Inter-Organisational Learning – A Review of Knowledge Sharing in Post-Exercise Reports

Johan Nordström
Swedish Defence Research Agency
Department of C4ISR
johan.nordstrom@foi.se

Björn J E Johansson
Swedish Defence Research Agency
Department of C4ISR
bjorn.j.e.johansson@foi.se

ABSTRACT

Inter-organisational learning from exercises and response operations is crucial for improving overall response capacity for coping with cross-domain crisis events. In order to compile and disseminate experiences and acquired knowledge in the form of lessons learned, post-exercise reports are written. This paper presents a review of 17 Swedish post-exercise reports. The review was conducted with the aim to investigate whether such reports contained enough information to support inter-organisational learning, i.e. if learning goals were stated, how the event was described, method for evaluation, conclusions, and whether recommendations for change were sufficient for supporting inter-organisational learning. It was found that most reports did not support organisational development and lacked recommendations that were useful outside the own organisation or the context of the specific exercise scrutinised. The results indicate the need for an exercise evaluation framework for inter-organisational exercises.

Keywords

Crisis management, post-exercise reports, inter-organisational learning, double loop learning.

INTRODUCTION

Learning is crucial for any kind of organisation and is in some respects particularly challenging for actors in the domain of crisis management (Baroutsi, 2018; Wang, 2015). Exercise evaluation has been described as crucial for learning, as it in many cases is the only feedback given to exercise participants (Baroutsi, 2018). Inter-organisational learning from actual crisis events or inter-organisational exercises is perhaps the most challenging type of learning as this demands the identification of lessons learned that are applicable not only to specific organisations or actors, but to a collective of such organisations or actors. Exercise reports, particularly those based on inter-organisational exercise events, are consequently very important sources of knowledge that need to be disseminated within the crisis response community in order to facilitate learning about inter-organisational crisis management. In 2014, the Swedish Civil Contingencies Agency (Myndigheten för Samhällsskydd och Beredskap, MSB) published a guideline introducing the fundamentals of exercise planning (MSB, 2017a), and in 2016 the agency published a guideline on exercise evaluation (MSB, 2017b). According to MSB (2017b), the purpose of exercises is to generate knowledge of strengths and deficiencies to systematically develop or maintain the ability to handle emergencies, i.e. organisational learning. Another purpose is to test plans, actions and organisations in a simulated environment, where failures will not lead to injuries or serious consequences in the society. Post-exercise reviews should answer the questions 'how did it go?' and 'why did it happen?'. In order to do so, the exercise evaluator needs to consider the conditions for the exercise, the participants' actions and the outcome of the participants' actions (ibid.). Furthermore, MSB (2017b) recommends the use of

indicators for exercise evaluations. An indicator indicates whether a certain ability has been demonstrated during the exercise. Indicators are often defined in terms of observable data, such as questions that can be answered by true/false, or measurable data such as time, numbers or percentages.

Granåsen et al. (2018) have pointed out that a systematic approach to identifying lessons learned for inter-organisational crisis management, in other words, the crisis management *system*, is needed. In a literature review by Granåsen et al., it was found that different aspects of crisis management capability can be organised into nine thematic clusters: interaction, relationships, coordination/C2, system performance, preparedness, situation awareness, resilience, decision-making and information infrastructure. The review surveyed 150 papers that had been identified as relevant out of a total of 1197 publications found in a database search. The nine clusters identified highlight the ambiguity of what capabilities are important in inter-organisational work as well as the potentially confusing usage of terminology within the field. Nevertheless, the authors suggested that these clusters can be used to support analyses of exercise reports in the sense that lessons learned that are applicable to one or more of the identified clusters can most likely be of importance to actors.

