Enhancing Learning from Incidents by Reconstruction of Events: Using the SQUARE Tool for Evaluation

Jaziar Radianti
Department of Information Systems,
University of Agder
jaziar.radianti@uia.no

Sofie Pilemalm
Department of Management and Engineering,
Linköping university
sofie.pilemalm@liu.se

Kristine Steen-Tveit
Department of Information Systems,
University of Agder
kristine.steen-tveit@uia.no

Kjetil Rustenberg
Department of Information Systems,
University of Agder
kjetil.rustenberg@uia.no

ABSTRACT

Evaluation after exercise and incidents—often called After Action Review or AAR—are important to enhance the emergency management (EM) response capability. However, evaluation support tools for event reconstruction after incidents and exercises are not yet fully available. We engaged EM stakeholders in a tabletop exercise based on snow chaos and car accidents scenarios to evaluate a fully functional prototype. The tool assists information sharing in real-time and enhances debriefing process of any EM response and exercise. Analysis of reflective discussions and an open question survey shows that the EM participants consider the features of the prototype to facilitate better learning from incidents. However, they have different attitudes concerning the adoption, management, and application of the tool in organizations. For instance, more security features are required to comply with regulations. We argue that the tool is an important first step to fill the gap on the need for "event reconstruction-based evaluation."

Keywords: Learning from Incidents, Event Reconstruction-Based Evaluation, Tabletop Exercise

INTRODUCTION

In recent decades, we have seen an increase in the frequency and severity of disasters, requiring complex response operations involving extensive multi-organizational collaboration. In these settings, emergency response organizations must collaborate to take effective decisions, execute life-saving measures, and prevent major damage. The extant studies concur on interconnectedness between situational awareness (SA), common operational pictures (COP), and common situational understanding as crucial factors for effective multi-organizational emergency management (EM) (Bunker et al., 2015 ; Danielsson et al., 2014 ; Giaoutzi and Scholten, 2017 ; Steen-Tveit and Munkvold, 2021 ; Wolbers and Boersma, 2013). The organizations involved must make quick and correct decisions at the different command and control structures, which is a challenging task (Bharosa et al., 2010 ; Boin and 't Hart, 2010).

Previous studies signalize the crucial role of training, EM exercise and learning from incidents (LFI) to apprehend the actions, decisions and COP occurred during the events that often constitute agency-specific elements and different perspectives. The theory of organizational learning proposed by Argyris and Schön (1978) pinpoints the importance of learning as a way to detect and respond to unwanted situations. There is a continuous development of organizational learning in the literature that highlights the multidimensionality of the concept e.g., sensemaking (Weick et al., 2005), knowledge flow and transfer (Nonaka and Takeuchi, 1995), learning culture (Yanow, 2016),

reflective practices (Greenwood, 1998) and more. Regardless of this advancement, Argyris and Schön (1978) introduce the idea of learning as a process for acquiring and storing information, which occurs in multiple levels, known as single-loop and double-loop learning. The latter is the learning process in which an individual or organization is able to reflect upon, question, and modify the goals, values, assumptions, and policies that led to certain actions. This information can be manifested as lessons-learned from others. In an extensive survey on LFIs and organizational learning theories, Drupsteen and Guldenmund (2014) argue on lacking attention towards LFIs literature that illuminates the need to share and store lessons learned.

Similarly, Pilemalm et al. (2021) argue on the existence of continuous challenges on LFI practices for EM, to make sure the lessons are actually learned, and the organizations are capable to reflect upon the LFIs, and able to act accordingly. That is, to modify the goals, values, assumptions, and policies that lead to certain actions- or a double-loop learning. Thus, in this study, LFI implies gathering information from the individual(s) and organization(s) involved in an incident, and from the incident itself, or from EM exercises and converting it into general knowledge for the whole organization (Jacobsson et al., 2011). This can be accomplished through systematic, documentable processes and reflective practices.

In practice, during an emergency response, the message exchanges between the responders can be recorded, e.g, in the secure radio TETRA network (Borglund and Granholm, 2020). However, using only audio-log capability in the evaluation is not trivial and not obvious concerning information elements to extract, time-consuming, unpractical for short-term purpose (Steen-Tveit and Radianti, 2019 ; Steen-Tveit et al., 2020), and indeed, it has not yet embraced in current AAR practice.

This study is motivated by a workshop result conducted in Norway for eliciting requirements from national, regional, and local stakeholders EM stakeholders for better COP and information sharing where respondents pinpoint the issue of missing "fact-based evaluation" in current LFI practices (Munkvold et al., 2019). This is also confirmed in previous research (Andersson et al., 2008). Fact-based here refers to the ability for one to go back in time and observe what has actually happened during the exercise or real response, including what has been shared by different actors, i.e., not solely rely on fresh or modified memories of what has happened and what has been done. Thus, the ability to "share" and "store" as a part of learning as suggested by organizational learning theory, seems to be a crucial issue to be concretized in EM practices. In the study context, a fully functional prototype of COP-based evaluation has been realized. The innovation value of the prototype tool called SQUARE, is among other things, its ability to address "fact-based" evaluation which we refer to as "event reconstruction-based evaluation". This would imply a replay functionality, including, time stamps, maps, logs, symbols, and descriptions of what factual elements users can access quickly in retrospect, with double-loop learning as a goal in mind (see study context).

This study reports a tabletop exercise as a medium to test and evaluate the prototyped SQUARE tool with EM stakeholders from several organizations. Information sharing, event reconstruction-based evaluation, and enhanced LFIs were the core purpose of the exercise. The aim of the study is thus to examine if the existence of a new tool with event reconstruction-based evaluation features can facilitate better LFIs and potentially trigger double-loop learning. This study thereby addresses two research questions. First, how do emergency management responders experience the SQUARE tool for enhancing LFI and map-based COPs? Second, what enablers or barriers can facilitate or hinder its adoption and use in EM day-to-day work?

