effective, it would also be useful to integrate an analysis tool with the overall F-REX Studio. The analysis tool should support quick coding of transcribed communication data using predefined categories. For transcriptions, several more operators than used previously are needed to speed up the transcription process for voluminous data.

Integration with other research methods

The study demonstrates several advantages of the R&E approach and the F-REX Studio as compared with traditional quantitative and qualitative research methods when used to capture large contexts in real time. First, the HCI study performed during the exercise substantiated several of the difficulties generally associated with quantitative research. The questionnaires that were used only measured estimations of own behavior, and the degree of correctness in the respondents' perceptions of different operational pictures. When the researchers' predefined hypotheses were not confirmed they were confused to not find any explanations to this in the questionnaires. With an approach like R&E, explanations can be found to quantitative results that otherwise would be difficult to explain. For instance, in the communication data there was much reasoning about certain errors in the operational picture such as it displaying double units and targets that came and went. This has certainly influenced the participants' responses and estimations in the questionnaire. The HCI questionnaires did not capture any information about the information chain problems and the misunderstandings concerning the different phases of the exercise that may have affected the respondents' perceptions. Meanwhile, this was rather easily acknowledged by using F-REX Studio.

PROOF COPY ONLY**DO NOT DISTRIBUTE**

PROOF COPY ONLY**DO NOT DISTRIBUTE**

Some of the surprising tendencies in the questionnaires were in part due to miscalculations made by the HCI researchers themselves. But this took time to find out and illustrates how vulnerable the quantitative “nonexploratory” data are to errors and thereby to produce incorrect interpretations. R&E, on the other hand, is based on what actually happened out at the field. Whereas quantitative methods presume an experimental environment that is not easily transferred to real, large-scale complex world settings, the F-REX Studio reconstructs these real settings and shows the effect of dynamic courses of events on predefined estimations and hypotheses.

The perceived deficiencies of the quantitative approach could, for sure, be overcome also by using traditional qualitative methods such as participant observation. Still, these qualitative research methods have to deal with subjectivity and the retrospective reconstruction of events. Meanwhile, what is logged by F-REX has indeed happened and this also facilitates presentation of less faltering and politically

controversial results. Also, even if participant observation and contextual inquiry catch the real-time aspect, the R&E approach is still superior in getting the overview and multiple perspectives of an exercise/operation and in capturing voluminous data from the field, identifying and working with several situations simultaneously. In the study, the “manual” participant observers appeared lost and reported only fragmented data from different locations. Further, in outdoor operations and exercises, space is sometimes very limited (for instance, in a vehicle) and it may be difficult to observe at the spot. Here, the R&E is a good alternative, capturing data from another location. Finally, with traditional observations, it is not possible to overhear everything said because voices are simultaneous and many participants use headsets to communicate. The missing gaps can be filled by F-REX that has the entire logged conversations of individuals.

In Table 1, the characteristics of the R&E approach combined by the F-REX Studio in relation to a variety of quantitative and qualitative research

Table 1. The F-REX approach in comparison to a number of traditional quantitative and qualitative research methods

Research approach	Reaches large group of respondents	Provides in real-time data	Provides graphical representations	Objective measures	Subjective interpretations	Provides answers to the question “why”	Support multi-perspectives
Quantitative (experiments)	/	/	/	X			
Quantitative (attitude surveys)	X		X	X			/
Surveys (CIT)	X				X	X	/
Interviews (CDM)					X	X	
Participant observation/ contextual inquiry/process tracing		X			X	X	/
F-REX	X	X	X	X	X	X	X
“X” indicates that the method fully embraces the characteristics; “/” indicates that the method sometimes embraces the characteristics.							

PROOF COPY ONLY**DO NOT DISTRIBUTE**

methods are summarized. As illustrated, F-REX is the most inclusive research approach in terms of embracing the listed characteristics.

