An AI-Supported Method for Scalable Micro-Exercises for Crisis Training

Martina Granholm
Mid Sweden University
martina.granholm@miun.se

Erik Borglund
Mid Sweden University
erik.borglund@miun.se

ABSTRACT

This article presents an ongoing study on AI-supported micro-exercises for crisis management training. The research is conducted across three municipalities and incorporates insights from educational contexts. It involves parallel tabletop exercises simulating Emergency Operations Centres (EOC). Audio recordings from these exercises were transcribed using OpenAI Whisper and analyzed with AI tools such as ChatGPT. The AI-generated transcripts facilitated comparisons of the groups' working methods and enabled reflective organizational learning.

Preliminary findings indicate the potential of AI to offer a scalable, data-driven exercise format that supports documentation, analysis, and after-action review. Our study advocates the advantages of flexible and repeatable micro-exercises and underscores the importance of ethical considerations when employing AI in crisis training.

Keywords

Emergency operation centre, micro-exercises, AI-support, method.

INTRODUCTION

Societal crisis preparedness places high demands on cooperation, analytical ability, and adaptation to rapidly changing events (Boin, 2009; Comfort, 2007). Crisis management typically involves multiple actors who must develop a shared understanding of evolving situations and make decisions under conditions of uncertainty and time pressure. This is often achieved in temporal organizations referred to as situation rooms, control rooms or Emergency Operation Centres (EOC). In this paper, we use the concept EOC (FEMA, 2025; Quarantelli, 1978), which is widely discussed in crisis management literature. EOCs can be understood as both a physical location and an organizational structure that facilitates communication, information sharing, and strategic planning. To function as such during a crisis, training and exercises are necessary. Crisis exercises play an important role in supporting preparedness and capability development, and are widely used by organizations to train staff, test procedures, and strengthen coordination (Eriksson & Hallberg, 2022; Perry, 2004). Ansell et al. (2010) emphasize the necessity of flexible training approaches to enable adaptation and improvisation during a crisis. A higher level of training correlates with a higher proportion of crisis response activation (Fast et al., 2016).

However, many traditional exercise formats—particularly larger or more comprehensive exercises—are resource-intensive and require substantial planning, personnel, and logistical support. There are also studies questioning the usefulness of cross-organizational exercises (Berlin & Carlström, 2015). Berlin and Carlström (2015) found that these large-scale exercises offer few opportunities to practice strategies, and that the format of the exercise was more useful for commanders than for operational staff. These large-scale exercises are also characterized by long waiting times. Magnussen et al. (2017) found that the size of the exercise may affect collaboration between actors, suggesting that smaller exercises could instead improve collaboration. While these exercises can provide valuable learning opportunities, their outcomes are not always structured in ways that enable systematic comparison across organizations or more detailed post-exercise analysis. Berlin and Carlström (2014) suggest that participants should be given the possibility to identify weaknesses and explore alternative approaches.

Against this background, there is growing interest in more flexible and data-driven exercise formats that can complement existing practices. Such approaches aim to support scalable and repeatable scenarios while enabling richer documentation and analysis of how crisis management unfolds over time. In this paper, "scalable," refers to an approach designed to accommodate exercises conducted by a single individual, a small team, or multiple teams simultaneously, ensuring flexibility and applicability across varying organizational sizes and resource levels. This scalability allows for consistent training and evaluation regardless of the number of participants. This

work aims to lay the foundation for a method that can be used for both those who are new to their roles and those with extensive experience, thereby accommodating a range of expertise levels. This inclusive design ensures that all participants benefit from the exercises, fostering continuous learning and improvement across different experience levels. Our goal is to lay the foundation for a method that can be used by both exercise leaders and participants to create conditions conducive to reflective learning. By enabling exercises based on the same scenario, the method promotes comparative reflection, which in turn strengthens the organizational learning process and contributes to a deeper understanding of how crisis management evolves. For this exploratory study, the research question is how AI can be used to support micro-exercises and thereby serve as a methodological approach to crisis training by enabling parallel exercises based on a shared scenario and fostering comparative reflection to enhance organizational learning. This paper contributes by establishing conditions that enable organizations to conduct exercises that are not necessarily costly but can be easily implemented within a standard workweek. At the same time, the method is scalable and adaptable for larger exercises.