THEORETICAL BACKGROUND

The study's theoretical background is inspired by organizational learning and, in particular, double-loop learning, and LFI. Early work on organizational learning distinguishes between single- and double-loop learning. The latter is the learning process in which an individual or organization can reflect upon, question, and modify the goals, values, assumptions, and policies that led to certain actions (Argyris and Schön, 1978). Subsequently, a four-stage spiral model of organizational learning was developed by Nonaka and Takeuchi (1995). They include the aspects of tacit knowledge vs explicit knowledge and describe a process of alternating between the two of them. Tacit knowledge is personal, specific, and subjective; whereas explicit knowledge is codified, systematic, formal, and easy to communicate. The tacit knowledge of key personnel within the organization can be made explicit by processes of socialization, i.e., acquiring the tacit knowledge of others through interaction. Next, the acquired knowledge is articulated and created in a corresponding process of externalization. Finally, the knowledge can be incorporated into new products and processes for later internalization, i.e., embodying externalized knowledge in employees and/or applying it in practice.

Organizational learning has been repeatedly portrayed as a continuous cycle of action and reflection as a part of

double loop learning (Argote and Miron-Spektor, 2011), that often involved the lessons-learned approach. The "lessons-learned approach" is one such example of knowledge management practice applied to the wider concept of organizational learning. It has been referred to as learning from incidents (Drupsteen and Guldenmund, 2014), from crisis (Renå and Christensen, 2020), or from disasters (Birkland, 2009 ; Choularton, 2001). In emergency management, lessons-learned as an approach concern "wider" learning from experience: whether it be our own experience or others', and whether it be from real events or simulations, improves practice and minimizes losses of human and properties (Savoia et al., 2012).

Scholars have articulated the relationship between disasters and learning in the lessons-learned context (Donahue and Tuohy, 2006). Various mechanisms for sharing experience have emerged. The "lessons-learned" mechanisms include tools like in-progress reviews, AAR, "hotwashes," and various kinds of debriefings. Regardless of these process variations, their sole goal is to share performance information to prevent the recurrence of adverse events and to better act if the situations are repeated in the future. Most processes involve systematic evaluation (what happened, why), *identifying lessons* (strengths weaknesses of actions), and *learning* i.e., sparking behavioral changes consistent with the lessons. However, questions such as how to evaluate without leaning entirely on memory, and how lessons-learned is preservable, retrievable, and reusable in organizations to promote preparedness for next incidents, have not been fully addressed in the literature (Drupsteen and Guldenmund, 2014). Our research offers the idea of event reconstruction-based evaluation to fill this gap.

Moreover, Greenwood (1998) underlines a reflective practice as a core of organizational learning, which contains diverse interpretations of this term meaning. Schön (1987) suggests that reflective practice should involve *reflection-in-action* and *reflection-on-action*. Reflection-in-action implies thinking about what one is doing while one is doing it which can be triggered by surprise or by something puzzling. Moreover, *reflection-on-action* involves a 'cognitive post-mortem' where an actor reviews her/his actions to reestablish new insights to be used in the actions in connection with their outcomes (Greenwood, 1998). Scholars have suggested additional elements such as reflection-before-action (Greenwood, 1993) and returning/ attending to feeling and re-evaluating the experience (Boud et al., 2013). Summarily, reflective practices for single and double-loop learning insinuate preparation for experience: what the learners bring to the event and what they want from it (the personal aspect), what constraints and opportunities the event provides (the context) and how learners can acquire what they need from the event (the learning strategies). However, Greenwood (1998) also points out that the aim of double-loop learning is not apparent in many frameworks for reflections (i.e., series of questions or guidelines to structure reflections on actions).

Relating, this general notion to information systems for crisis management, previous research has claimed that tools for EM evaluation that are able to record and replay events and chains of action in an exercise or real crisis operation, will enhance the prospects for reconstruction-based evaluations. This will change the focus in evaluations, e.g., AARs and debriefings to explore *why* something happened rather than *what* happened. This, in its turn, will enhance the prospects for double-loop learning (Pilemalm et al., 2014 ; Pilemalm et al., 2021). In addition, such tool serves as a medium for triggering reflective learning in all aspects, identifying the lessons, and maintaining lessons-learned.

Of relevance to this study, we thus find that previous studies acknowledge the importance of single-loop and double-loop learning for organizational learning in general and learning from crisis and incidents, specifically. However, how to actually implement the learning in practice, such as reflections for lessons to be learned (Greenwood, 1998) remains a challenge. Previous studies rarely suggest evaluation technological support and frameworks that allow organizations to evaluate outcomes based on reconstruction of events, i.e., actions taken and their timestamps in real-time that can be reused quickly in the aftermath of an incident or exercise, and then stored and retrieved. Our study sets out to address this gap by proposing a prototype of evaluation support technology that can facilitate crisis-induced single and double-loop learning.

STUDY CONTEXT

This study reports the results from the INSITU (Sharing incident and threat information for common situational understanding) project. In earlier related research (Pilemalm et al., 2021), we have described the requirements gathering, needs, and requirements related to learning from EM incidents. The requirement gathering involved various EM organizations from regional and national levels and across sectors. We have also described the general design principles for the SQUARE tool for information sharing and learning. In this study, we move on to prototype implementation, user testing, and evaluation of the tool. The requirements on LFI included are event reconstruction-based evaluation objective information, a repository of previous evaluations, and a replay function including a dynamic incident timeline, with time scales, time logs and stamps, and continuously updated COPs that monitor the crisis development, to be able to reconstruct events and do systematic follow-ups. It also includes

a decision repository and a COP with the capability to aggregate reports, generate action plans from reports, and extract statistical data and a repository for lessons identified. At the organizational level, requirements concerned inter-organizational after-action reviews, vertical and horizontal evaluations at the management and/or operational level, and evaluations across sectors. At the structural level, requirements related to simplified national guidelines to develop common and regular routines for evaluations, to focus more on best practices, and to develop standards for evaluation, regulations of how information should be stored, owned, and distributed, and similar evaluation methodologies across organizations for synergy effects.

