# Training, Test and Experimentation: A classification of command post exercises

# Erich Heumüller

## Sebastian Richter

Universität der Bundeswehr München erich.heumueller@unibw.de

Universität der Bundeswehr München s.richter@unibw.de

## Ulrike Lechner

Universität der Bundeswehr München ulrike.lechner@unibw.de

#### **ABSTRACT**

Based on a state of the art analysis of exercise classifications and empirical findings from an Action Research approach this paper presents a classification for command post exercises, that addresses identified shortcomings concerning practical support for goal-achievement and evaluation in exercise design. The authors' classification distinguishes between Training, Test and Experimentation Exercises, which are characterized by the following aspects: goal and purpose of exercise, participants, evaluation content and output, evaluation methodology, role of observers, scenario complexity, potential for organizational innovation and results. The classification was developed in an Action Research approach with an empirical basis of four command post exercises. Results indicate high benefit for exercises and the approach is perceived rather simple, easy to understand and to apply.

### **Keywords**

Exercise classification, exercise goal, exercise evaluation

## INTRODUCTION

Disaster management exercises are valuable instruments for disaster preparedness in a vulnerable complex world. They facilitate several functions like the validation and adaption of emergency plans, the evaluation of equipment and preparedness, the training of processes and staff cells and the demonstration of disaster response skills (US Dept of Homeland Security, 2007). However, successful exercising for disaster response is limited by several factors as e.g. lack of financial resources or the opportunity to exercise, particularly in Germany, where disaster response is characterized by a voluntary service. Thus, exercises are itself invaluable means and it is necessary to think of design exercises to be efficient and useful.

The authors developed a classification approach for command post exercises to support practitioners in exercise design as the existing guidelines and classifications for exercise design (e.g. Australian Emergency Management Institute, 2012; Emergency Management Division - Michigan State Police, 1998; Oregon, 1997; Payne, 1999; US Dept of Homeland Security, 2007) lack of a respective perspective of goal- and evaluation-orientation.

The classification is part of ongoing research. To increase efficiency and benefit of exercises the authors develop a framework for exercise design. The concept and structure - a V-Model with a content and an evaluation part in three levels of granularity - were presented in ISCRAM 2012 (Heumüller et al. 2012a). The key objective of this framework is to design exercises goal-oriented and evaluation-driven. This means that the framework offers a systematic way for a scenario-based exercise design to ensure, that all scenarios and injects follow the exercise goal and enable a specific evaluation. The framework provides guidelines and templates to support exercise design. Another part of this research project is a conceptual model of staffs to guide performance assessment of staffs (Heumüller et al., 2012b).

Subsequently, research design, exercise design challenges and results of a state of the art analysis are presented. Then, the new classification and findings from its application are offered. The paper finishes with a discussion.

## **METHOD**

The authors follow a Canonical Action Research approach (Davison et al., 2004; 2012) to design a framework for exercise conceptualization emphasizing exercise goal-achievement and exercise evaluation. This paper

concentrates on exercise classification, which is one component of the framework. The research is based on four command post exercises (Table 1) and the collaboration with staffs of two different disaster response organizations in Germany. FRANKENSTURM involved six staffs and the others five staffs.

Data collection was done mainly in participatory observation (e.g. during workshops, discussions or exercise conducting) as Baskerville and Wood-Harper (1998) describe it in Action Research context. Additionally, the exercise participants' performance was assessed on the judgment by experienced staff members, who acted as observers or referees. The authors prepared evaluation reports, which contain the evaluation of the participants and the findings regarding problems, processes, lessons learnt, etc.

Exercise	Organization	Date	Involvement	Topic	Scenario	Phase in Action Research
GERETSRIED	THW	22 <sup>nd</sup> -24 <sup>th</sup> October, 2010	Observer during conducting Evaluation	Change of shifts during an on-going mission	Flood Scenario in southern Bavaria	Diagnosing Action Planning
GROSSER KREIS	KVK	15 <sup>th</sup> /16 <sup>th</sup> July, 2011	Observer and advicing during conceptualization and conducting Evaluation	Not specified	Fire disasters in the forests around Nuremberg (northern Bavaria)	
FEUERBALL	THW	7 <sup>th</sup> -9 <sup>th</sup> , October, 2011	Conceptualization and conducting Evaluation	Management of an assembly area	Flood scenario in Munich	Action Taking Evaluation Specify Learning
FRANKEN- STURM	THW	5 <sup>th</sup> -7 <sup>th</sup> October, 2012	Conceptualization and conducting Evaluation	Change of shifts	Flood and whirlwind near Selb (northern Bavaria)	

**Table 1: Examined Command Post Exercises** 

The coordination and communication units (*Fachgruppe für Führung und Kommunikation* (FK)) of our research partner – the Federal Agency of Technical Relief (THW) a German disaster response organization under superior of the Federal Ministry of the Interior – are such staffs. A FK is assigned to coordinate THW-forces or forces of other disaster response organizations and to provide communications tools. The second partner is the *Kreisverbindungskommando* (KVK) FÜRTH, located in northern Bavaria and responsible to coordinate Bundeswehr-forces in case of disaster. KVKs are liaison units at county level and help to integrate Bundeswehr-forces in disaster response operations during disaster. KVKs are organized as staffs with reservists serving in a typical military structure. In a KVK-staff there are usually three to six members engaged and in a FK seven up to 15 members.