According to Argyris (1977, p. 116), organisational learning is 'a process of detecting and correcting errors'. As emergencies are quite rare, exercises are important opportunities for detecting and correcting problems within the crisis management system. In addition, they are vital for organisational learning within the crisis management domain. Argyris (1977) denotes the cases where it is possible to detect and correct errors without changing the organisation's underlying objectives or policies as single loop learning. Cases that require a change of the organisation's objectives or policies are referred to as double loop learning. Further, Argyris (1991) argues that highly skilled professionals are very good at single loop learning but are often relatively bad at double loop learning. According to this line of reasoning, organisational change cannot be expected to take place as a consequence of everyday operations, as such operations rarely challenge established work practices. Rather, it is deviating, unusual events that challenge organisational practice and demand inter-organisational cooperation (Quarantelli, 1998; Bigley & Roberts, 2001; Trnka & Johansson, 2009). Such crisis events are fortunately rare and some organisations will only ever encounter them in exercises or simulations. This fact stresses the importance of capturing knowledge and experiences generated during such exercises in order to support both organisational and inter-organisational learning, particularly when knowledge that could support double loop learning is concerned. How is then such knowledge captured, stored and disseminated? In most cases, the main output after an exercise, be it a single-organisation exercise or an inter-organisational exercise, is a report written by an evaluator who is either internal or external to the participating organisation(s).

The objective of this study is to explore if post-exercise reports identify and describe conditions and experiences that support inter-organisational learning, i.e. is it possible to identify lessons learned, suggestions for change and action from these kinds of reports and apply it to other organisations? The study was carried out as a review of exercise reports within a Swedish context. The review investigates how regional inter-organisational exercises in Sweden have been evaluated in terms of:

- Whether the exercise reports actually describe inter-organisational aspects, i.e. do they reflect the nine thematic clusters of inter-organisational learning identified by Granåsen et al. (2018)?
- How the exercise goals and evaluation approach were described in the reviewed reports – are they suitable for inter-organisational learning or even organisational learning?
- Whether the reviewed reports support single or double loop learning as described by Argyris (1977)?
- Whether the results are applicable outside the realm of the specific exercise or organisation?
- Whether there are recommendations in the report, and to what extent these are applicable to other organisations?

METHOD

The present study is based on reviews of post-exercise reports from 17 regional cooperation exercises in Sweden during 2008-2017. The search for post-exercise reports was performed partly in a document database in MSB's integrated decision support system (RIB), and partly on the internet. The inclusion criteria were that the report must concern an exercise describing an event that in a real situation would affect several organisations, even if only one organisation participated in the reported exercise. There should thus be a potential for reporting on issues of interest in an inter-organisational context, i.e. show potential for learning about such events. Furthermore, a time limit of ten years was applied (2008), and reports were selected from what was deemed to be a representative set of counties. Due to demographical aspects, such as funding for exercises and reporting tradition/culture, some counties were found to have carried out more exercises than others. Consequently, more reports were available in those counties.

The search for post-exercise reports on the internet was carried out in two steps. In the first step, a search for 'samverkansövning' (i.e., cooperation exercise, normally indicating an inter-organisational exercise) was carried out on the web pages of all county administrative boards in Sweden. In the second step, three different searches were made using Google and the search terms 'samverkansövning' (i.e., cooperation exercise[1]), 'regional övning' (i.e., regional exercise), and 'utvärderingsrapport & övning' (i.e., evaluation report & exercise). The search in RIB generated 35 results, of which nine contained post-exercise reports that met the inclusion criteria. The internet searches yielded four additional post-exercise reports that met the inclusion criteria. As only one post-exercise report from a county with a nuclear power plant was found, one county administrative board (county 4 in Table 1) was contacted directly. Four post-exercise reports were obtained in this way. In total, exercise reports from seven of 21 counties in Sweden were included in the analysis.

All reports were written in Swedish. The exercises covered emergencies such as the release of radioactive material from nuclear power plants, heavy rainfalls, landslides, shortage of fuel, interference with payment/financial systems, pandemic flu, epizootics, large oil spills at sea, air accidents and threats of violence against schools. The reviewed post-exercise reports were written by eleven authors representing six organisations, i.e. one municipal rescue service, three county administrative boards, one university and MSB. One author, employed at MSB, wrote seven reports, one author employed at a county administrative board was a co-writer of three reports, and another author employed at a county administrative board was a co-writer of two reports (Table 1). All other authors wrote one post-exercise report each.

Table 1. Reviewed post-exercise reports

Exercise	Year	County	Author(s), employed at
1	2016	1	A, municipal rescue service
2	2017	2	B, county administrative board
3	2013	2	C, MSB
4	2010	2	D, county administrative board
5	2008	2	E, county administrative board
6	2012	3*	C, MSB
7	2017	4*	F & G, county administrative board
8	2016	4*	H & G, county administrative board
9	2016	4*	F & G, county administrative board
10	2015	4*	I, county administrative board
11	2011	5	C, MSB
12	2016	6	J, county administrative board
13	2011	7	C, MSB
14	2014	8	K, university
15	2012	8	C, MSB
16	2010	9	C, MSB
17	2008	9	C, MSB

* County with a nuclear power plant.

