Realitätsnah oder genau? Triangulative Übungsevaluation: Erfahrungen aus dem REBEKA-Projekt

Patrick Drews; Johannes Sautter; Marcel Kübel

Zusammenfassung

Vorherrschende Standards der systematischen und fairen Evaluation, basierend auf dem Qualitätsmanagement, dem Notfallmanagement und der Evaluationsforschung, sollten auch für die Auswertung von Katastrophenschutzübungen gelten. Dennoch wird bei bisherigen Übungen die Planung, Auswertung und Nachbereitung eher unsystematisch durchgeführt. Im Projekt REBEKA wurden zwei Übungen mit Hilfe eines angepassten triangulativen Methodenmixes nach Mayring (2001) ausgewertet. Dieser Beitrag beschreibt den Aufbau der Übungsevaluation, bestehend aus standardisierter Befragung, Übungsbeobachtung sowie Prozesskennzahlen und gibt Einblicke in die Ergebnisse. Er wirft einen kritischen Blick auf das Vorgehen und reflektiert systematisch verwandte Methoden und die Bedeutung für die Übungsevaluation als Ganzes.

Einleitung

Der Bevölkerungsschutz in Deutschland stützt sich auf etwa 1,7 Millionen ehrenamtliche Einsatzkräfte (BMI 2017). Um Einsatzkraft im Katastrophenschutz werden zu können, vermitteln die einzelnen Organisationen die Grundfertigkeiten und grundlegenden Verfahren in Basis- oder Grundausbildungen. Darüber hinaus werden Weiterbildungen für bestimmte Funktionen vorausgesetzt (Mitschke & Karutz 2017). Neben Aus- und Fortbildungen absolvieren Einsatzkräfte auch **Übungen**, die von hoher Relevanz für die Bevölkerungsschutzausbildung sind. Denn für viele Einsatzkräfte sind sie die einzige Möglichkeit, das vorhandene Wissen in die Praxis umzusetzen (Mitschke & Karutz 2017).

Übungen dienen dem Zweck (1) der Erlernung und Vertiefung von Verfahren (Training), (2) der Überprüfung von Einsatzplänen, Verfahren oder dem Zusammenspiel von Organisationen (Mitschke & Karutz 2017) oder (3) der Anwendbarkeit bestehender Planungen auf andere Szenarien (ISO 22398: 2013). Somit bieten Übungen in Form simulierter Schadenfälle die Möglichkeit, die „Readiness“ (McCreight 2011) zu messen sowie Verbesserungspotentiale von Prozessen und Verfahren der Gefahrenabwehr zu

identifizieren. Darüber hinaus unterscheidet die ISO 22398: 2013 insgesamt neun verschiedene Arten[20] von Übungen. Glass (2006, zit. in: Mitschke & Karutz 2017) unterscheidet Übungen nach den übenden Personen: Stabs-, Stabs-Rahmen und Rahmenübungen, die die Entscheidungsebene betreffen sowie Fach- und Vollübungen, die im Gegensatz dazu auch die operativen Kräfte beinhalten.

Für die Evaluation von Übungen haben sich bislang vor allem Verfahren etabliert, die ausschließlich qualitative Bewertungen von Experten umfassen (DRK 2016). Dabei besteht in diesen Berichten die Gefahr der Subjektivität und Selektivität (Karsten 2016). Die Nutzung von Kennzahlen oder die systematische Befragung der Übungsteilnehmenden finden nur selten statt – ungeachtet deren Anwendbarkeit (Djalali et al. 2014). Das könnte auch an der fehlenden, wissenschaftlichen Basis für eine systematische Übungsauswertung liegen (Perry 2004). Es besteht die Forderung unabhängig von bestehenden Einsatzkonzepten Übungen zu evaluieren (Brauner et al. 2014). Dabei ist es unerheblich, ob es sich um wissenschaftlich-begleitete Übungen oder Übungen der operativen Einsatzkräfte auf Standortebene handelt. In der Praxis bringen Übungsverantwortliche aufgrund hoher Kosten und Aufwand im Übungsdesign möglichst viele Übungsziele und Punkte zur Validierung in eine Übung ein. Eine Kombination verschiedener Methoden zur Übungsevaluation scheint somit aufgrund unterschiedlicher Einflussfaktoren und einer Vielzahl zu erfassender Variablen sinnvoll, um ein ganzheitliches Ergebnis zu erlangen. Damit ließen sich verschiedene Perspektiven der Einsatzbewältigung abdecken. Ein Evaluationsmodell angelehnt an das Triangulationsmodell nach Mayring (2001) könnte vor diesem Hintergrund ein geeigneter Ansatz zur Erfüllung der Anforderungen an eine systematische Vor- und Nachbereitung sowie Auswertung von Übungen im Katastrophenschutz darstellen.