Are the R&E approach, F-REX, and similar multimedia tools sufficiently mature for evaluating large-scale DTOs? Are traditional quantitative and qualitative methods not needed at all? The answer to both questions is no. Quantitative-controlled experimental designs and formalized studies indeed become increasingly difficult to carry out with the growing size and complexity of DTOs. On the other hand, quantitative methods do have some advantages such as their capacity to capture data from a very large number of respondents and determination of attitude tendencies from these data. Such an approach can be coordinated with the use of R&E and F-REX for dressing quantitative findings with qualitative data and thereby increase the understanding of them. As for qualitative research methods, the study clearly illustrated mutual dependencies. R&E/F-REX aided participant observation, by providing means for more effective reporting and by providing a rich and authentic data set that can be compared with other findings. On the other hand, when applied on a large-scale (dynamic DTOs) the approach will not be successful without the domain experts/observers who collect qualitative data out at the field and support the study design and the analysis of data. An ideal case might be a study design where the early phases of the R&E methodology is more emphasized, where R&E and F-REX support quantitative research, and where qualitative methods generate in-data to F-REX and support the analysis process by providing qualitative concepts and techniques for coding and interpreting data. Such a case is outlined in Figure 3. Hence, facilities for the coding of data need to be integrated in the F-REX Studio.

DISCUSSION

In this study, an R&E approach and a multimedia tool were applied to a large-scale nationwide exercise taking place in the Swedish armed forces. The study showed that the R&E approach in combination with the F-REX Studio has several advantages compared with traditional research methods, both quantitative

Figure 3. Intended R&E/F-REX approach. The colored area indicates the previous usage of the approach. For the future, more focus on prestudy issues (eg, the determination of hypotheses or research questions) and on retrospective data analysis (using qualitative and/or quantitative methods) is suggested.

and qualitative. They can support interpretation of voluminous attitude data and provide explanations to achieved results and rejected or confirmed hypotheses. They can complement retrospective qualitative surveys and interviews (eg, CIT and CDM) with factual data produced in real time. And, in the case of participatory observation and contextual inquiry, they speed up efficiency and provide a multitude of perspectives that can complement the observations made by individual observers.

As for resources and costs, a multimedia tool such as F-REX is not very resource intensive. It is portable and data can easily be stored on modern powerful personal computers. For visualization, powerful but regular computers with the possibility to run three to four screens are desired. Further, even though the R&E methodology requires a number of operators

PROOF COPY ONLY

PROOF COPY ONLY**DO NOT DISTRIBUTE**

and observers/experts to be effectively used in large-scale contexts, this number is still less than it would be to capture the same amount of data with traditional participatory observation or contextual inquiry.

On the other hand, the study showed that to apply R&E and F-REX in very large-scale contexts further development is needed. The approach is in certain respects pioneering; but there exist related tools both within media/television and within the military domain in the form of battle training systems or AAR systems.^{21,22} While the similarities include replay of events using maps, GPSs, timelines, and videos, these systems have the individual user in focus rather than embracing network, context, and communication. Further, none of the systems include a structured methodology, neither guidelines nor tools aiding analysis. R&E has a methodology, though in need of formalization. Carrying out such a formalization and adding clear guidelines and a solid analysis tool to the F-REX Studio would make it an even more useful and competitive approach when evaluating large-scale DTOs and exercises. In this sense, R&E and F-REX will also pave the way for future developments of similar systems, all designed to improve evaluation of training and real operations in the military and emergency management domain; thereby, in the long run contributing to a society that more effectively restores security in case of upcoming crises or attacks on human safety.

In this study, the R&E approach and the F-REX Studio were applied to a large-scale training event. The approach has earlier also been applied in real-live operations, but then in smaller emergency-related contexts (fire fighting operation). It is, for sure, possible to apply R&E/F-REX in large-scale real DTOs, but it is important to acknowledge some major differences between training sessions and real-life operations. Most crucial, there are severe security risks associated with real operations that do not exist in training, and a real operation cannot afford interruptions, breakdowns, and failures. This, of course, restricts the availability of material and personnel for evaluative purposes and further limits the possibility to intrude in activities or events. But also in this respect the R&E approach has some advantages compared with qualitative research methods in terms of

being less intrusive and interruptive. One might even argue that real operations enhance the benefits of using R&E because what happens during an operation in many cases is different from what happens during training and unexpected events will happen—the R&E approach enables capturing this reality.