Based on the most reported requirements among stakeholders, the web-based application SQUARE was developed. The focus is on the requirements for replaying incidents to enable event reconstruction-based evaluation. A replay function for information sharing, chat communication, and decisions taken during the incident, displayed at various speeds based on a timeline indicator is included. It is possible to freeze the timeline and take screenshots if you want to explore something in more detail. When a user replays the event, s/he can register evaluation notes (e.g., something went wrong here, why?) that can be used in later evaluations, thus reflecting the requirement for a repository of lessons-learned. At a more detailed level, the functions in SQUARE are described in Table 1. A snapshot from the replay function is displayed in Figure 1.

Table 1. Realization of the SQUARE Solution for Evaluation and Event Reconstruction-Based Evaluation Learning

Features in the SQUARE Solution. A user can:
Share locations, make use of symbols representing organizations, actions, or situations. SQUARE contains a repository of EM symbols for information sharing. Any objects that have been placed underpin various important facts (events, decisions, actions). A chat function allows real-time textual documentation of the users´ SA and the reasons for their actions, which support event reconstruction in evaluation.
Invite other relevant stakeholders to the emergency response quickly when they receive a notification (confirmation code) through his or her mobile phone/PC and login instantaneously to the event response. Such a rapid process ensures the synchronization of information across agencies vertically and horizontally with little delay, also granting access to the evaluation repository, whenever it is made from time to time.
Use a slider that can bring the user back in time and replay the timeline of response and critical decisions and examine previous action at a specific point in time, or overall overview.
Save the decisions represented by symbols and timestamps, and possible textual information. Thus, whenever an organization places an object in a specific location in the SQUARE map, it will be automatically saved and can be seen later.
Document the points being discussed during the debriefing and evaluation meeting (e.g., after-action-review) in the SQUARE application.

Figure 1 Replay function in SQUARE

METHODOLOGY

We applied a qualitative research approach which is common research methods when the aim is to gain deeper understanding of complex phenomena (Myers and Avison, 2002). The study can be considered as interpretive seeking to produce and understand the context of IS for emergency management (here: the SQUARE tool for evaluation) and the process whereby the IS influences and influenced by the context (here: is it feasible to apply in organizations?). We have implemented the process of requirement gathering the requirement from EM stakeholders concerning the needs for event reconstruction-based evaluation /map-based evaluation tool that supports debriefing after EM exercises and response, and how it was transformed into the SQUARE evaluation tool that is intended to support reconstructions of events and thereby contribute to better LFIs. Figure 2 illustrates

the overall research process, depicted as Phase 1 (Requirement gathering and development activity), and Phase 2 (Evaluation combined with data collection activities). The research reported in this article derived from the activities in Phase 2.

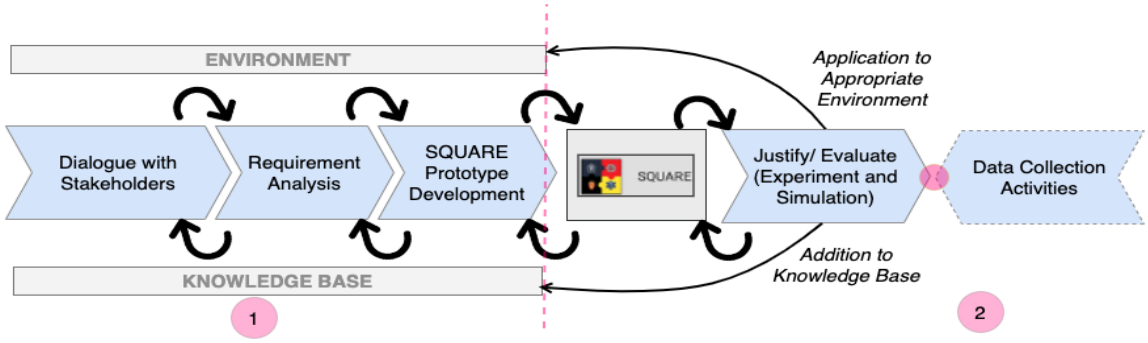

Figure 2 Research Phases and Study Coverage

Data Collection: Scenario-based Table-top Exercise

To evaluate the tool, we created a scenario-based table-top exercise that serves as a simulated environment and engaged professional practitioners to allow them to assess the value of the enhancement of the evaluation process. Thus, two parallel processes were prepared for data collection: To design a tabletop exercise that allow the participants to experience a tool that provide a possibility to reconstruct events and decisions, and thus promote LFI instead of relying solely on memory on actions taken. Also, to collect data on the exercice using common techniques in qualitative research, i.e., focus group discussions (FGD) and survey.

The purposes of the tabletop exercise in our study are twofold: First, to allow the practitioners *to test new technologies/SQUARE and practice new ways of working (together across organizations/sectors)*. Scholars have argued on appropriateness of tabletop exercises in the EM context e.g., to test and measure performance, preparedness, procedures, and competencies in rare events (Agboola et al., 2013 ; Savoia et al., 2009a ; Savoia et al., 2009b) to teach and promote new ways of thinking and acting, and to train the interaction of multiple actors in complex systems (Lasky, 2010 ; Tobergte et al., 2022)—to name a few. Our tabletop exercise was thus designed to facilitate familiarization of technology for reconstruction of events/evaluation testing and enable reflections of participants on the applicability of the features described in Table 1. Second, *tabletops can serve data collection activities*. The potential of collecting data in simulated settings has been discussed widely in the literature including tabletop exercises, simulations/ role playing/ serious games etc. (Radianti et al., 2015 ; Smith et al., 2015). Smith et al (2015) pointed out that interviews, FGD, questionnaire, direct and indirect observations have been used for data collection simulated environments, which is also applicable for tabletop exercise settings.