RESULTS

The reviewed post-exercise reports generally had the same structure: (1) background to the exercise, (2) description of the exercise scenario, (3) description of the purpose and goals for the exercise, (4) description of the evaluation process, (5) evaluation, and (6) areas for improvement. In ten of the post-exercise reports there was also a description of the exercise planning process and planning organisation (included between (1) and (2)). Two post-exercise reports lacked a description of areas for improvement.

The capability clusters identified by Granåsen et al. (2018) were applied in an effort to see what type of inter-organisational aspects the reports addressed. It was found that most reports (more than ten) addressed the clusters interaction, relationships, situation awareness, coordination/C2, preparedness, and information

[1] All translations from Swedish have been made by the authors.

infrastructure, while largely ignoring (less than ten) system performance, resilience, and decision-making. Table 2 describes the distribution of reports for each cluster. *Situation awareness* positions itself on the borderline as it occurs in nine reports, which is just about half of the reports included in this review.

Table 2. Number of reports in each cluster

Cluster	Capabilities found in post-exercise report goals or areas of improvement that are also mentioned by Granåsen et al. (2018)	Number of post-exercise reports
Interaction	Communication, collaboration, information sharing, collaboration practices, inter-organisational crisis management, participant interaction, managing collaborative actions, information flow.	16
Relationships	Formal and trust based relationships, resource allocation (shared), formal and informal networks, predictability, awareness of tasks, roles, material and personal resources, creating a collaborative culture, building a collaborative institution, organisational structure and practices.	15
Coordination/C2	Hierarchies between actors, C2 coordination, coordination centres, hierarchical multi-agency organisational structure, inter-organisational coordination, strategy.	13
System performance	Team performance, system/organisational capacity, capability and function, general organisational performance, operational efficiency.	5
Preparedness	Training and evaluation of disaster preparedness plans, contingency planning, preparedness (institutional, individual, collective, local emergency), emergency plan, learning related work activity.	15
Situation awareness	Situation assessment, shared situation awareness, situational picture.	9
Resilience		0
Decision-making	Analysis of crisis situation.	1
Information infrastructure	Information and communication technology (ICT), use of coordination technology equipment, inter-organisational IT use.	14

The main focus of the reviewed post-exercise reports was on interaction, relationships, coordination, preparedness and infrastructure. This was expected as the inclusion criteria directed the search towards inter-organisational exercises. Somewhat more surprisingly, *system performance* was noticeably absent in the goal descriptions (stated as goals in five reports), even though it was mentioned as a purpose in eleven of the exercises,. This is surprising as the overall performance of the organisation probably could be turned into an assessable indicator. *Decision-making* only occurred in one report, which may be less surprising as the focus of the identified exercises was inter-organisational cooperation, indicating that decision-making aspects probably were not the primary focus. If they were, they were probably approached from the angle of team decision-making or cooperation. *Resilience* was not found at all. However, this term is relatively new in a Swedish context, and resilience-related aspects may very well have been exercised although this terminology was not used. Information systems for supporting information sharing between organisations were rarely mentioned. The only system that was consistently referred to was WIS, an IT-system designed to support diary-like status

reports posted into a shared database. However, the post-exercise reports usually described technical difficulties with WIS rather than evaluating the potential benefits of using it. No reports provided suggestions for improving WIS. In general, the reviewed reports seem to reflect capabilities that are crucial for inter-organisational work in crisis response.

The stated goals for each reported exercise differed, although all but three stated goals that were classified as 'qualitative' in the analysis. The remaining three exercises did not report any stated goals whatsoever. No exercise was given goals that could be classified as quantitative, such as, for example, 'all fire brigades must reach the incident area within X minutes'. According to the post-exercise reports, indicators were used to evaluate the goal achievement in seven of the exercises. For each of the exercise goals, between one and 14 indicators were used for evaluation, with an average of 5.3 indicators per goal. Table 3 presents four examples of goals from two different exercises (exercise 10 and 14) that demonstrate a great difference in the number of indicators.