RELATED RESEARCH

Organizational Learning and Reflective Practice

The term "organizational learning" is used when learning progresses to the organizational level, resulting in the establishment of new routines and standards within the organization. Organizational learning is characterized as a dynamic process comprising four stages: intuiting, interpreting, integrating, and institutionalizing. The interpreting phase functions as a bridge between the individual and group levels, whereas integrating serves as a connection between the group and organizational levels (Crossan et al., 1999). Argyris (1977, p. 3) states that "Organizational learning is a process of detecting and correcting errors". Learning is enacted by individuals, and an organization learns by merging each member's understanding of the organization as a construct with shared routines, norms, and governing variables. Argyris (1977) distinguishes between single-loop and double-loop learning. Single-loop learning can be understood as the detection and correction of errors without understanding the cause of the error. Double-loop learning, on the other hand, questions assumptions and governing variables. Double-loop learning leads to behavioral changes as a result of cognitive changes (Auqui-Caceres & Furlan, 2023). In this manner, organizational learning becomes embedded in routines, plans, governing variables, and theories-in-use.

Knipfer et al. (2013) contend that organizational learning is fundamentally rooted in both individual and team learning, with reflection serving as the primary catalyst for this learning process. Reflection provides the opportunity to better understand one's own work and is thereby a driver of behavior change. This can be executed at the three levels of organizational learning: individual through feedback, group through assessment, and organization through evaluation, which are intertwined with each other (Järvinen & Poikela, 2001). Thereby, reflective processes serve as tools for making implicit knowledge embedded in contexts explicit (Järvinen & Poikela, 2001).

Crisis Management Exercises

Crisis management exercises are a central mechanism for building resilience (Weick & Sutcliffe, 2015) and developing preparedness and capabilities necessary for crisis management (Mitroff et al., 1988). Exercises are an important tool for strengthening the participants' capability to act during a crisis (Grunnan & Fridheim, 2017). There are various forms of exercises, such as tabletop, simulation, functional, and full-scale exercises. Tabletop exercises are a discussion-based form of exercise. Instead of acting, participants discuss how they intend to act in the exercise scenario (Van Laere & Lindblom, 2019). Through exercises, participants can, among other things, rehearse roles and test plans and procedures. Van Laere and Lindblom (2019) found that varying the training formats over time creates a variety of learning that impacts individuals, groups, and organizations when these are part of a coherent series of exercises. Brief and focused collaborative exercises can significantly contribute to organizational learning while simultaneously enhancing the perceived utility of such exercises and increasing motivation to engage in preparedness activities (Venemyr, 2025). Training for emergency management often takes place in the form of tabletop exercises, functional exercises, or full-scale exercises (Perry, 2004). Tabletop exercises are a good exercise format for learning skills such as understanding roles, forming interagency teams, and dividing responsibility (Elvegård & Andreassen, 2023). However, there is less research and guidance on how the learning from these exercises can be operationalized and translated into practical capabilities and organizational learning. Thus, although tabletop exercises are perceived as valuable, there is a lack of clear methods for systematically transforming insights and experiences from the exercises into concrete improvements in crisis management capacity.

Gleason (2014) argues that smaller and more focused exercises are important for improving preparedness and

enhancing inter-agency coordination, as they afford a deeper understanding of strategic issues. The exploratory study at hand investigates the feasibility of using technology as a methodological tool in tabletop exercises to create small and focused exercises. Building on the work of Borglund et al. (2025), who demonstrated that Automatic Speech Recognition (ASR) can produce detailed transcripts from recorded communications within Emergency Operations Centres (EOCs), we explore how ASR, combined with AI-assisted tools, can systematically capture and analyze interactions during these exercises. The aim is to enhance the capacity of emergency preparedness actors by transforming exercise interactions into actionable insights, thereby improving organizational learning and crisis response capabilities through systematic, technology-enabled evaluation and reflection.

RESEARCH METHOD

This section explains the overall design of the research, including the data collection and the technologies used, though as a work in progress, the paper does not cover all aspects of the study. The preliminary findings are presented as results. This research is an exploratory study consisting of two separate data collections, which are described in chronological order. Both phases apply a qualitative research approach (Creswell, 2013), as the results are not measured using metrics and the collected data are not analyzed quantitatively, for example using mathematical measures. The research is carried out by researchers with expertise in information management, command and control, and crisis communication.

This research is based on the initial evaluation of how small, parallel micro tabletop exercises can be enhanced and supported by off-the-shelf AI-tools. Below we begin by describing the technology used, followed by the setup and scenario. Finally, we describe how the data have been collected.