Dieser Beitrag beschreibt die Übungskonzeption (1) und den gewählten Evaluationsansatz (2) sowie die daraus resultierenden Erfahrungen aus zwei Übungen innerhalb des REBEKA-Projektes. Die eigentlichen Übungsergebnisse sollen hier jedoch nicht detailliert dargestellt werden. Vielmehr soll ein Überblick auf den Verlauf der Übungen und die resultierenden die Erkenntnisse aus Evaluationssicht (3) gegeben werden und ein kritischer Blick (4) auf das Vorgehen geworfen werden.

Übungskonzeption im REBEKA-Projekt

Innerhalb des REBEKA Projekts wurden zwei Übungen auf unterschiedlichen Ebenen der Gefahrenabwehr im Rhein-Kreis-Neuss durchgeführt. Ein Starkregen-Ereignis mit

[20] Die allerdings nicht sehr trennscharf unterschieden werden. So werden z.B. Übungsserien als eigener Typ dargestellt, die allerdings eher eine Aneinanderreihung einzelner oder mehrerer der acht weiteren Übungstypen beschreibt.

einer signifikanten Beeinträchtigung der Verkehrsinfrastruktur wurde als Übungsszenario herangezogen. Dabei wurde zusätzlich ein hoher Wasserstand des Rheines angenommen und die Schadensbewältigung durch fehlende Einsatzkräfte im Übungsszenario erschwert.

Übungsziele

Die Übungen im Projekt hatten zum Ziel, die Praxistauglichkeit der im Projekt entstandenen Lösungsansätze[21] zur Steigerung der Resilienz zu testen. Innerhalb der Übung sollte u. a. untersucht werden, inwieweit die Annahmen des Tätigkeitenkatalogs zutreffend sind. Hierzu wurde die Tätigkeit „Füllen von Sandsäcken" ausgewählt und im Szenario Deichverteidigung evaluiert. Zudem sollte der Einfluss beruflich-vorgebildeter Spontanhelfender, sogenannter Rollenerweiterer, auf die Einsatzbewältigung untersucht werden. Darüber hinaus wurden weitere Fragestellungen des Rhein-Kreises Neuss aufgenommen. Im Vordergrund stand dabei die Frage, ob die Einsatzkonzepte auch für das gewählte Szenario mit eigener Betroffenheit anwendbar sind oder ob Schwächen bestehen.

Übungsanlage

Die **Stabs[22]-Rahmenübung** fand am 23.06.2018 im Kreishaus in Grevenbroich mit insgesamt 60 Beteiligten statt. Die Einsatzleitung – der operativ-taktische Stab des Landkreises – und der Krisenstab – die politisch-administrative Komponente – wurden in dieser Übung durch eine gemeinsame Übungsleitung gesteuert. Diese wies die Stabsmitglieder zu Beginn in die Lage ein und spielte im Verlauf weitere Lagemeldungen oder Reaktionen auf Entscheidungen des Stabes der simulierten Einsatzabschnitte ein[23]. Die beiden nach FwDV 100 gleichberechtigten Stäbe sollten in dieser Übung Entscheidungen vorbereiten und umsetzen. Die Übungsleitung fungierte dabei als Gegenstelle, die sowohl das Szenario (repräsentiert durch die untergebenen Einsatzabschnitte) als auch die übergeordneten oder parallelen Führungsstrukturen darstellte. Beide Stäbe sollten sich, wie in der realen Einsatzbewältigung auch, untereinander abstimmen und die Entscheidung in ihrer eigenen Zuständigkeit und Kompetenz treffen. In Abbildung 17 ist das Zusammenspiel der beiden Stäbe graphisch dargestellt.

---

[21] Das Einsatzhandbuch und der Tätigkeitenkatalog (s. bd. Ges. Beitrag).