## **CHALLENGES IN EXERCISE DESIGN AND EVALUATION**

The authors observed several challenges during their involvement in the exercises. Some exemplary findings about exercise design are listed below:

- A one-size-fits-all-approach for exercise design is insufficient and inefficient for exercise evaluation.
- The role of observers depends on the exercise goal and has to be clarified early in exercise design.
- Evaluation has to be integrated and organized in exercise design from the beginning.
- A continuous calibration of all measurements during exercise design to measure exercise goal precisely and support exercise evaluation is necessary.

Based on these findings the authors assumed a classification approach to be helpful for exercise planners to categorize exercise goals in general and to support a respective goal-orientation in the exercise design (e.g. for trying something new, to train or to test the participants).

## STATE OF THE ART

After GERETSRIED current literature about exercise classifications and guidelines was analyzed. The authors found out, that exercise classifications are often incorporated into several guidelines. Basically, nine different exercise types were identified (Table 2). All approaches have in common a distinction according scale and complexity despite some differences in notation or detail. However, no evaluation aspects within the classifications could be identified. Although existing classifications often provide recommendations for purpose orientation, guidelines for exercise evaluation are missed.

Authors/ Articles		- US Dept of Homeland Security, 2007	- Renner, 2001	- Callan, 2009, - Australian Emergency Management Institute, 2012;	- Payne, 1999	United States Federal Emergency Management Agency, 2012, - Trnka and Jenvald, 2006, - Klein et al., 2005, - Perry, 2004, - Peterson and Perry, 1999, - Ohio Emergency Management Agency, 1999, - Texas Department of Public Safety, Division of Emergency Management, 2000
Seminar	Orientation Exercise	Seminar	Information Session	Seminar, Agency Presentation		
Workshop		Workshop		Workshop		
Walkaround			Walkaround			
Paper-Feed Exercise					Paper-feed Exercise	
Tabletop Exercise	Tabletop Exercise	Tabletop Exercise	Tabletop Exercise	Syndicate, Hypothetical Exercise	Table-top Exercise	Tabletop Exercise
Game		Game				
Drill	Drill	Drill				
Functional Exercise	Functional Exercise	Functional Exercise	Centre Simulation	Functional Exercise	Communications Simulated Exercise	Functional Exercise
Full-Scale Exercise	Full-Scale Exercise	Full-Scale Exercise	Field Exercise	Field Exercise	Live Exercise	Full-Scale Exercise

Table 2: State of the Art of exercise type classifications

In literature research an analogy in the context of transformation of armed forces was identified, where experiments are conducted to explore new concepts. These experiments have predefined goals like the identification of opportunities for improvement, testing of hypotheses or demonstration of capabilities or processes. Hence, goal-orientation and evaluation are also important issues. Codes of Best Practices classify experiments according aspects of purpose: *Discovery, Hypothesis testing, Demonstration Experiments* (e.g. Alberts and Hayes, 2002; 2005) This approach was transferred to exercises in a disaster response context.

#### A GOAL-ORIENTED AND EVALUATION-DRIVEN APPROACH FOR EXERCISE CLASSIFICATION

The classification distinguishes between Training, Test and Experimental Exercises (Table 3). These exercise types are characterized by the following aspects: goal and purpose of exercise, participants, evaluation content and output, evaluation methodology, role of observers, scenario complexity, potential for organizational innovation and results. Based on empirical findings of the examined exercises, these aspects became apparent as crucial for exercise design. Because of paper limitations, the aspects cannot be presented in detail. In the following, two aspects are briefly presented as examples.

*Goal and Purpose:* Exercises fulfill different goals and purposes. The goal of a Training Exercise is to train or educate participants in a special topic of disaster response as e.g. rescue, water supply or emergency aid to fulfill emergency plans or regulations and to achieve a higher training level in this topic.

The purpose of a Test Exercise is to test participants regarding their practical competences when standard operation procedures are defined or benchmarks are available. The focus of an Experimental Exercise is to analyze and improve procedures, behavior patterns or equipment interactively together with the participants to contribute to the knowledge of the organization to improve disaster preparedness.

The different goals or purposes call for different requirements in exercise design and evaluation methodology.