Technical and Practical Setup

The technical and practical setup for this research was as follows.

This study is based on conducting a tabletop exercise involving multiple groups working in parallel, and the staff/groups that were to be trained were placed in a small room. The room and work assignments were intended to simulate work in an Emergency Operations Centre. In the room, a high-quality recorder was placed. The recorder was started once the staff who were to work in the EOC had received the instructions for the exercise and an initial introduction to the scenario. The participants were fully aware that their work was being recorded and that it would be transcribed using AI and subsequently analyzed using AI tools. To create a safe environment, they were instructed to work as a fictional municipality, Allmänsta. After the training day, OpenAI Whisper was used as an open-source automatic speech recognition (ASR) system to transcribe the recordings from crisis and emergency management work in EOCs into different text files. Whisper was installed locally on the computers used by the researchers.

OpenAI Whisper provides the following files after processing an audio file in the version used in this study:

- Transcription Text File (.txt) – a plain text file containing the transcribed speech without timestamps;
- SubRip Subtitle File (.srt) – a subtitle file with timestamps, commonly used for video captions;
- WebVTT File (.vtt) – another subtitle format similar to .srt, but often used in web applications;
- JSON File (.json) – a structured file containing timestamps, confidence scores, and segment details.

The text files were later uploaded to different AI chatbots, primarily OpenAI ChatGPT-4 and ChatGPT-5 to analyze the work carried out by the groups. Llama 3, a large language model from Meta that can be used on standalone computers, was also tested. Llama was not powerful enough for this purpose. One aim of this research is to explore what can be achieved with publicly available solutions without developing a custom solution. This means that the research was designed to transcribe recordings, rather than using a solution in which Whisper transcribes speech in real time.

Data Collection

Data were collected during two separate occasions targeting two different audiences and groups.

Data Set One

Data set one included a total of six recordings from two tabletop exercises with three municipalities, which was organized by the research team. In December 2024, three municipalities were invited to The Swedish Civil

Defence and Resilience Agency (MCF) Sandö[1] training facility. The three municipalities have a crisis organization that is activated during a large crisis, and, if needed, they will organize themselves to establish command and control. They were offered the opportunity to participate in two shorter tabletop exercises (FEMA – Emergency Management Institute, 2003) in which we were going to record their work in the fictional EOC. When designing the research, the plan was initially to use ChatGPT-4, and to minimize the risk of sharing sensitive data, a fully fictional scenario was created. An advantage of this was that the same scenario could be used for the three municipalities. The scenario was about a suspected sabotage on a water tower in the municipality Allmänsta. The representatives from the three municipalities were expected to act according to their roles in their municipal crisis organization. Each crisis organization was responsible for providing a basis for decision-making to the fictional head of Allmänsta municipality. The representatives were also asked to maintain their own documentation for further evaluation of the micro-exercises.

The scenario was divided into two parts, Part 1 started before lunch, and Part 2 after lunch. Each part lasted 90 minutes, during which the researchers acted as participants and introduced new information throughout. The representatives from the three municipalities were placed in separate rooms, each equipped with whiteboards and large computer screens. When they started their work, a voice recorder began recording all voice activity in the room and stored it in a raw format. After the two exercises, the recordings were processed in Audacity to minimize noise and other unwanted sounds. The three municipalities all used the same scenario and no researcher interfered except to provide additional input to the scenario when needed.

Data Set Two

Data set two was collected in March 2025 using the same scenario as in data set one. This time, it was not municipal crisis management professionals who participated in the exercise. Instead, the aim was to examine how well this form of micro-exercise could be used with individuals who do not have the same level of experience in managing major crises. Students from the three-year Risk and Crisis Management program at Mid Sweden University were therefore selected. It is a bachelor's program, and the students who participated were in their fourth semester. The program aims to provide students with a broad social science knowledge base, enabling them to work in positions responsible for crisis management in both public and private organizations. The class was divided into three groups of five students each and informed that they were now representing staff from a municipal crisis management organization. The groups were allowed to organize themselves within the EOC based on their own prior understanding of how a municipal EOC can be structured. Just as in the exercises conducted with the municipalities, the only involvement of the researchers during the exercise occurred when scenario injects were introduced.