Scenarios and Flow

We developed two different emergency scenarios. The first scenario was called “Snow Chaos” capturing a real local emergency occurring in the winter of 2021 (Cantero et al., 2021). The first scenario started two days before the exercise date. The first alert sent to the participants included a weather forecast. Following this, different alerts consisting of situational-related information were distributed to the participants over the next two days. This approach provided participants the opportunity to plan for possible emergency events in the SQUARE application. The participant could develop a common plan for EM early and establish SA and a common situational understanding in advance. On the actual day of the exercise, we played out the “Snow Chaos” scenario as it developed into an emergency situation and several decisions and actions were required by the participants. Examples of elements included in the scenario were a closed road for several hours, electric cars being out of electricity, and the need for evacuation.

The second scenario is called “Traffic accident -explosive cargo” and involves a front-to-front collision between a truck and a passenger car. The truck´s cargo was loaded with explosives which can cause a fire hazard. After the alert phase, an animated film was distributed to the participants and indicated an overview from a drone (Figure 4). Examples of elements included in the scenario are injured people, liquid natural gas cargo, demanding availability, and the need for evacuation.

The exercises started with the “Snow Chaos” scenario and lasted for 1.5 hours including AAR. The AAR was following the evaluation format in SQUARE and was conducted with all participants present. The key elements

for the evaluation were situational awareness during the build-up of the scenario e.g., sequence of events, how the stakeholders interpreted the information in SQUARE, critical information, and how the participants can act/share/enter things in the tool.

Figure 3 Animated film of "Traffic accident

The second scenario also lasted 1.5 hours and incorporated the evaluation features of SQUARE. The key elements for the evaluation were SA, critical information such as where the vehicle's equipment is located, and how the participants can act/share/enter things in the tool. The participants were given access to SQUARE and a brief tutorial on how to use it a week before the actual exercise.

Participants, Survey and FGDs

The participants were divided into two groups and physically sat together in two separate rooms. The incident commander team included stakeholders from the first-responder agencies (see Table 2) and occupied a simulated command and control center. The stakeholders from the supportive organizations utilized a classroom with their personal computers and a big screen. The exercise involved two different scenarios (see previous section), lasted for four hours, included AAR for both scenarios and a common evaluation of SQUARE.

Table 2: Participants of the Tabletop Exercise

Organization	Role	Role in the exercise
Fire and rescue services	Emergency dispatcher and coordinator	First responder agency
Police services	Officer	First responder agency
Ambulance services	Emergency dispatcher and coordinator	First responder agency
Ambulance services	Incident commander	Incident commander
Road authority	Advisor and emergency manager	Emergency manager
Road authority	Advisor and emergency manager	Emergency manager
County governor`s office	Emergency manager	Emergency manager
Municipality	Emergency manager	Emergency manager

The data collection was performed through three FGDs encompassing three parts: discussion/AAR after the first scenario, discussion/AAR after the second scenario, and the reflection session on overall learning during the tabletop exercise. In total, we had around 2 hours of discussions. The verbal conversation was transcribed. In addition, we circulated a survey that covered questions that we did not ask during the discussion sessions.

Examples of questions used during the AAR that are related to functionalities of the tool are: "how could the replay feature be used for evaluation purposes?" Many of the exemplary questions were intended to require *reflective answers* e.g., "what knowledge could you get that you don't have today? How to take the knowledge further into your organization? What in the situation picture/replay do you see specific possibilities in? What functions are not needed for evaluation/learning purposes?

The survey questions included open questions. The survey comprised: 1) questions related to *the preparation phase before using the tool*, such as: "were the training videos effective for learning how to use the system?"; 2) questions on *the learning experience and perception change*, such as "how did the ability to communicate both visually via the map and the chat affect your situational awareness?" 3) questions related to *the degree of reflection*, e.g., "how do you think the SQUARE evaluation module can be used in cooperative exercises?"; 4) question concerning *the opinion on possible change in organization through the adoption of event reconstruction-based evaluation tool,* such as: "do you have any suggestions for how the evaluation function can be incorporated

further into your organization?". We also sent the participants the exemplary SQUARE-generated evaluation reports, to facilitate answering the survey. Because, during the actual tabletop exercise, they have limited time to see deeper all the innovative features of the SQUARE tool.

Data Analysis Method

We applied thematic analysis, which typically involved the interpretation of meaning in the text. In specific, we used exploratory thematic analysis in which specific codes or analytical categories are not predetermined but directly derived from the data (Guest et al., 2011). The categories were then coded into themes. In presenting the results, we disregarded if the information emanated from the first, second or the after-action evaluation sessions. We put them together whenever they are in the same theme.

RESULTS

In this section, we present the findings into two subsections: First, the experience working with event reconstruction-based evaluation tools. Second, enabler and barrier factors for adoption, where we clustered the information related to the possible use of the SQUARE technology on a day-to-day basis in the participant's own organization.

Experience of Working with SQUARE

The experience of working with SQUARE can be divided in different sub-themes, presented below.

Working in Real-time with Quick Information Sharing

The tabletop participants regarded SQUARE as meritorious for giving full-transparency of the emergency event management, enabling users to place map-objects when needed, and thought it could potentially provide better SA. Indeed, not all features were deemed excellent, as they experienced some inconvenience. An appreciated feature, for example, was that they managed to get in touch and could confirm or deny the placement of symbols from the command control center (e.g., 110, 112 or 113) based on the map, while chatting with the operator. Relating to confusion, no features that allowed them seeing a continuous list of events/ activities added to the map. It was difficult to see which objects were new, or what objects/ events the user needed to pay extra attention to. One of the participants also said that "This was a bit difficult, since you can only choose one thing at a time. The chat should be linked to the events on the map. Otherwise, the idea works very well to be able to form a situational understanding." They also added that the solution was optimal for the web. However, it was more difficult when using mobile phone as they would like to chat while working on the map (multitasking on phone). It had been useful to get input from other actors along the way, they said. On the other hand, they regarded a quick chat function as not yet being realistic at the operational level for those who physically and directly handle the situation.