Table 3. Examples of goals and the number of related indicators found in the exercise reports

Exercise	Goal	Number of indicators
10	*The actors have a good ability to understand their own role in a nuclear accident*[2].	2
	The actors have a good ability to communicate internally and externally, as well as explain the decisions taken and their related consequences.	13
14	*The actors have a basic ability to work in accordance with the county's guidelines for coordination in a crisis situation.*	14
	The actors have a good C2 ability and a good ability to cooperate with other actors in: (a) search and rescue operations at sea, (b) oil spill clean-up operations at sea, (c) oil spill clean-up operations in coastal environments.	3

Table 3 also presents examples of how goals were formulated in the exercise reports. As can be seen, goals were formulated in terms of *abilities*, often related to several different situations and skills ('good ability to communicate internally and externally, as well as explain the decisions taken and their related consequences'). However, there does not seem to be a direct relationship between the complexity of the goal statement and the number of indicators used to assess whether the goals were reached. For example, the second goal in Table 3 relates to 13 indicators, while the fourth goal, the description of which is by no means less complex than the second goal, only relates to three indicators. In turn, indicators manifest themselves in different ways in different reports. Table 4 shows four randomly selected goals from the various reports and how the indicators were described. As can be seen, most indicators were defined as observable events that can be answered by *yes* or *no*.

[2] All citations and examples from exercise reports have been translated from Swedish by the authors.

Table 4. Examples of indicators for different exercise goals

Exercise	Goal	Indicators
2	*Achieve reporting of operational picture and a shared situational awareness*	• Requested information is delivered to the county administrative board within the stipulated time. • Requested documentation is delivered to the county administrative board with the requested information. • WIS (an MSB system used for information sharing on the internet) is used among actors as a tool for reporting operational picture. • A shared situational awareness is created between the involved actors during orientation and coordination meetings, using the regional methodology for achieving a shared situational awareness. • The shared situational awareness provides sufficient support for orientation and coordination on a regional level.
3	*Basic ability to cooperate according to the regional cooperation agreement.*	• If necessary, the regional TIB group ('officials on standby group') is contacted by the alarm centre within 30 minutes. • If necessary, a first orientation and coordination meeting is held within one hour after the first alarm. • Within two hours, the TIB group has made a decision on the continued orientation and coordination of the actions taken due to the event.
7	*The county administrative board has a basic ability to provide relevant means of communication at the forward command post*	• There is access to Rakel (a TETRA-based Swedish communication system) and it can be used. • There is access to Skype and it can be used. • It is possible to connect to the county administrative boards' intranet.
14	*Basic ability to cooperate based on local and regional oil spill response plans in the event of a major oil spill at sea.*	• The actors assess the need for support from other organisations. • The actors formulate the need for support from other actors. • The actors request support from other actors. • The actors receive a response to their request for support from other actors.

As can be seen in Table 4, some indicators, for example 'There is access to Rakel (a TETRA-based Swedish communication system) and it can be used' is relatively easy to assess. Others, such as 'A shared situational awareness is created during orientation and coordination meetings between the involved actors, using the regional methodology for achieving a shared situational awareness', can be very challenging to assess, if at all possible. In terms of evaluation methods, various qualitative methods, such as interviews, observations, evaluator judgement, surveys, protocols, and after-action reviews were applied. Interestingly, all but one report stated that methods were combined in different ways to evaluate the concerned exercise. Applied in twelve of

the evaluations, surveys were by far the most common method used. . However, surveys were in all cases combined with other methods, most commonly observations and evaluator assessments. In three cases, interviews were utilised. 'Evaluation seminars', probably resembling after-action reviews, were relatively common, with five occurrences in our data set. The least common method was 'protocols', which evidently was a form of pre-arranged checklists used to assess whether or not the participants in the exercise reached the specified goals.

Table 5. Goal type, evaluation methods, and use of indicators identified in the reviewed reports.