Alternative usages of the F-REX Studio

Several suggestions as to the alternative usage of the F-REX Studio were put forward by exercise participants. One such suggestion concerned the possibility for individual actors to browse a filtered subset of the massive dataset with some assistance by F-REX operators to clear out obscurities, find explanations, and get fast feedback on their own performance during or directly after the exercise. In the future, this suggestion could evolve into a rapid analysis, which may be a more effective alternative to early AARs when handling masses of data in the large-scale context. Another suggestion concerned the use of the F-REX video display where four locations can be shown and listened to simultaneously. The suggestion was to place one of the observers in front of this display and let him or her have a real-time overview of what is happening out at the field. By doing this, individual observers can listen around, compare what happens at different locations simultaneously, and thereby capture the real-time course of events from a multi-perspective, something that is impossible with traditional observation methods. If one location is specifically interesting, it can be zoomed in and the whole conversation at that location can be listened to. However, it should be pointed out that “rapid analysis” and potential AARs should always be followed by in-depth analyses. Even if longer loops for analysis may initially involve certain additional time and costs, their outcome makes it possible to construct presentation material, which can be re-used for educational purposes. Through such lessons learned some mistakes in carrying out DTOs can be avoided, thus providing additional benefits in the long run.

CONCLUSIONS

The past decade’s frequent catastrophes world wide (eg, the tsunami in Thailand, 2004, the terror

PROOF COPY ONLY**DO NOT DISTRIBUTE**

attacks in New York, Madrid, and London, 2001-2005, the Katrina hurricane in New Orleans, 2005, and the forest fires in Greece, 2007) clearly remind of the need for the authorities, military, and emergency organizations to repetitively train, evaluate, and enhance their skills. Considering the fact that crises are becoming increasingly complex, involving numerous people, and requiring interorganizational and transnational collaboration, the demands on training and evaluation of large-scale exercises increase. Therefore, reconstruction approaches and multimedia tools like F-REX are most valuable and those costs that still are associated with applying them are highly motivated.

The unique combination of the R&E methodology and the F-REX Studio has an explicit focus on context, communication, and multiple perspectives that is in several aspects pioneering. Even if this study took place in a military setting, it focused on aspects that are relevant for DTOs, in general, and the results are thus highly applicable for civil crisis and emergency management operations. Others can learn from this study in pursuing their own development of tools and techniques that support R&E of DTOs or training sessions to create a better understanding of what is happening out there, thereby being able to perform more high-quality evaluations. The practical implications, in the long run, include improved, more effective and high-quality emergency management operations alleviating consequences in terms of material damage and human suffering.

FUTURE WORK

Future developments of the F-REX approach should involve more powerful and sophisticated technology, and a more formalized methodology. Presently, technical development and extension of F-REX are taking place using the C#.net. The development includes replacing the previous logging mechanism with a more powerful database-based framework for automatic data collection. Thereby, the capacity to capture data from multiple systems simultaneously will increase significantly. The F-REX replay mechanism has also been developed to enable more efficient navigation in large datasets to enable easier and faster reconstruction of mission histories and to provide a

quicker overview of large, complex, multiperspective events. Other technical issues to be solved include optimization of the logging tools and the need to overcome the issues with sound quality when many simultaneous voice sources are present. Related to this, more efficient ways to store sound, ie, by use of compression techniques like mp3, to remove the need of preprocessing incoming sound streams and to reduce disk usage, are investigated.