[22] Stäbe als unterstützendes Gremium von Einsatzleitenden zur Koordination von größeren Einsatzlagen sind ein unverzichtbarer Bestandteil der Führungsorganisation (Heumüller, Richter & Lechner 2014). Den grundsätzlichen Aufbau von Führungsstäben beschreibt die FwDV 100 bzw. deren Äquivalente der anderen Hilfeleistungsorganisationen.

[23] Dieses Vorgehen der Übungsleitung wird als Einspielung bezeichnet.

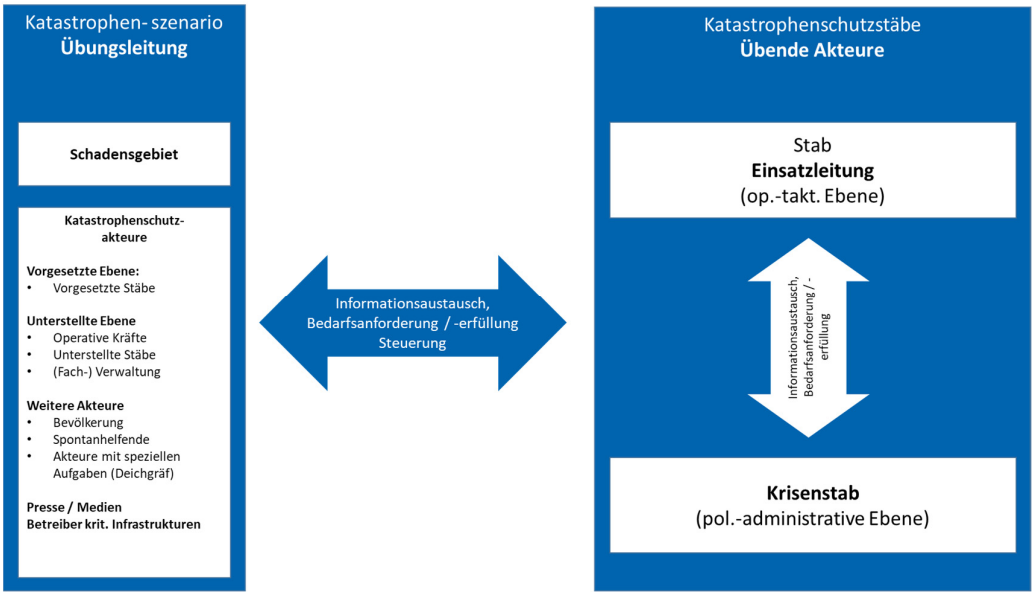

Abbildung 17: Zusammenspiel zwischen den Stäben und der Übungsleitung (erg. Darstellung n. Heumüller, Richter & Lechner (2014))

Die **Feldübung** wurde am 20. Oktober 2018 mit insgesamt 250 Beteiligten an vier Standorten im Rhein-Kreis-Neuss durchgeführt. Der jeweilige Standort repräsentierte dabei einen der drei Aufgaben- und damit Übungsschwerpunkte: (1) Einen Massenanfall von Verletzten (MANV) aufgrund eines Windteufels in einem Zeltlager einer Jugendgruppe auf einem Campingplatz, (2) die Verteidigung eines Deichstücks mit vorherigem Füllen von Sandsäcken und (3) die Errichtung eines Bereitstellungsraumes für 500 Einsatzkräfte als Vorbereitung auf anrückende Einsatzkräfte aus anderen Bundesländern [24]. Am Übungstag wurden zudem als weiterer Schwerpunkt (4) in allen drei vorgenannten Szenarien insgesamt 25 Spontanhelfende[25] eingesetzt. In den Szenarien **Deichverteidigung** und **Bereitstellungsraum** mussten zur Erfüllung der Übungsaufgabe einfache, manuelle Tätigkeiten durch Einsatzkräfte und Spontanhelfende durchgeführt werden. Diese bestanden im Füllen von Sandsäcken und deren Transport in einer Sandsackkette sowie der Vorbereitung von Lunchpaketen während der Übung und der Ausgabe einer Warmverpflegung am Ende der Übung. Im Teilszenario MANV wurden aus dem Spontanhelfendenpool medizinisch-vorgebildetes Personal, mehrheitlich Krankenpflegeschüler ausgewählt, um als Spontanhelfende die Einsatzkräfte zu unterstützen.