RESULT

The application of Automatic Speech Recognition (ASR) and AI-assisted summarization produced transcripts that enabled systematic comparison of how different trained EOC groups constructed meaning during a shared crisis scenario. Across the total of nine recorded tabletop sessions, the AI-generated documentation provided detailed empirical material for identifying both similarities and differences in sensemaking strategies.

Using ChatGPT for EOC Work Assessment

From the transcripts, it was evident that all groups collectively, through iterative interpretation of incoming information, created a mutual understanding and provided recommendations for decisions. The summaries generated by ChatGPT-4 reflected the overall storyline of the scenario and described key decisions, discussions and questions raised. This indicates that AI-generated transcripts can provide a reconstruction of the group sensemaking process without direct observations in the room.

Another similarity between the EOCs was the identification of small but potentially important details. When asking questions about the data, analysis showed that participants in the groups handled fragmentary cues through collaborative clarification. Querying the transcripts enabled rapid retrieval of such details (such as information about a suspected vehicle), indicating that automatic transcription can support analysis of how groups worked during the exercise.

All groups engaged in temporal sensemaking through their iterative workflow making sense of new cues, making decisions, and thereby shaping how the situation was interpreted. The availability of timestamped subtitle-format

[1] MCF Sandö refers to a training and education facility operated by The Swedish Civil Defence and Resilience Agency. https://www.mcf.se/sv/utbildning--ovning/utbildning/var-utbildningsverksamhet/sando/

outputs enabled an understanding of when a particular issue was raised, when actions were initiated, and how long deliberations lasted. This indicates that AI-generated transcripts can also be used to compare processes across groups, for example how different groups interpreted situations or changed their decisions.

The transcripts enabled the production of coherent event summaries. The transcript quality was sufficient to capture both what was occurring as well as to follow the decision-making process. Importantly, the summaries aligned closely with the participating municipalities' own documentation provided during the exercise and with their after-action reflections, suggesting that the generated material is a reliable representation of the collective sensemaking.

The transcripts from the exercises have been used to evaluate similarities and differences in working methods. Since the scenarios were fictional, ChatGPT was used to compare the different exercise groups. For example, the working methods described in the transcripts, which consist of all spoken communication from the EOC work, were evaluated. The different groups were compared with one another to identify similarities and differences in how they approached a similar scenario. Questions were also posed about how well the described working methods corresponded to the normative approaches outlined by the Swedish Civil Defence and Resilience Agency in the document "Common Guidelines – A Framework for Command and Collaboration in Societal Disruptions."

Using AI in this way allows for the identification of patterns in working methods that would otherwise have required researchers to observe all activities of the exercise groups. Acting as a non-participant observer during an exercise creates a risk of indirectly influencing the group's decision-making. By recording their work, it was possible to analyze it afterwards and identify patterns and decisions that might otherwise have been overlooked. One example is how the three participating municipalities interpreted the information they received in Scenario 2. It became apparent that one of the municipalities chose to adopt a more restrictive and precautionary approach, which is not incorrect from an exercise perspective; however, when their interpretation of the situation was compared with the others, differences emerged that can serve as a basis for discussion after the exercise.

It was also notable that ChatGPT was able to summarize the actions taken by the three student groups and identify deviations from the Swedish Civil Defence and Resilience Agency's "Common Guidelines – A Framework for Command and Collaboration in Societal Disruptions." There were weaknesses in the groups' strategic thinking, their planning for endurance, and in their leadership cycles, such as situation awareness → direction → action → follow-up. However, considering that these were students with nearly half of their education still remaining, they handled a difficult scenario with limited preparation better than expected. In this case, ChatGPT contributed to enabling a more holistic assessment of their performance and helped avoid excessive focus on details.

A key aspect of the analysis involved examining similarities and differences in how the three municipalities and the three student groups approached the task when experimenting with different prompts in ChatGPT, the focus was on how various clues were interpreted and handled. One such clue was that a witness had seen a car leaving the area around the water tower at high speed. Another detail was the information that several plastic containers had been found near the water tower. Nothing indicated that these were connected to the break-in, but there was also nothing to suggest otherwise.

Differences were also identified in how decisions were formulated. Some groups—particularly one of the municipalities—were very formal and explicit in their decision-making, clearly separating formal decisions from their internal reasoning. Of course, students cannot be expected to make decisions in the same way as municipal officials. Nevertheless, it became evident that several key decisions were made by all the participating groups. Since the scenario involved suspected sabotage of the water supply, one such decision was to initiate some form of emergency water solution. The municipalities used more formal language, whereas the student groups expressed the same intent in less formal terms, but they were equally swift in making decisions.