Making Inaccessible Information Accessible

We encouraged the participants to envision the benefits of having access to information that they normally would not be unavailable for the evaluation, such as shown in the tabletop exercise. Timeline, chat possibility and positioning Point of Interests (PoIs), timestamp of the actions from other actors and pictures from the event were deemed as useful features. The participants were positive towards the evaluation template feature, its usefulness and comprehensiveness, especially in acquiring a clear timeline, both in the replay function or in the self-generated report. However, they wished for "live timeline" that flowed during the events and contained a flag when important events appear. Our tabletop exercise was deemed to give considerable information for users, in contrast to existing solutions in current EM practice where their information originated solely from their own agency. However, not all participants shared this opinion, as some of them spotlighted the prototype´s shortage such as the quality of evaluation heavily depended on the quality of the logged information. Relating to the participants´ previous experience on AAR, it showed that information about "when" and "who" often were deemed deficient. Here, SQUARE could overcome this limitation. Having a receipt scheme after someone posted information would improve the usefulness of SQUARE as one could easily track whether the information has been processed.

Self-generated Report and Fact-based Information for Evaluation

Most participants agreed that a self-generated report feature of SQUARE and information on the reconstructed events could contribute to the AAR evaluation. "When all the facts have been documented, it will be easier to

find learning points", claimed one of the participants, "especially if everyone writes the information in a SQUARE document at the right time". The SQUARE-generated report provided a better objective picture of what happened and a broader understanding of the whole situation, with logs from all agencies. Their previous evaluations would often be composed of several sources, where time and place were not specified, which provides risks for deliberate manipulations or errors. The participants believed that the tool could ease reporting issues, especially the compilation of various sources (emails, phones, logging systems) that come from an event. The tool would provide a useful value for overall exercise evaluation, and assessments in their own individual agencies.

Timeline Function: COP in the Evaluation

As previously described, the innovative SQUARE timeline function allows a user to return in time and see the development of the COP over time from the perspective of different agencies. Concerning this specific function, the participants deemed the timeline function useful, enabling users for moving back and forth and actually seeing the passage of time and graphical representation of the event progression. It also provided an overview of the COP development and enhanced the understanding of the situations during the evaluation of an event. "This makes it easier to see where the "time thieves" were or explain the reason for the time spent", explicated by one of the participants. They also saw the advantages of who has chatted (and entered a task) or Point of Interest (PoIs) in the map, and possibly when an event has ended, i.e., the PoI was removed again. In a long effort on the spot, it is easier to recreate the event with a timeline.

Chronological Order of Map Objects and Chat Logs

Relating to the possibility to list all map objects and logs from the chat in chronological order, the participants deemed such features as a help to the evaluation after exercises or events. However, they would like to have a summary of the overall setup from various actors in the event. The positive opinion for such objects was that these contribute to providing a good overview of the entire event, i.e., create a comprehensive SA of all actors: "Who knows what and when". Such a timeline of the information provided knowledge from the latest known information until action was taken and could tell if the measures were adequate and time spent was efficient. In short, they considered this timeline feature gave many advantages as listed above and gave "less vision and more facts" in an evaluation.

Barriers and Enablers for Adoption

As illustrated above, the participants considered many features of SQUARE useful, facilitating a continuous understanding of the situation. However, a doubt raised if the police would have the time/ resource to update such a system during the incidents. At the same time, they confirmed that the SQUARE map with information sharing capacity and chat feature was crucial for timely and correct SA. The participants should have familiarized themselves with the tool before they could assess the value of the tool in giving a better overview of time and space during an incident. The participants confirmed that the exercise has given them a greater understanding of the other different actors working on the same situation, including the cooperation between emergency services and municipalities.

The participants emphasized that the exercise would be more useful if it had been a full-day exercise or even for a longer period. The tool might be even more valuable in a longer crisis e.g., to simulate a forest fire that lasted for several weeks. In such a scenario all actors would be separated (except for those who usually would be co-located). It would potentially bring more points to the surface, as the participants could not talk together "across the table". Moreover, they would not overhear information that one would not have otherwise, unless this information has been communicated in maps or chat, which demonstrated in a tabletop exercise setting. The realism issue was also a point mentioned by the participants, especially concerning the speed of information appearing over time. There would be more delays or waiting time, thus, having a long exercise time would help to achieve this realism. It is worth mentioning that our tabletop exercise did not include higher rank personnel in the command control chain (e.g., 110 for fire services, 112 for the police, and 113 for the ambulance). The participants considered that if we could have involved these higher rank personnel, their presence would give extra advantage for the whole teams participated in our tabletop exercise, as it creates more realism to overall decision-making process from time to time.

Effective Use of Event Reconstruction-based Evaluation Module

The participants were asked to assess the situations where to use the event reconstruction-based evaluation module will be most effective. They mentioned the examples that can be relevant, e.g., to use it in complex events with many actors, and in all kinds of events that should be evaluated. It also can be applied in previous major event

experience that might not have been as successfully evaluated as one would like to see due to limited information, or in long, challenging mission where their own agency has been involved over time. The participants considered that the tool should be used right after the event and work best in exercises where there are several collaborators who have worked together (different schools of evaluation) to provide a holistic way of evaluating.

While the participants showed positive attitude toward the report module including the documentation possibility to not forget the important learning points, they also wished to have a simplified version of the template before starting of the evaluation. Then they would be able to choose between complete and simplified versions of the evaluation template.