As for the methodology, work to incorporate meta-data with raw F-REX data has recently started. By adding the possibility to tag data from different locations, the analysis process will be much smoother because items can be matched and compared when trying to make sense of large datasets. The metadata component is a first step toward developing and integrating a solid analysis tool with F-REX. Work has also been initiated to make the R&E methodology as such further formalized with regard to the preparatory study design, the data collection, and the analysis phases. Finally, because the original version of F-REX is somewhat complex and requires training of its operators, work is presently done to provide a lighter version stored on a DVD and showing examples. Here, actors will be able to use the tool completely themselves after a tutorial. The light version will enhance the prospects for rapid analysis using F-REX and will support distribution of the F-REX suite to reach an increased number of people involved in large-scale DTOs and exercises.

Sofie Pilemalm, PhD, Department of Systems Development and IT Security, Swedish Defence Research Agency (FOI), Sweden.

Dennis Andersson, MS, Department of Systems Development and IT Security, Swedish Defence Research Agency (FOI), Sweden.

Niklas Hallberg, Department of Systems Development and IT Security, Swedish Defence Research Agency (FOI), Sweden.

REFERENCES

1. Thomas A, Kopczak L: From logistics to supply chain management: The path forward in the humanitarian sector. *Fritz Institute*. 2005. Available at www.fritzinstitute.org/PDFs/Programs/FromLogisticsto.pdf. Accessed January 5, 2008.
2. Albinsson P-A, Wong W, Pilemalm S, et al.: Using a reconstruction-exploration approach to empower cognitive analysis methods: Possibilities and challenges. In *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL, 2005: 224-228.

PROOF COPY ONLY

Journal of Emergency Management
 Vol 6, No 4, July-August 2008
DO NOT DISTRIBUTE

PROOF COPY ONLY**DO NOT DISTRIBUTE**

3. Wong WBL, Blandford A: Situation awareness in emergency medical dispatch. *Int J Hum-Comput Stud.* 2004; 61, 421-452.
4. Morin M, Albinsson P-A: Exploration and context in communication analysis. In Bowers C, Salas E, Jentsch F (eds.): *Creating High-Tech Teams: Practical Guidance on Work Performance and Technology.* Washington, DC: APA Press, 2005.
- AQ8 5. Andersson D, Pilemalm S, Hallberg H: Evaluation of crisis management operations using reconstruction & exploration. In *Proceedings of the International Community on Information Systems for Crisis Management Response and Management (ISCRAM)*, 2008.
- AQ9 6. Morin, 2002.
- AQ10 7. Pigeau R, McCann C: Redefining command and control. In McCann C, Pigeau R (eds.): *The Human in Command: Exploring the Modern Military Experience.* New York: Kluwer/Plenum, 2000: 163-184.
8. Byrne D: *Interpreting Quantitative Data.* London, UK: Sage Publications, 2002.
9. Baron S, Kruser DS, Huey BM (eds.): *Quantitative Modeling of Human Performance in Complex, Dynamic systems,* Washington, DC: National Academy Press, 1978.
- AQ11 10. Denzin NK, Lincoln YS (eds.): *Handbook of Qualitative Research.* Thousand Oaks: Sage Publications, CA, 2000.
11. Flanagan, JC: The critical incident technique. *Ment Bull.* 1954; 51, 327-358.
12. Dean PJ: A qualitative method of assessment and analysis for changing the organizational culture. *Perform Improv.* 1998; 2(37): 14-23.
13. Hoffman RR, Crandall B, Shadbolt N: Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis. *Hum Factors.* 1998; 40(2), 254-276.
- AQ12 14. Beyer H, Holtzblatt K: Contextual design. *Interactions.* 1996; (1): 32-42.
15. Woods DD: Process-tracing methods for the study of cognition outside of the experimental psychology laboratory. In Klein GA, Orasanu J, Calderwood R, et al. (eds.): *Decision Making in Action: Models and Methods.* Norwood, NJ: Ablex, 1993: 228-251.
- AQ13 16. Sanderson P, Scott J, Johnston T, et al.: MacSHAPA and the enterprise of exploratory sequential data analysis (ESDA). *Int J Hum-Comput Stud.* 1994; 41: 633-681.
17. Mackay WE, Beaudouin-Lafon M: DIVA: Exploratory data analysis with multimedia streams. In *Proceedings of the International Conference for Human Computer Interaction.* 1998: 416-423.
18. Albinsson et al., 2004.
19. Vasend G: After action review system development trends. In *Proceedings of the 27th Conference on Winter Simulation,* Arlington, VA, 1995; 1262-1266.
20. Darling MJ, Parry CS: After-action-reviews: Linking reflection and planning in a learning practice. *Reflections.* 2001; 2(3): 64-72.
21. Allen G, Smith R: After action review in military training simulations. In *Proceedings of the 26th Conference on Winter Simulation,* Orlando, FL, 1994.
22. Gratch J, Mao W: Automating after-action-review: Attributing blame or credit in team training. In *Proceedings of the 12th Conference on Behavior Representation in Modeling and Simulation,* Scottsdale, AZ, 2003.
23. Ford R, Gallagher V, Giuli R. et al.: The JTEP takehome package: An AAR tool for distributed training. In *Proceedings of IEEE 2004 Fall Simulation Interoperability Workshop,* 2004.
24. Stake, 2000.
25. Yin RK: *Case Study Research. Design and Methods.* Thousand Oaks, CA: Sage Publications, 1994:1.
26. Bernard, RH: *Social Research Methods. Qualitative and Quantitative Approaches.* Thousand Oaks, CA: Sage Publications, 2000.
27. Taylor SJ, Bogdan R: *Introduction to Qualitative Research Methods.* New York: Wiley, 1998.
28. Seidman I: *Interviewing as Qualitative Research. A Guide for Researchers in Education and the Social Sciences.* New York: Teachers College Press, 1998.
29. Stewart DW, Shamdasami PM: *Focus Groups. Theory and Practice.* Sage Publications, 1990.
30. Thorstensson M, Albinsson P-A, Johansson M, et al.: *MARULK 2006—Utvecklingsmetoder för funktioner, förband och system [MARULK 2006—Development Methods for Functions, Military Units and Systems].* Linköping, Sweden: FOI-T-2188-SE, 2006.
31. Courtney JD: Automatic, object-based indexing for assisted analysis of video data. In *Proceedings of the 4th ACM international conference on Multimedia,* Boston, MA, 1996: 423-424.
32. Kiekel PA, Cooke NJ, Foltz PW, et al.: Some promising results of communication-based automatic measures of team cognition. In *Proceedings of the Human Factors and Ergonomic Society 46th Annual Meeting,* Santa Monica, CA, 2002: 298-302.
33. Deutsch S: International experiences with technological change. *Mon Labor Rev.* 1986; 106: 35-40.