ChatGPT also helped clarify and trace the pathways leading up to specific decisions. By prompting questions about the underlying basis and reasoning behind certain decisions, it was possible to gain a clearer understanding of how these were handled. Because the text files were of manageable size, it was also possible to return to the original transcribed material to verify aspects highlighted by ChatGPT.

DISCUSSION

Reflection serves as the primary catalyst for the organizational learning process (Knipfer et al., 2013). However, reflection demands time, memory, and the availability of shared materials. In this endeavor, the objective was to develop a method that facilitates both single-loop learning and organizational learning through double-loop learning. At this initial stage of method development, the results indicate single-loop learning. However, there is clear potential to advance towards double-loop learning as the process evolves and deeper organizational insights are incorporated. The method of using AI-supported micro-exercises could facilitate double-loop learning by identifying systemic flaws and thereby establishing a robust foundation for organizational learning. Although

current results predominantly reflect single-loop learning, there is potential to refine this method to achieve double-loop learning. For instance, the results show that ChatGPT could identify deviations from standard guidelines, which may indicate systemic issues beyond mere errors. Such deviations could prompt organizations to critically evaluate and potentially revise the frameworks or assumptions underpinning their crisis management strategies. Furthermore, the transcripts generated enable an understanding not only of the decisions made but also of the rationale behind them. This capacity to delve deeper into decision-making logic allows for the examination and challenge of underlying mental models and assumptions.

Effective documentation is essential for learning from crisis exercises, as it facilitates reflection, comparison, and after-action review. However, documentation practices, whether for exercises or real incidents, are often incomplete. This shortcoming limits an organization's ability to analyze decision-making processes, collective sensemaking, and the governing variables that underpin actions. Using audio recordings and transcripts can address this issue by providing a broader data foundation for evaluating exercises, offering feedback, and conducting assessments.

The transcripts serve as a reflective resource, functioning as a soundboard that facilitates structured reflection on action. In this context, the groups conducted their exercises simultaneously but in separate rooms, without collaborating during the exercise. Although their experiences may vary, they all encountered the same challenge. Comparing different groups addressing the same scenario enhances organizational learning by fostering a deeper understanding of diverse approaches and outcomes. The transcripts, along with their summaries and the questions posed to this material, provide a robust foundation for discussion. Within an organization, such discussions can reveal variations in sensemaking strategies, highlight best practices, and identify areas where assumptions differ among teams. This method facilitates the exploration of various interpretations of the assignment. While the interpretations themselves may be of less interest, the reflective discussions they mirror are of significant value. This internal scrutiny affords knowledge sharing, reveals strengths and weaknesses, and helps to define the organization's objectives during crisis management. The current study examined reflections at the group level; however, the potential application of this method at both individual and organizational levels appears promising. The findings illustrate the effectiveness of AI-supported micro-exercises as a versatile and adaptable training format, consistent with Van Laere and Lindblom's (2019) emphasis on the significance of diverse training methods.

This reflective process can also be conducted between organizations. Such comparisons between organizations are rarely supported due to a lack of time and comparable documentation. When multiple organizations engage, working with the same scenario, a unique opportunity arises. It will be possible for organizations to train their staff and, during the evaluation of the exercise, discuss the actions taken and the decisions made not based on individuals but from the perspective of the organization's objectives, routines, plans, governing variables, and theories-in-use. Despite the dataset's limitations, the results indicate that this method effectively enhances the ability to train staff, align actions with organizational plans, and foster continuous improvement through reflective evaluation grounded in shared understandings rather than individual actions. Furthermore, when this method is applied across multiple organizations, it can uncover differences stemming from each organization's unique goals and objectives.

It is acknowledged that the study involving students is distinct from the dataset collected from the municipalities. Nonetheless, the application of the method in an educational context is compelling. However, further work is necessary at this stage. These preliminary results from this ongoing work demonstrates the method's flexibility and its potential to adapt to diverse data sources. To enhance validity, future research should focus on refining the approach when applied in educational settings.

The use of AI in crisis exercises requires careful consideration to ensure ethical and secure implementation. All participants must be fully informed about the use of AI throughout the exercise, including its capabilities, limitations, and the nature of the data involved. Transparency helps build trust and ensures that participants understand how AI contributes to the exercise, avoiding misunderstandings or unintended consequences. Additionally, organizations must implement strict data governance measures to safeguard all information generated or used during these exercises. This includes controlling access, securing data storage, and preventing data leakage that could compromise sensitive information or organizational integrity.