Promoting Fact-based Evaluation in the Own Organization

We asked the participants to reflect on how it would be possible to promote useful evaluation features to be adopted by each organization of the respective exercise participant. One of the most important factors mentioned was that the tool should be presented in the management meeting and be used by default for all events by all sections or departments in the organization. It could be used for refining EM plans and ideally be included in the quality system. In addition, a new tool like SQUARE should also be practiced by those who are responsible for evaluation after an event or disaster. One of the skepticisms related to the more formalized evaluation through the application was situations where there would be many players who should enter their inputs for the evaluation as they rarely got long time, e.g., with leadership, in such processes.

Replay Function and Other Issues

On the opportunity of using the replay function to improve AARs and evaluations, one participant did not directly discuss it, and rather highlighted the importance of assessing what is useful for other resources to know: "Whether the information was relevant to other actors, and whether it possibly answered other actors' needs for information. What information is relevant to the other agencies, or concerns only your own?"

In relation to this statement, they wanted the possibility to first take the preview for assessing the usefulness of the information for others outside his or her agency, before actually sharing them. They are also concerned about the relevance, level to distribute the evaluation, and timeliness. Moreover, handling sensitivity of information being shared to other agencies, authentication issues, the duty of confidentiality, and verification of information should be made possible. Indeed, such detailed features should be addressed properly in the mature version of the application.

DISCUSSIONS AND LIMITATIONS

Having a tool that facilitates documentation of actions, evaluation and confronts the participants on what really happened would lay-down the foundation for the organization's double-loop learning. Organizational changes after exercises are often a challenge to implement because the perceived effect of the exercises itself can be hard to conclude on (Sorensen et al., 2018). Earlier, we referred to the literature that have underscored single loop and double-loop learning as the core of organizational learning (Argyris and Schön, 1978). Single loop learning is achieved when the organizational members detect and correct discrepancies and flaws in organization and procedures, without inquiring into basic organizational premises and norms. Basically, in a single loop learning, thus, an organization can continue carry out their present policies and objectives. This is possible because of slow changes of the environment.

However, in times of rapid changes, including crisis management, there is often (an unsustained) need to achieve double-loop learning where learning inquiries may trigger the need for restructuring organizational norms strategies and assumptions of such norms. The actors need to to detect and correct errors by modifying or even discarding the underlying norms, policies and objectives in organizations—if necessary, and adding new ones. Despite our tabletop exercise did not directly detect errors, but in one of the findings, the participants expressed positive reaction when they actually can see more information than they used to have during the even (i.e., information on own organization only).

We have suggested a novel tool to support evaluation in the organizational learning context as can be seen in Figure 4, as typical of the double-loop learning. In other words, we introduce the SQUARE tool and argue that it can contribute in different parts of the organizational learning and support both single and double-loopdouble-loop learning. The *evaluation formalization loop* is the requirement for being able to reflect and learn, not only at the "after event" level, but also at organizational level (better planning).

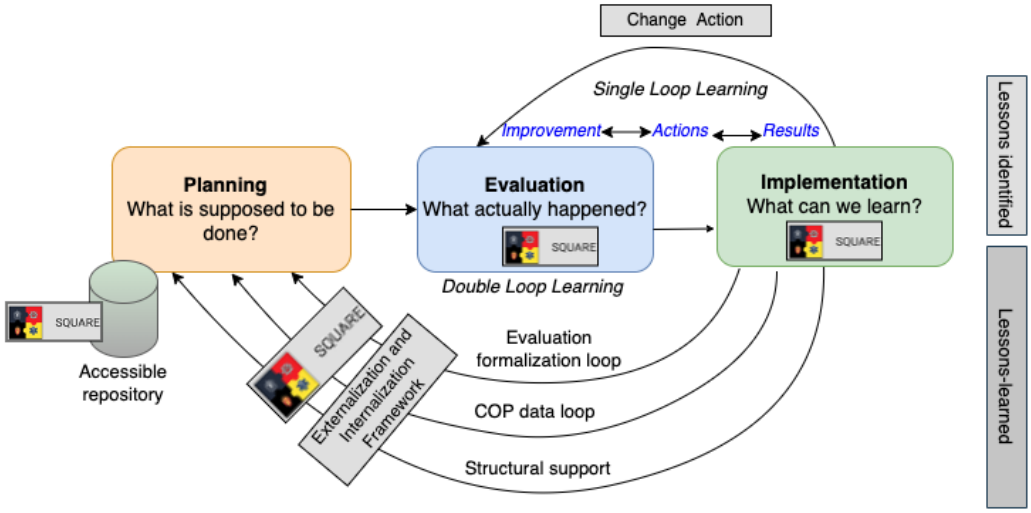

Figure 4 Double-loop learning with the support of SQUARE

In double-loop learning, we may ask questions such as what is supposed to be done (Planning), what actually happened (Evaluation) and what can we learn (Implementation). The proposed tool is prepared not only to answer the set of questions concerning the events, but actually for documenting real-time decisions taken during the event response, such as decisions to mobilize resources including their locations, timestamps and textual chats displaying when actions were carried out by different organizations involved in the emergency response. Further, the possibility for quick information sharing between the organizations involved can facilitate learning on each other's organizational perspectives which can be important aspects when performing the evaluation. It is represented by the *COP data loop.* As the real-time data can be stored, it can serve as a structural support for an organization by making the results accessible as a repository. In this way, it may facilitate reflections on multi-organizational aspects and provide enhanced possibility for the challenging task of lessons to be learned on each other's tasks and perspectives (Greenwood, 1998).

The evaluation stage of the learning loop can support typical AAR questions, where the participants` statements will be strengthened by evidence collected by the tool. The graphical representation of the event progression can help the participants remember the actions by seeing the different developments of decisions from time to time documented by the SQUARE application which will be generated when using the tool in the real-time. This can be a crucial element when converting the lessons-learned from this particular incident and further into general knowledge for the organizations involved (Jacobsson et al., 2011).