Evaluation Methodology: Evaluation methodology captures all ideas of measuring, data gathering and data analysis and contains the selection of appropriate methods and the definition measurement moments. The specifications defined in evaluation methodology call for different requirements in exercise content, course and organization. Basically, the exercise content comprises injects, which trigger a specific behavior of the staff. The behavior, which should be analyzed (e.g. trained or tested), is defined by the exercise goal. Therefore, the challenge is to ensure, that exactly this behavior is triggered, which should be analyzed and that it is possible to analyze this behavior at all. The exercise course should be orientated on the defined moments of measurements and the exercise organization accounts for the fulfillment of the planned tasks and measures of evaluation.

A Training Exercise's purpose is to increase the participants' training level. Therefore, it is sound to measure the training level before and after the exercise to identify exercise success. The overall performance assessment in a Test Exercise is usually composed of many single measurements assessing defined evaluation criteria. In an Experimental Exercise the evaluation methodology depends strongly on the evaluation area of interest. To identify the surplus of an innovation, comparison measurements with Test Exercises are an option as well as qualitative assessments. Note that, the innovations one wants to identify in an Experimental Exercise, which is typically resource intensive, are disruptive. Such innovations usually take time to mature and unfold their complete positive potential (Rogers, 2003) and are therefore hard to evaluate in a single exercise.

	Training Exercise	Test Exercise	Experimental Exercise
Goal and purpose	To provide operational competence or authority and handling confidence to participants through practical application of imparted theoretical knowledge.	To test participants regarding practical competences (e.g. Standard Operating Procedures).	To analyze and improve procedures, behavior patterns or equipment interactively together with the participants from an organizational view.
Participants	−Normally novices −Immature regarding process and output	-Trained personnel	-Trained personnel, experts
Evaluation Content and Output	−Clearly defined in regulations or emergency plans −Measureable depending on specifications in regulations (times, patterns, actions, etc.)	–Exact descriptions and measurability otherwise risk of arbitrariness	–If possible, described –Rather qualitative –Probably new measurement criteria are needed
Evaluation Methodology	-Before and after measurement of training level to assess exercise success -Meta analyses of various exercises -Continuous improvement of the participants	–Single-measurements against defined criteria	–Comparison measurements with Test Exercises –Innovations have seldom direct positive effects
Role of Observers	-Trainer or instructor -Actively involved in the exercise -Guide the participants -Have to be familiar with standards	–Act as referees –Supervise compliance of standards –Neutral and passive during exercise	-Active part of the exercise, support experimental process -Advice participants and analyze results and behavior -Qualitative evaluation
Scenario complexity	–Easy or simple scenarios –Focus on training content	–Full bandwidth of complexity (simple to complex) –Complexity increases with iterations	-Full bandwidth of complexity -Simple "implementations" of changes till to the "experiencing" of the potentials in complex processes
Potential for innovation	-Low	-Low	-High
Results	-improvement of operational skills of the participants -Practicing and increase of familiarity with regulations and emergency plans -improvement of quality -Findings about training methodology	-Assessment of equipment, plan, policies, procedures, agreements, training level -Identification of gaps in resources	-Identification of opportunities for improvement -Findings about how things are made and how they can made better

Table 3: Classification of exercises: Training, Test and Experimental Exercises

Subsequently, the grounding of the classification is described. The first presentation of this approach took place at a conference of all Bavarian FK-commanders in May 2011. The authors were actively involved in the conceptualizations of FEUERBALL and FRANKENSTURM, which were guided by the classification. The following impact of the classification was observed:

#### FEUERBALL:

- The regulations about the management of an assembly area built the basis for exercise design: exercise contents and evaluation criteria were derived from them. At GROSSER KREIS, the planning team was uncertain about evaluation criteria or training contents, respective regulations were seldom available.
- Evaluation criteria and moments of measurement were defined in advance. Screenplay and injects
  followed these specifications (e.g. times for fuel requests or briefings). Neither in GERETSRIED nor in
  GROSSER KREIS such specifications were defined sufficiently.
- The planning team developed evaluation guidelines based on regulations. This is in contrast to GERETSRIED (were no checklists were prepared by the planning team) and to GROSSER KREIS (were checklists were rather general). Thus, a valid and comprehensive evaluation was enabled.

# FRANKENSTURM:

- For the first time, exercise course followed evaluation methodology: the core was a before-and-after measurement of identified criteria to analyze the shift change.
- The scenario was conceptualized to enable evaluation, not to produce stress (as in former exercises). The actual scenario content and its amount of injects were rather less important, as it was only used to analyze the shift change (in GROSSER KREIS the attractiveness of the injects was important).
- New processes were identified, which can be further analyzed and documented to contribute to disaster preparedness. In GERETSRIED a new process should be tested, but after exercising no information about the process was available as the evaluation followed the traditional scheme.