Exercise	Type of Goals	Methods applied	Use of Indicators
1	Qualitative	Evaluator's rating, Documents and logs from the exercise, Observations, Evaluation seminar.	No
2	Qualitative	Evaluator's rating, Observations, Documents and logs from the exercise, Evaluation seminar.	Yes
3	Qualitative	Survey, Evaluator's rating, Review of documentation and logs from the exercise, Evaluation seminar, Observations.	Yes
4	Qualitative	Evaluator's rating, Survey	No
5	Qualitative	Observations based on Observation protocols, Review of documentation and logs from the exercise, Evaluation seminar, Survey, Interviews.	No
6	Qualitative	Reports from evaluators, Evaluation seminar	Yes
7	Qualitative	Evaluator's rating, Survey, Evaluation seminar.	Yes
8	No goals	Survey, Observations.	No
9	No goals	Survey, Documentation from exercise	No
10	Qualitative	Survey, Evaluator rating based on Observation Protocol, Evaluator's rating.	Yes
11	Qualitative	Survey, Evaluator's rating, Observations, Review of documentation and logs from the exercise, Evaluation seminar	No
12	Qualitative	Survey, Review of documentation and logs from the exercise, Evaluator's rating, Observations	Yes
13	Qualitative	Survey, Evaluator's rating, Review of documentation and logs from the exercise, Evaluation seminar.	No
14	Qualitative	Surveys, Evaluator's rating, Interviews	Yes
15	No goals	Survey, Observations	No
16	Qualitative	Survey, Review of documentation and logs from the exercise, Observations, Interviews.	No
17	Qualitative	Documentation and logs from the exercise, Survey	No

The outcome of these evaluations were in all cases descriptive, independent of whether goals were articulated and/or indicators were used. No statistical or other form of numerical analysis was conducted in order to

compare the outcome of the exercise with set goals, earlier exercises, or other exercises.

However, the results were discussed in a variety of ways. Six evaluations discussed whether or not the exercise goals were fulfilled. Three discussed the strengths and weaknesses of the exercise. One presented lessons learned. The other evaluations discussed the participants' experiences, whether or not the design of the exercise was appropriate, the number of participants in the surveys conducted, etc.

Based on the outcome of our analysis, an effort was made to determine whether or not the exercise reports could be considered to support double loop learning, as described by Argyris (1977). It was found that only five reports seemed to achieve this goal (Table 6). This analysis was based on the following criteria: Reports were deemed to support single loop learning if the recommendations focused on improving existing procedures, structures or functions, such as exercising more to perfect a procedure, improve individual skill, or improve an existing technical system. Reports were deemed to support double loop learning if the report questions revolved around the way work was conducted, and if there were recommendations going beyond conclusions in the style of 'more exercises are needed', or other suggestions that can be labelled as single loop learning. The type(s) of goal(s) formulated for each exercise was also taken into account. Table 6 illustrates the number of reviewed reports that support single loop learning vs. double loop learning.

	Type of learning	
	Single loop	Double loop
Indicators used for evaluation	6	1
Indicators not used for evaluation	6	4

Lastly, we performed an analysis of the recommendations provided in the reviewed reports in an effort to answer the question whether the reports support learning outside the specific exercise context. Eight of 17 reports did to some extent support learning outside the own organisation or specific exercise context. In some cases, the recommendations require a degree of interpretation on behalf of the reader to achieve this. In general, recommendations may be hard to implement even within the realms of the own organisation or exercise context. Table 7 presents a set of randomly selected suggestions for improvements from different reports. As shown, some of the recommendations were written in such a way that they include a motivation as well as suggestions on how to implement the recommendation. In other cases, the recommendation consists of a few words, which implies that readers need to provide a context as well as a way to implement the suggested recommendation themselves. It can also be concluded that there seems to be no agreed format for providing recommendations in post-exercise reports of the kind reviewed in this paper.

Table 7. Random examples of conclusions/future improvements in the post-exercise reports

Exercise	**Conclusion/suggested future improvements**
2	*One of the goals with the exercise was to achieve a shared situational awareness. Today, there is a defined method for achieving a shared situational awareness. Simplified, the method is based on all actors providing information, which is then compiled into a regional shared operational picture. During the exercise the information was distributed by the WIS system, except for restricted information. As there was a new version of WIS, which was not fully functional during the whole exercise, problems occurred with both the submission of information and the compilation of a common operational picture* • *Further development work should be carried out based on the established method. This work should highlight information gathering, compilation, updates and dissemination of the common operational picture. Knowledge about and the ability to use WIS is important. The work should include routines for both the regular approach and for a backup system in case the regular system fails. It is also important to highlight how restricted information is to be handled and distributed.*
3	*Develop the work on shared situational awareness by:* • *Determining what information shared situational awareness*