AQ14

PROOF COPY ONLY**DO NOT DISTRIBUTE**

Author Queries

- AQ1: Kindly provide the degree/educational qualification for “Niklas Hallberg.”
- AQ2: Kindly check whether ALL section heads and subheads in the manuscript have been set as intended.
- AQ3: Please note that the reference citations have been changed from name-year system to citation order (per journal style).
- AQ4: Please check the usage of “cost some” in the sentence beginning “Using qualitative methods...”.
- AQ5: Please check the edits made to the sentence beginning “The study was performed...”.
- AQ6: Please check the edits made to the sentence beginning “It is typical...”.
- AQ7: Kindly spell out GPSs.
- AQ8: Kindly provide the complete page range for reference 4.
- AQ9: Kindly provide the location where these proceedings/workshops were held (for refs. 5, 17, and 23).
- AQ10: Please note that references 6, 18, and 24 were cited in the text and the details were not provided in the list. Kindly provide the complete details of these references.
- AQ11: Please note that in the original name-year reference citation “Baron et al., 1990” was provided, whereas the year “1978” was given in the list (ref. 9). Kindly check.
- AQ12: Kindly provide the volume no. (if possible/available) for ref. 14.
- AQ13: Kindly provide all the editor names for this reference (ref. 15).
- AQ14: Please note that “Deutsch 1996” (ref. 33) is not cited anywhere in the text. Kindly insert its citation at an appropriate place or delete it from the reference list.

PROOF COPY ONLY**DO NOT DISTRIBUTE**