[24] Der vierte Übungsort bestand aus dem Krisenstab des Rhein-Kreises-Neuss, der eine eigene, interne Evaluation durchführte und deshalb nicht weiter erwähnt wird.

[25] Diese wurden durch einen im Vorfeld getätigten Aufruf in der Presse und in sozialen Medien für die Übung gewonnen.

Evaluationsansatz

Die Auswertung beider Übungen erfolgte in einem Methodenmix. Dabei wurde ein triangulativer Ansatz (Mayring 2001) auf die Übungsbedürfnisse zugeschnitten: Beobachtung als qualitative Betrachtungsebene, standardisierte Fragebögen sowie die Erhebung von Prozesskennzahlen als quantitative Betrachtungsebene (s. Abbildung 18). Dieses bestand aus einem auf Zeit-, Mengen- und Boole'sche Werten[26] basierenden Kennzahlensystem. Die Beobachtung erfolgte durch erfahrene und in die Aufgabe eingewiesene Einsatzkräfte. Die an der Übung beteiligten Einsatzkräfte und Spontanhelfenden wurden zudem mit Hilfe standardisierter Fragebögen befragt. Durch diesen Ansatz sollten möglichst viele Perspektiven auf das Übungsgeschehen ermöglicht werden.

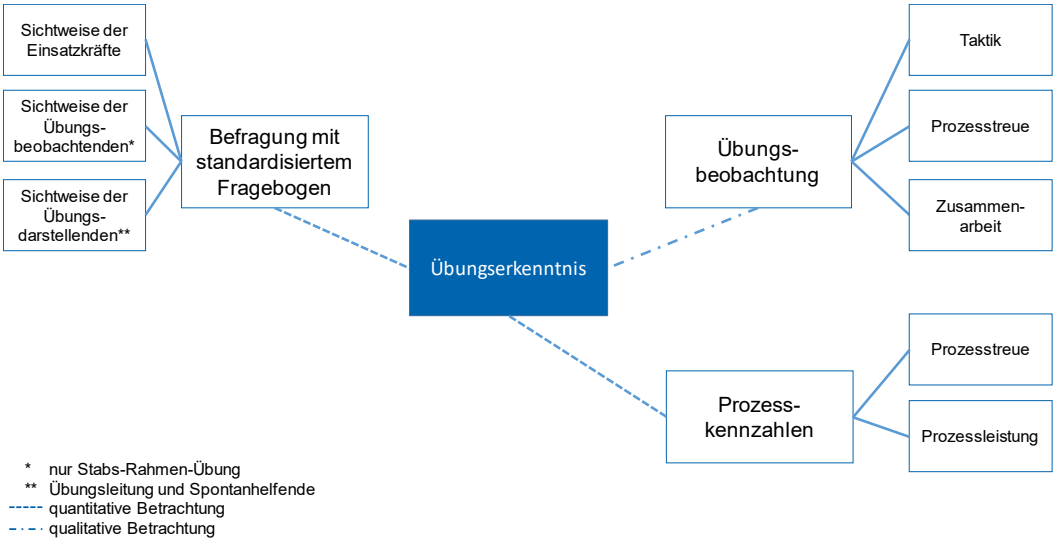

Abbildung 18: Triangulärer Methodenmix (eig. Darstellung)

Beobachtung

Die Beobachtenden wurden vor der Übung in die Evaluation, die Übungszielsetzung und das Evaluationsinteresse eingewiesen. Die Reduzierung von Subjektivität und Selektivität des jeweiligen Übungsbeobachters erfolgte durch Formblätter, auf denen Interpretation und Beobachtung getrennt voneinander notiert werden sollten. Zeitgleich wurde eine Foto- bzw. Kurzvideo-Dokumentation eingesetzt. Die Beobachtenden konnten damit besonders augenfällige Situationen per Smartphone-App direkt dokumentie-

[26] Die Erfüllung eines vorher definierten Kriteriums ist TRUE und die nicht-Erfüllung FALSE.