	should include. • *Providing procedures for reporting.*
6	*The following points need to be addressed to increase the county's ability to handle a nuclear accident:* • *Confusion about concepts* • *Clarification of roles, responsibilities and mandate* • *Communication with units at sea* • *Alarm numbers and messages* • *Receiving stations* • *Synchronisation of plans* • *Transport issues and traffic diversion* • *Endurance* • *Development of exchanges of liaison officers* • *Increase of the ability to command county administrative board staff.*
17	*The three closing questions in the questionnaire discussed the future. Some answers that stood out are reported below.* ***Have you discovered a need for constitutional changes to better manage a pandemic? If so, please specify.*** • *The Working Hours Act may constitute an obstacle. A review of possible exceptions from the act is needed.* • *A legal prerequisite for distributing fuel or taking action with mandates similar to those of an incident commander, according to the Civil Protection Act (SFS 2003:778,) is needed* ***What support does your organisation expect from the county administrative board for the continued planning for pandemics?*** • *A handbook on the legal aspects of crisis management.* • *Exchange of experiences from the planning process within the county.* • *Epidemic intelligence and exercises.* ***What support does your organisation expect from other actors than the county administrative board for the continued planning for pandemics?*** • *Support from the Swedish Institute for Infectious Disease Control in the production of instructions and advice for contingency planning.* • *Knowledge from the county council's infection protection unit.*

In general, the exercise reports focused on describing what happened and what actions were carried out during the exercise. Only a small number of the reports were devoted to analysing the problems that arose during the exercise, i.e. finding the underlying reasons for the problems that occurred, how the problems were solved during the exercise, and what measures that need to be taken to prevent the problems from occurring again. It is also difficult to follow the link between the problems that arose during the exercise and the suggested improvements.

Naturally, the evaluation of the exercise took its starting point in the goals of the exercise. However, as most of the exercise goals were formulated as abilities, there is not much in the way of analysis, and there are only a few suggested improvements in the cases where the exercise goals were met, even if the evaluation showed that

there was room for improvement.

DISCUSSION

When analysing the 17 reports in the light of the nine thematic capability clusters identified by Granåsen et al. (2018), it was shown that the reports largely reflect capabilities important to inter-organisational crisis response work. This is promising as the thematic clusters reflect a large set of academic writings on inter-organisational work, and thus indicate that the core of the exercises were conducted in such a way that there was potential for learning and improving overarching crisis response system capabilities. However, some challenges and room for improvement remain, as suggested by the results presented above.

The 'indicator approach' suggested in the MSB (2017b) guidelines is positive in the sense that indicators, at least in some respects, can be assessed objectively and with well-defined measures, such as time, distance, frequency, etc. However, the connection between indicators and theory in terms of learning, particularly organisational and inter-organisational learning, is weak, even in a best-case scenario. Hollnagel (1997) discussed that researchers and evaluators often grasp for *what is measurable* rather than *what should be measured*. This implies that many of the measures performed using the MSB guidelines may capture experiences that are less useful for supporting aspects of double loop learning, or even single loop learning. In several cases, indicators of the type 'was IT-system X used during the exercise? ' were found. Even if it is possible to count the number of times IT-system X was used or not, such an indicator, nor reporting the outcome of it, does not explain *why* it was/wasn't used. Furthermore, the present review shows that the type of indicators used vary greatly between similar goal formulations, indicating that the theoretical or methodological support for assigning indicators is poor. A plausible hypothesis is that the number of indicators for a given goal reflects the possibility to identify observable events rather than a reflection of the complexity of the goal. Some goals are related to only two indicators, whereas others are related to as many as 15. In addition, the descriptions of the various indicators suggest that there is a great variation in the types of method needed to assess whether they have been fulfilled or not, ranging from system log files to subjective self-estimates. This is not a problem in itself, but reflects the ambiguity in the way indicators are chosen and applied in exercises.

The present review showed that when indicators were used to evaluate an exercise, the focus for the evaluation tended to be on whether the goals of the exercise had been met or not. However, as most exercise goals focused on ability and not learning, the result of the evaluation tended to present a 'test result', rather than a starting point for future improvements of the crisis management system. This view indicates that the exercise was viewed as a confirmation of the quality of an existing ability, rather than a learning event. The level of detail in indicators and goals also varied greatly, ranging from the ability to handle a specific IT-system to more generic abilities to establish inter-organisational cooperation. This variety of indicators and goals suggests that challenges remain when it comes to creating guidelines and training for exercise planners and evaluators, if useful exercise reports in an inter-organisational context are to be assured.