In this context, deploying a local installation of AI offers significant advantages by keeping all data within the organization's own infrastructure. This approach minimizes risks associated with data transmission over external networks and helps maintain compliance with privacy regulations and internal security policies. For organizations interested in testing AI in crisis simulations, this localized setup provides a practical model for balancing the benefits of AI assistance with the imperative of data confidentiality. It also allows for tailored customization of AI tools to fit specific organizational needs and security requirements, thereby enhancing both the effectiveness and safety of crisis management exercises. At present, the required level of computing power for running a local

installation of an open LLM such as LLaMA 3 or Mistral has not been fully evaluated. Using this method to formulate more advanced questions requires careful consideration of the large language model and its training data. The use of a large language model was intended to support the analysis of the work as a whole, not merely to help identify anomalies in how the groups operated. Therefore, it is likely that small-scale instances of open LLMs will not be sufficient, but this requires further evaluation.

Discussion About the Used Setup

In this study, Whisper, ChatGPT-4 and ChatGPT-5 were used. The focus has not been on prompt engineering, nor on conducting detailed analyses of whether errors in Whisper's transcriptions lead to inaccuracies. As this article is being finalized, additional transcription technologies have emerged. For example, the National Library of Sweden has developed an extension to Whisper incorporating 50,000 hours of Swedish data to improve accuracy.

It is important to recognize that this is ongoing research, and that significant methodological challenges remain. One such challenge is determining how the quality of both transcriptions and the formulation of prompts affects the outcomes. It is also likely that, for this method to gain wider adoption, either a locally deployed LLM or one that can guarantee that no information is exposed to unauthorized parties will be required. This effectively excludes most commercial LLMs, as few organizations will trust assurances from Meta, OpenAI, Google, or Mistral when it comes to handling sensitive information. It is also important to emphasize that the proposed tool should not be seen as providing a single truth, but rather as highlighting nuances and contributing to other forms of training and practice.

In a crisis exercise, the work involves much more than what is actually said, and this is clearly a limitation of which we are aware. For example, how situational awareness is shared and visualized on screens, how whiteboards are used, and similar aspects are not considered within the scope of this article. Therefore, it does not provide a complete picture that can be understood; rather, the main contribution of this study lies in the decisions made and the reasoning behind them.

When exercise leaders are not physically present in the room observing, a different kind of interaction also takes place, in which participants do not seek confirmation from exercise leaders through glances and body language. This means that the proposed method should be seen as a complement for training and evaluating organizations in crisis management. The method has its shortcomings, just as traditional tabletop exercises with facilitators physically present do.

CONCLUSION

The article demonstrates how AI can be used as a method to enable scalable, comparable, and reflective micro-exercises to enhance organizational learning. Limited time and scale make a systematic learning mechanism essential for extracting organizational value from training activities. Future work should explore how the method can be integrated into regular municipal training activities and how automated analysis can be combined with human facilitation. The method should also be further developed to support all three levels of organizational learning.

ACKNOWLEDGMENTS

We thank our reviewers for insightful comments that helped improve the paper.

REFERENCES

Ansell, C., Boin, A., & Keller, A. (2010). Managing transboundary crises: Identifying the building blocks of an effective response system. *Journal of Contingencies and Crisis Management*, *18*(4), 195–207. https://doi.org/10.1111/j.1468-5973.2010.00620.x

Argyris, C. (1977). Double-loop learning in organizations. *Harvard Business Review, 55*(5), 115–125.

Auqui-Caceres, M.-V., & Furlan, A. (2023). Revitalizing double-loop learning in organizational contexts: A systematic review and research agenda. *European Management Review*, *20*(4), 741–761. https://doi.org/10.1111/emre.12615

Berlin, J. M., & Carlström, E. D. (2014). Collaboration exercises—the lack of collaborative benefits. *International Journal of Disaster Risk Science*, *5*(3), 192–205. https://doi.org/10.1007/s13753-014-0025-2

Berlin, J. M., & Carlström, E. D. (2015). Collaboration exercises: what do they contribute? *Journal of Contingencies and Crisis Management*, *23*(1), 11–23. https://doi.org/10.1111/1468-5973.12064

Boin, A. (2009). The new world of crises and crisis management: Implications for policymaking and research. *Review of Policy Research*, *26*(4), 367–377. https://doi.org/10.1111/j.1541-1338.2009.00389.x

Borglund, E., Granholm, M., Johansson, C., & Jonriksson, P. (2025). Using automatic speech recognition for documenting work in municipal emergency operations centers. *Proceedings of the 22nd International ISCRAM Conference.*

Comfort, L. K. (2007). Crisis management in hindsight: Cognition, communication, coordination, and control. *Public Administration Review*, *67*, 189–197.