In *the implementation stage*, users can make notes to identify the lessons, which also can be documented in the SQUARE application. In this way, SQUARE eventually can potentially support the double-loop learning through the evaluation formalization loop, common operational picture (COP) loop and structural support (represented by three arrows linking from implementation to the planning). In this case, a good learning process would trigger further restructuring of goals and planning for responding to a crisis, which typically is stated in the organizational goals and procedures as a part of *the planning stage*. In this case, the possibility to store the information and generate after action reports into a repository, indirectly support the planning stage, and results in several the lessons-learned that may be used for organizational double-loop learning and improvements. To repeat, our study emphasizes *how* SQUARE supports AAR, LFI, reflections after events, and potentially promotes double-loop learning. We do not claim that double-loop learning is automatically successfully accomplished through the tabletop exercise using SQUARE.

A limitation of the study is that SQUARE was developed under a strict project deadline and accommodated only parts of requirements and needs expressed by the multi-agency stakeholders. In the exercise, the participants mentioned additional wish-lists that were not included in the application to the date of this publication. These "wished features" are the possibilities to choose the timeline model by having short, medium and comprehensive levels of evaluation, to fasten the replayed timeline, and to remove irrelevant evaluation points, especially in the long-term storage in organizations. These new requirements have not yet been implemented in the prototype.

CONCLUSIONS

In this study we have presented the results from a table-top exercise for testing an innovative technology that can facilitate fact-based elements for evaluation and enhanced learning, i.e., the SQUARE tool. The tool and related functionalities were developed iteratively based on a set of requirements that have been extracted from a workshop with emergency management stakeholders in Norway, including regional and national levels. We have asked research questions regarding the experience of emergency management responders when using the SQUARE tool for enhancing learning from incidents and map-based COPs, as well as the enablers or barriers that can facilitate or hinder its adoption and use in EM day-to-day work.

The evaluation session of this table-top exercise shows that the stakeholders agree that the tool is promising in supporting event reconstruction-based evaluation reconstruction of events that enables learning and make it possible for practitioners to evaluate the development of the common operational picture through several features that has been included in SQUARE (e.g., the replay function and self-generated report from an event response). However, to enable organizational adoption of the tool, the participants consider the need of the top management to look and have further experience with it. It is also important that different units in organizations are familiarized with the tool/functions. Thus, to enable further adoption in the EM organizations, organization´s hierarchical process and decisions must be considered. However, we have already seen the impact that this project has made. The application has attracted the Norwegian Authority for Civil Protection on potential use of the tool for emergency management-related education that is managed by the authority. Thus, despite the limitations as to what features can currently be achieved in SQUARE, this interest opens up the possibility to develop the application further, e.g., as regards missing features specified during the tabletop testing.

ACKNOWLEDGMENTS

We thank all participants of the SQUARE tabletop testing who have engaged in the whole sessions and have provided most valuable feedback. The INSITU project has been funded with support from the Norwegian Research Council Grant#295848.

REFERENCES

Agboola, F., McCarthy, T., and Biddinger, P. D. (2013) Impact of Emergency Preparedness Exercise on Performance. *Journal of Public Health Management and Practice, 19*, S77-S83.

Andersson, D., Pilemalm, S., and Hallberg, N. (2008) *Evaluation of Crisis Management Operations Using Reconstruction and Exploration.* Proceedings of the 5th International ISCRAM Conference.

Argote, L., and Miron-Spektor, E. (2011) Organizational Learning: From Experience to Knowledge. *Organization science, 22*(5), 1123-1137. doi:10.1287/orsc.1100.0621

Argyris, C., and Schön, D. (1978) *Reading Organizational Learning: A Theory of Action Approach*: Addison Wesley, MA.

Bharosa, N., Lee, J., and Janssen, M. (2010) Challenges and Obstacles in Sharing and Coordinating Information During Multi-Agency Disaster Response: Propositions from Field Exercises. *Information Systems Frontiers, 12*, 49-65.

Birkland, T. A. (2009) Disasters, Lessons Learned, and Fantasy Documents. *Journal of Contingencies and Crisis management, 17*(3), 146-156.

Boin, A., and 't Hart, P. (2010) Organising for Effective Emergency Management: Lessons from Research 1. *Australian Journal of public administration, 69*(4), 357-371.

Borglund, E. A., and Granholm, M. (2020) *You Talk the Talk–but What Do You Talk About.* Proceedings of the Proceedings of the 17th ISCRAM Conference, Blacksburg, VA, USA.

Boud, D., Keogh, R., and Walker, D. (2013) Promoting Reflection in Learning: A Model. In *Reflection: Turning Experience into Learning* (pp. 18-40): Routledge.

Bunker, D., Levine, L., and Woody, C. (2015) Repertoires of Collaboration for Common Operating Pictures of Disasters and Extreme Events. *Information Systems Frontiers, 17*, 51-65.

Cantero, C., Damsgaard, E., Hamre, M., and Kolstad, T. N. (2021) Politiet Etter Snøkaoset På E18: – Dette Kan Vi Ikke Oppleve På Nytt. . Retrieved from https://www.nrk.no/sorlandet/politiet-etter-snokaoset-pa-e18_-_-dette-kan-vi-ikke-oppleve-pa-nytt-1.15762735

Choularton, R. (2001) Complex Learning: Organizational Learning from Disasters. *Safety Science, 39*(1-2), 61-70.

Danielsson, E., Alvinius, A., and Larsson, G. (2014) From Common Operating Picture to Situational Awareness. *International Journal of Emergency Management, 10*(1), 28-47.