Another aspect of the evaluation in the reviewed reports was the descriptive, qualitative, approach utilised. While a qualitative approach is useful for describing what took place during the exercise and why, it has limited value in terms of comparability between different exercises. For example, if no quantitative performance assessment is carried out it becomes impossible to analyse the development of performance over time. The same goes for comparisons between exercises or organisations. In sum, all kinds of statistical analysis are effectively rendered impossible.

Moreover, support of double loop learning was found in fewer post-exercise reports using indicators than in post-exercise reports where indicators had not been used. In the single case where a post-exercise report used indicators and the report resulted in double-loop learning, the double loop learning was supported by reports by observers rather than evaluation reports based on indicators from evaluators. Clear ways for incorporating double loop learning aspects in exercises, as well as methodological support for evaluating and reporting this, seems to be lacking or be outside the scope of most regional exercise evaluators' training.

The challenges posed by crisis events today also point to a more overarching question: given that most exercises aim to improve existing capacities or assure that abilities are kept on an acceptable or good level, will crisis management systems be able to handle increasingly complex and challenging events? The literature shows that crisis response organisations always need to prepare for extraordinary events, suggesting that such organisations need to be able to perform not only at an acceptable level but at an *exceptional level*. Consequently, this needs to be reflected in exercise goals, at least in some cases. This implies that failure during exercises must be a possibility, something that was noticeably absent in the reviewed reports.

Only about half of the reviewed reports contained recommendations that were found to be useful outside the context and the organisations taking part in each respective exercise. However, it should be noted that the

recommendations identified frequently require a fair degree of interpretation on behalf of the reader as they do not always motivate why the recommendation is given or how the recommendation should be implemented. It is naturally challenging to create such recommendations, particularly in an inter-organisational context, but our review shows that there is a need to investigate if it is possible to create a method or guideline concerning the format of such recommendations as they are crucial for knowledge sharing in an inter-organisational crisis response context. One suggestion could be to organise the lessons learned from an exercise according to the nine clusters suggested by Granåsen et al. (2018) and provide recommendations for each cluster in a structured way.

How should exercise reports then be improved to better support inter-organisational learning? The answer is multi-faceted and cannot be answered based on this investigation alone. However, some suggestions can be made based on the outcome of this study. Firstly, there is a lack of common terminology that could be supported by creating an inter-organisational ontology that describes phenomena that are important for inter-organisational exercises. Such an ontology could be based on the clusters provided in the paper by Granåsen et al. (2018). Secondly, the way data is collected during exercises must be aligned in such a way that it is possible to compare different exercise events with each other. Exercise evaluators should have similar training and be supported by a data collection framework, including observation protocols and other data collection tools. Similar indicators should be used across different exercises. All exercises cannot be treated in the exact same way when it comes to evaluation, but a small set of approaches to evaluation should be sufficient to cover most situations. Thirdly, exercise reports should have a similar structure and disposition, making it easy for readers to find the sections they are most interested in. Conclusions should preferably be arranged in the same way. Recommendations that apply to both immediate needs (single loop learning) and more fundamental challenges (double loop learning) should be given.

One way of supporting the effort to improve data collection during exercises, as well as the task of writing useful post-exercise reports, could be to introduce an IT support system that structures the exercise evaluators' work . Such an IT-system should preferably be possible to carry around, like a tablet, be connected to a database for storing and sharing exercise data between participating organisations in the inter-organisational context, and have the ability to record sounds and videos. The physical device needs to be rugged and have enough battery capacity to endure the realities of field observations. Essentially, the exercise evaluation support system should be a digital notepad, using semi-structured observation protocols and having the ability to document different situations that the exercise evaluator finds important. The database capturing the collected data should provide output that is structured according to the logic of the inter-organisational exercise reports suggested in the previous paragraph. This does not imply that analysis should be automated, only that a structured approach to data collection is supported and that a shared format for collecting the same data is assured.