Creswell, J. W. (2013). Qualitative inquiry and research design: Choosing among five approaches (3rd ed.). SAGE.

Crossan, M. M., Lane, H. W., & White, R. E. (1999). An organizational learning framework: From intuition to institution. *Academy of Management Review*, *24*(3), 522–537. https://doi.org/10.5465/amr.1999.2202135

Elvegård, R., & Andreassen, N. (2024). Exercise design for interagency collaboration training: The case of maritime nuclear emergency management tabletop exercises. *Journal of Contingencies and Crisis Management, 32,* e12517. https://doi.org/10.1111/1468-5973.12517

Eriksson, P., & Hallberg, N. (2022). Crisis management as a learning system: Understanding the dynamics of adaptation and transformation in-between crises. *Safety Science*, *151*, 105735. https://doi.org/10.1016/j.ssci.2022.105735

Fast, L. C., Weaver, J. F., Miller, S. D., & Ferrin, T. E. (2016). Training effects on emergency management activation response. *Journal of Organizational Psychology*, *16*(1), 71–77.

FEMA. (2025, December 12). NIMS components - guidance and tools https://www.fema.gov/emergency-managers/nims/components

Gleason, J. J. (2014). Getting big results by going small - the importance of tabletop exercises. *International Oil Spill Conference Proceedings*, *2014*(1), 114–123. https://doi.org/10.7901/2169-3358-2014.1.114

Grunnan, T., & Fridheim, H. (2017). Planning and conducting crisis management exercises for decision-making: the do's and don'ts. *EURO Journal on Decision Processes*, *5*(1), 79–95. https://doi.org/10.1007/s40070-017-0065-0

Järvinen, A., & Poikela, E. (2001). Modelling reflective and contextual learning at work. *Journal of Workplace Learning*, *13*(7–8), 282–290. https://doi.org/10.1108/13665620110411067

Knipfer, K., Kump, B., Wessel, D., & Cress, U. (2013). Reflection as a catalyst for organisational learning. *Studies in Continuing Education*, *35*(1), 30–48. https://doi.org/10.1080/0158037X.2012.683780

Magnussen, L. I., Carlstrøm, E., Sørensen, J. L., Torgersen, G.-E., Hagenes, E. F., & Kristiansen, E. (2017). Learning and usefulness stemming from collaboration in a maritime crisis management exercise in Northern Norway. *Disaster Prevention and Management*, *27*(1), 129–140. https://doi.org/10.1108/DPM-06-2017-0131

Mitroff, I. I., Pauchant, T. C., & Shrivastava, P. (1988). The structure of man-made organizational crises. *Technological Forecasting and Social Change*, *33*(2), 83–107. https://doi.org/10.1016/0040-1625(88)90075-3

Perry, R. W. (2004). Disaster exercise outcomes for professional emergency personnel and citizen volunteers. *Journal of Contingencies and Crisis Management*, *12*(2), 64–75. https://doi.org/10.1111/j.0966-0879.2004.00436.x

Quarantelli, E. L. (1978). Uses and problems of local EOCs in disaster*s* (Preliminary Paper No. 53). *Disaster Research Center*, University of Delaware. http://udspace.udel.edu/bitstream/handle/19716/428/PP53.pdf?sequence=3

Van Laere, J., & Lindblom, J. (2019). Cultivating a longitudinal learning process through recurring crisis management training exercises in twelve Swedish municipalities. *Journal of Contingencies and Crisis Management*, *27*(1), 38–49. https://doi.org/10.1111/1468-5973.12230

Venemyr, G. O. (2025). Making the most out of a little: How compact collaboration exercises for crisis management teams can lead to organisational learning. *International Journal of Disaster Risk Reduction*, 127, 105677. https://doi.org/10.1016/j.ijdrr.2025.105677

Weick, K. E., & Sutcliffe, K. (2015). *Managing the unexpected* (3rd ed.). John Wiley & Sons.