Donahue, A., and Tuohy, R. (2006) Lessons We Don't Learn: A Study of the Lessons of Disasters, Why We Repeat Them, and How We Can Learn Them. *Homeland Security Affairs, 2*(2).

Drupsteen, L., and Guldenmund, F. W. (2014) What Is Learning? A Review of the Safety Literature to Define Learning from Incidents, Accidents and Disasters. *Journal of Contingencies and Crisis management, 22*(2), 81-96.

Giaoutzi, M., and Scholten, H. J. (2017) A Common Operational Picture in Support of Situational Awareness for Efficient Emergency Response Operations. *Journal of Future Internet, 2*(1), 10-35.

Greenwood, J. (1993) Reflective Practice: A Critique of the Work of Argyris and Schön. *Journal of advanced nursing, 18*(8), 1183-1187.

Greenwood, J. (1998) The Role of Reflection in Single and Double Loop Learning. *Journal of advanced nursing, 27*(5), 1048-1053.

Guest, G., MacQueen, K. M., and Namey, E. E. (2011) *Applied Thematic Analysis*: Sage publications.

Jacobsson, A., Ek, Å., and Akselsson, R. (2011) Method for Evaluating Learning from Incidents Using the Idea of "Level of Learning". *Journal of loss prevention in the process industries, 24*(4), 333-343.

Lasky, M. (2010) The Value of Tabletop Exercises and One-Page Planning Documents. *Journal of Business Continuity & Emergency Planning, 4*(2), 132-141.

Munkvold, B. E., Radianti, J., Rød, J. K., Opach, T., Snaprud, M., Pilemalm, S., and Bunker, D. (2019) *Sharing Incident and Threat Information for Common Situational Understanding.* Proceedings of the 16th ISCRAM Conference, València, Spain.

Myers, M. D., and Avison, D. (2002) *Qualitative Research in Information Systems: A Reader*: Sage.

Nonaka, I., and Takeuchi, H. (1995) *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford: Oxford University Press.

Pilemalm, S., Andersson, D., and Mojir, K. Y. (2014) Enabling Organizational Learning from Rescue Operations: The Swedish Rescue Services Incident Reporting System. *International Journal of Emergency Services*.

Pilemalm, S., Radianti, J., Munkvold, B. E., Majchrzak, T. A., and Steen-Tveit, K. (2021) *Turning Common Operational Picture Data into Double-Loop Learning from Crises–Can Vision Meet Reality?* Proceedings of the 18th ISCRAM Conference–Blacksburg, VA, USA May 2021.

Radianti, J., Lazreg, M. B., and Granmo, O.-C. (2015) Fire Simulation-Based Adaptation of Smartrescue App for Serious Game: Design, Setup and User Experience. *Engineering Applications of Artificial Intelligence, 46*, 312-325.

Renå, H., and Christensen, J. (2020) Learning from Crisis: The Role of Enquiry Commissions. *Journal of Contingencies and Crisis management, 28*(1), 41-49.

Savoia, E., Agboola, F., and Biddinger, P. D. (2012) Use of after Action Reports (Aars) to Promote Organizational and Systems Learning in Emergency Preparedness. *International journal of environmental research and public health, 9*(8), 2949-2963.

Savoia, E., Biddinger, P. D., Fox, P., Levin, D. E., Stone, L., and Stoto, M. A. (2009a) Impact of Tabletop Exercises on Participants' Knowledge of and Confidence in Legal Authorities for Infectious Disease Emergencies. *Disaster Medicine and Public Health Preparedness, 3*(2), 104-110.

Savoia, E., Testa, M. A., Biddinger, P. D., Cadigan, R. O., Koh, H., Campbell, P., and Stoto, M. A. (2009b) Assessing Public Health Capabilities During Emergency Preparedness Tabletop Exercises: Reliability and Validity of a Measurement Tool. *Public Health Reports, 124*(1), 138-148.

Schön, D. A. (1987) *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*: Jossey-Bass.

Smith, S. P., Blackmore, K., and Nesbitt, K. (2015) A Meta-Analysis of Data Collection in Serious Games Research. *Serious games analytics: Methodologies for performance measurement, assessment, and improvement*, 31-55.

Sorensen, J., Magnussen, L., Torgersen, G., Christiansen, A., and Carlström, E. (2018) Perceived Usefulness of Maritime Cross-Border Collaboration Exercises. *Arts Social Sciences Journal, 9*(361), 2.

Steen-Tveit, K., and Munkvold, B. E. (2021) From Common Operational Picture to Common Situational Understanding: An Analysis Based on Practitioner Perspectives. *Safety Science, 142*, 105381.

Steen-Tveit, K., and Radianti, J. (2019) *Analysis of Common Operational Picture and Situational Awareness During Multiple Emergency Response Scenarios.* Proceedings of the Proceedings of the 16th ISCRAM Conference, València, Spain.

Steen-Tveit, K., Radianti, J., and Munkvold, B. E. (2020) *Using Audio-Logs for Analyzing the Development of a Common Operational Picture in Multi-Agency Emergency Response.* Proceedings of the Proceedings of the 53rd Hawaii International Conference on System Sciences.

Tobergte, P., Knispel, A., Landsberg, L., and Mudimu, O. A. (2022) *Evaluation of Tabletop Exercises in Emergency Response Research and Application in the Research Project Sortie.* Proceedings of the Proceedings of the 19th ISCRAM Conference, Tarbes, France.

Weick, K. E., Sutcliffe, K. M., and Obstfeld, D. (2005) Organizing and the Process of Sensemaking. *Organization*

science, 16(4), 409-421.
Wolbers, J., and Boersma, K. (2013) The Common Operational Picture as Collective Sensemaking. *Journal of Contingencies and Crisis management, 21*(4), 186-199.
Yanow, D. (2016) Seeing Organizational Learning: A "Cultural" View. In *Knowing in Organizations: A Practice-Based Approach* (pp. 32-52): Routledge.