The authors analyzed the influence of the classification, reflected exercise design together with the planning teams and questioned the approach (usability, comprehensibility, etc.) in interviews and online surveys after the exercises. Exemplary findings are:

- The approach supported common understanding and served as a first orientation for exercise design.
- The terms and definitions of Training, Test and Experimental Exercise were adopted quickly.
- The approach provides a frame for the planning team to prepare exercise content and evaluation.
- The planning team of FEUERBALL recommended the classification approach to the planning team of FRANKENSTURM. This team perceived a high value using this approach.
- Although the classification was perceived to be logically and comprehensible by the planning teams, professional guidance is required to apply the classification conclusively in exercise design.

#### **DISCUSSION**

This paper presents a classification approach for exercises, which addresses a shortcoming of existing exercise classifications: support for goal-achievement and evaluation-orientation. The classification is rather simple, easy to understand and to apply for practitioners and increases the benefit of exercises. Different planning teams applied this classification, which enabled a valid exercise evaluation and simplified the design. The classification has proved to be successful in two exercises. This classification is part of a more comprehensive approach for exercise design, which provides applicable templates and guidelines for practitioners (Heumüller et al., 2012a). Nevertheless, our empirical base is limited to four command post exercises. Additional research is needed to improve applicability of the classification and to strengthen the empirical validation.

#### **REFERENCES**

- 1. Alberts, D. S., and Hayes, R. E. (2002) "Code of best practice: experimentation," Command and Control Research Program, DTIC Document, pp. 436.
- 2. Alberts, D. S., and Hayes, R. E. (2005) "Campaigns of experimentation: Pathways to innovation and transformation," Command and Control Research Program, DTIC Document, pp. 227.
- 3. Australian Emergency Management Institute (2012) "Managing Exercises Handbook 3," Australian Emergency Management Handbook SeriesBarton, pp. 1–128.
- 4. Baskerville, R., and Wood-Harper, A. (1998) "Diversity in information systems action research methods," European Journal of Information Systems (7:2), pp. 90–107.
- 5. Davison, R. M., Martinsons, M. G., and Ou, C. X. J. (2012) "The roles of theory in canonical action research," Management Information Systems Quarterly (36:3), pp. 763–786.
- 6. Davison, R., Martinsons, M., and Kock, N. (2004) "Principles of canonical action research," Information Systems Journal (14:1), pp. 65–86.
- 7. Emergency Management Division Michigan State Police. (1998) "Disaster Exercise Manual."
- 8. Heumüller, E., Richter, S., and Lechner, U. (2012a) "Towards a framework for command post exercises", Proceedings of the 9th International ISCRAM Conference, Vancouver, Canada.
- 9. Heumüller, E., Richter, S., and Lechner, U. (2012b) "Towards a conceptual model of staffs in disaster response organizations". Proceedings of the 25th Bled eConference, Bled, Slovenia.
- 10. Klein, K., Brandenburg, D., Atas, J., and Maher, A. (2005) "The use of trained observers as an evaluation tool for a multi-hospital bioterrorism exercise," Prehosp Disaster Med (20:3), pp. 159–163.
- 11. Ohio Emergency Management Agency. (1999) "Planning and Exercise Guidance Booklet".
- 12. Oregon, S. O. (1997) "Oregon Version Exercise Design and Evaluation Course,," pp. 1–208.
- 13. Payne, C. (1999) "Contingency plan exercises," Disaster Prevention and Management (8:2), pp. 111–117.
- 14. Perry, R. (2004) "Disaster exercise outcomes for professional emergency personnel and citizen volunteers", *Journal of Contingencies and Crisis Management* (12:2).
- 15. Peterson, D. M., and Perry, R. W. (1999) "The impacts of disaster exercises on participants," Disaster Prevention and Management (8:4), pp. 241–255.
- 16. Renner, S. (2001) "Emergency exercise and training techniques", Australian Journal of Emergency Management (16:2).
- 17. Rogers, E. M. (2003) "Diffusion of Innovations", (5th ed.), Free Press.
- 18. Texas Department of Public Safety, Division of Emergency Management (2000) "Preparedness Standards for Emergency Management in Texas," Austin.
- 19. Trnka, J., and Jenvald, J. (2006) "Role-playing exercise—a real-time approach to study collaborative command and control", *International Journal of Intelligent Control and Systems* (11:4).
- 20. US Federal Emergency Management Agency (2012) "IS-139: Exercise Design," *Independent Study Progr.*
- 21. US Dept of Homeland Security (2007) "Homeland Security Exercise and Evaluation Program (HSEEP) Volume I: HSEEP Overview and Exercise Program Management."
- 22. Virginia Dept. of Emergency Mgmt. (2001) "School Crisis Management Exercise Development Guide"