ren, die dann zentral mit Zeit-, Ortsstempel sowie der jeweiligen Kamera versehen gespeichert wurden. Diese Dokumentation stand auch der jeweiligen Übungsleitung zur Verfügung. Die Beobachtungen während der **Feldübung** wurden mit unterschiedlichen Schwerpunkten durchgeführt. Eine Hälfte der Beobachtenden wurde mit der Beobachtung des allgemeinen Einsatzablaufs betraut. Die Dokumentationsbögen enthielten daher bspw. Fragen zur gewählten Taktik und dem Vorgehen der Einsatzkräfte (Prozesstreue). Die andere Hälfte konzentrierte sich auf das Zusammenspiel zwischen Einsatzkräften und Spontanhelfenden. Die Leitfragen bezogen sich daher allgemein auf den Umgang, die gewählten Tätigkeiten und die Zielgerichtetheit der Einbindung sowie das Verhalten gegenüber Spontanhelfenden (Prozesstreue).

Befragung mit standardisierten Fragebögen

Zur Erhebung der Sichtweise der Einsatzkräfte auf die Übungen wurden im Vorfeld standardisierte Fragebögen erstellt, die am Übungstag zu Übungsende zur Online-Beantwortung freigeschaltet wurden. Auf eine Papierfassung wurde verzichtet, da davon auszugehen war, dass alle Einsatzkräfte über einen Internetzugang – zu großen Teilen über das Smartphone – besitzen. Der Fragebogen bestand in beiden Übungen aus vier Teilen[27], die allerdings unterschiedliche auf die jeweilige Übung angepasste Items enthielten. Alle Fragebögen enthielten Fragen zum Übungsschwerpunkt (Einsatzleitung oder Krisenstab bzw. Einsatzabschnitt, in dem die Einsatzkraft tätig wurde), Erleben der Übung, Übungsanlage und Übungsleitung.

Für die **Stabs-Rahmenübung** orientierte sich der Fragebogen am Stabsmodel von Heumüller, Richter und Lechner (2014) und umfasste die sechs Bereiche[28] Führungsprozess, Taskmanagement, Respondermanagement, Informationsmanagement, Ressourcemanagement und Standardprozeduren, die als Cluster von vier bis zwölf Items abgebildet wurden. Jedes Item wurde als Fünf-Punkt-Likert-Skala operationalisiert (trifft vollkommen zu (5) bis trifft überhaupt nicht zu (1)). Zusätzlich wurde den Teilnehmenden die Möglichkeit eingeräumt durch Freitextanmerkungen weitere Punkte zu ergänzen. Am Ende der Übung wurde den Beobachtenden wie auch den Teilnehmenden der gleiche Fragen- und Antwortsatz vorgelegt. Zielsetzung hierbei lag darin, die beiden Sichtebenen auf Basis des gleichen Items-Sets miteinander zu vergleichen. Durch dieses Vorgehen sollte auf Basis von Wahrnehmungsunterschieden diejenigen Bereiche, in denen eine genauere, objektive Betrachtung des Übungsgeschehens notwendig wird, identifiziert werden.

[27] Ein weiterer Teil bestand aus Fragen des Stressempfindens und persönlicher Ressourcen und wurde durch die FU Berlin gesondert ausgewertet.

[28] Für die Definition dieser Bereiche wird auf die genannte Publikation von Heumüller, Richter und Lechner verwiesen.

Der operative Charakter der **Feldübung** erforderte eine andere Ausrichtung der Fragebögen. In den drei Übungsschwerpunkten mussten unterschiedliche Aufgaben absolviert werden. Allerdings sollten in allen Teilszenarien Spontanhelfende eingesetzt werden und bildeten somit auch den Schwerpunkt des Fragebogens. Als Antwortmöglichkeiten wurde die bereits beschriebene Likert-Skala verwendet. Im Vorfeld der Übung wurden Einsatzkräfte im Umgang mit Spontanhelfenden geschult. Um die Auswirkungen dieser Schulungen auf das Übungsergebnis zu ermitteln, enthielt die Befragung ein Item zur Schulungsteilnahme.

Ebenso wurden die im Szenario eingesetzten **Spontanhelfenden** nach Ihren Übungserfahrungen befragt. Die Zielsetzung bestand darin die Perspektive auf die Zusammenarbeit zwischen Spontanhelfenden und Einsatzkräften diametral auszurichten. Die Feldphase und die Auswertung erfolgten analog zu der Befragung der Einsatzkräfte.