CONCLUSIONS

This review has identified a number of challenges for knowledge sharing from exercises in an inter-organisational context. In the analysis, it was found that:

- Exercise goals are in many cases formulated as 'abilities', indicating that exercise goals focus on assuring that the involved organisations have the ability to perform certain tasks, rather than relating to learning or development.
- The type and number of indicators used to assess the exercise goals vary greatly between exercises. In many cases, the indicators concern specificand easily assessable issues, such as 'was system X used?', without a clear description explaining why this is important or how it should be interpreted in an inter-organisational context. Theoretical support for the selection and usage of indicators was largely absent in the reviewed reports. In addition, there were no discussions about whether the indicator(s) used were sufficient to determine if a specific goal had been met.
- No structured data collection of quantitative measures was applied in the reviewed exercise evaluations. Neither were statistical analyses performed on the collected data. This limits the possibility to assess development over time, as well as the possibility to compare exercise outcomes or organisational performance between exercises.
- Five of 17 reports could be viewed as supporting double loop learning, i.e. supporting recommendations that could lead to change on a systems level, such as process, organisation, training or method. However, these recommendations did in most cases lack sufficient motivation and description of how such a change should take place.
- About 50% of the reports were deemed to have the potential to be useful in other contexts than that of the specific exercise or involved organisation(s). However, as in the case with double loop learning, the

recommendations and lessons learned generally did not provide suggestions on how to implement them in other contexts than the reported exercise.

- With the exception of one nuclear energy emergency exercise where an 'internet diary' was used to share information between the actors, WIS was the only IT system for information sharing mentioned in the post-exercise reports. Regarding the use of WIS, conclusions in the post-exercise reports were limited to the need for education and training of individual users, and the need for routines concerning how the system should be used for sharing an operational picture. None of the examined post-exercise report suggested technical improvements to WIS, despite the fact that several reports mentioned that there are deficiencies in the system and that users experience difficulties when using the system.

An investigation is needed in order to formulate a framework that supports planners and evaluators of inter-organisational exercises in conducting exercises and evaluations that are useful, not only for the exercised participants/organisations, but also for other actors. The authors of the present study therefore suggest an outline for such a framework based on the clusters identified by Granåsen et al. (2018). This should include training provided to all inter-organisational exercise evaluators, and be supported by a rugged, field-worthy IT system, such as a tablet, that aligns data collection by providing observation protocols and the ability to collect sound and video recordings. Inter-organisational crisis response capabilities are crucial, but training and exercising such capabilities is costly and resource-demanding. It is therefore highly important to ensure that experiences and lessons learned from such exercise events can be utilised to their full extent. In most cases, the primary product capturing lessons learned from an exercise is the post-exercise report, which thus must be written in such a way that it is relevant beyond the scope of the specific exercise or organisations involved, and include recommendations that are useful for a wide range of actors in the crisis response community.

ACKNOWLEDGEMENTS

This paper presents a study carried out within the research project KOMET, sponsored by the Swedish Civil Contingencies Agency.

REFERENCES

Argyris, C. (1977) Double loop learning in organizations, *Harvard business review*, 55, 5, 115-125.

Argyris, C. (1991) Teaching Smart People How to Learn, *Harvard business review*, 69, 99-109.

Baroutsi, N. (2018) A Practitioners Guide for C2 Evaluations: Quantitative Measurements of Performance and Effectiveness. In *Proceedings of the 15th ISCRAM Conference*, Rochester, NY.

Bigley, G.A., & Roberts, K.H. (2001) The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal,* 44, 6, 1281-1299.

Granåsen, M., Olsén, M. & Oskarsson, P.-A. (2018) Assessing Inter-organizational Crisis Management Capability - Initial Results of a Systematic Literature Review, *Proceedings of the 15th ISCRAM Conference*, Rochester, NY.

Hollnagel, E. (1997) Measurements and models, models and measurements: You can't have one without the other. *Collaborative crew performance in complex operational systems,* 14, 2.

MSB. (2017a) *Exercise guidance: basic manual - an Introduction to the fundamentals of exercise planning*, Swedish Civil Contingencies Agency, Stockholm.

MSB. (2017b) *Exercise guidance: method booklet - exercise evaluation*, Swedish Civil Contingencies Agency, Stockholm.

Quarantelli, E.L. (1988) Disaster Crisis Management: A summary of research findings. *Journal of Management Studies*. 25, 4, 373-385.

Trnka, J., & Johansson, B. (2009) Collaborative Command and Control Practice: Adaptation, Self-Regulation and Supporting Behavior. *Int. J. Information Systems for Crisis Response and Management*, 1,2, 47-67.

Wang, W. T. (2015) Evaluating organisational performance during crises: A multi-dimensional framework, *Total Quality Management*, 23, 6, 673–688.