Prozesskennzahlen

Zusätzlich zu den Beobachtungen und Befragungen, die eher die Perzeption der Übung durch Beobachtende und Teilnehmende widerspiegelten, wurde ein prozessbasiertes Kennzahlensystem eingesetzt. Einfach zu ermittelnde Messwerte wie Zeitpunkte, Anzahlen oder Bool‘sche Werte bilden so die Grundlage für die Beurteilung von Ergebnisqualität, Prozessgüte sowie Zielerreichung. Das Übungsszenario gab einen nicht-optimalen Material- und Personalfluss vor. Die Basis bildete die Hypothese, dass dies neben fehlenden, unzureichenden oder falschen Informationen zur Beeinträchtigung der Prozessleistung und somit der Ergebnisqualität führt.

Um prozess- und ergebnisbasierte Messwerte innerhalb der **Stabs-Rahmenübung** zu erhalten, wurden die drei Kennzahlen Reaktionszeit[29], Erledigungszeit[30] und Erledigungsgrad[31] im Vorfeld definiert. Diese sollten gezielt durch Faktoren beeinflusst werden. In diesem Fall waren dies fehlende bzw. unzureichende Informationen bzw. Ressourcen. Damit wurde implizit auf die Prozessschritte Informationsaufnahme und -verarbeitung, Maßnahmenplanung und -umsetzung sowie Qualität[32] der Stabsarbeit Bezug genommen. Es bestand die Hypothese, dass die Manipulation des jeweiligen Faktors zu einer Verschlechterung der Performance innerhalb der Übung führt. So würde ein Informationsmangel zu einer verzögerten Abarbeitung und Verringerung der Qualität des Ergebnisses der Einsatzbewältigung führen. Die Erhebung der Daten sollte durch die

[29] Zeitpunkt der unmittelbaren Reaktion des Stabes durch eine Nachfrage oder eine Maßnahme.

[30] Zeitpunkt der kompletten Erfüllung der Erwartungen.

[31] Prozentualer Anteil der Erfüllung der Erwartungen zum Übungsende.

[32] Im Sinne der Qualitätsdefinition der DIN EN ISO 9000: 2015 als Erfüllung von Anforderungen.

Übungsleitung erfolgen, die als „Katastrophenszenario“ die Reaktionen und Aufträge der Stäbe erhalten würde.

Die drei unterschiedlichen Übungsschwerpunkte der **Feldübung** (s. Übungsanlage) machten ein differenziertes Vorgehen notwendig.

Zur Erfassung des Ausgangs (Anzahl der Sandsäcke und Lunchpakete) wurden in den **Szenarien Deich und Bereitstellungsraum** mechanische Zählgeräte verwendet. Die Dokumentation erfolgte in einem fünfminütigen Raster. Die Anzahl der gefüllten Einheiten je Zeitraster bildete die Kennzahl Prozessleistung. Zusätzlich sollten auch die Personenanzahlen im Prozess zu diesen Zeitpunkten sowie besondere Vorkommnisse dokumentiert werden. Um einen direkten Vergleich zwischen beiden Gruppen zu erhalten, wurde angenommen, dass jeweils zwei Gruppen gebildet werden; eine ohne Spontanhelfende und eine mit Spontanhelfende. Die Hypothese bestand darin, dass sich die Prozessleistungen beider Gruppen nach einer gewissen Zeit angleichen.

Für das **Teilszenario MANV** wurde der Qualitätsindikatoren-Ansatz von Sautter et.al. (2014) mit dem ergänzten Indikatorenset von Drews et.al. (2018) verwendet, die auf folgende kritische Erfolgsfaktoren beim MANV Bezug nehmen: Priorisierung der Behandlung (Sichtung), Geschwindigkeit der Behandlung kritischer Patienten und Behandlungsqualität. Referenzzeitpunkt zur Messung der darin definierten Zeitindikatoren war das Eintreffen des ersten Rettungsmittels an der Sperrstelle, etwa ein Kilometer vor der eigentlichen Einsatzstelle auf einem Campingplatz. Jeder Patientendarsteller wurde im Vorfeld mit Übungsidentität, Verletzungsmuster und hinterlegten Behandlungsmaßnahmen erfasst. Die Dokumentation erfolgte mittels Versorgungsbewertungskarte[33]. Diese musste bei Ankunft des Rettungsmittels an einem unweit von der Einsatzstelle etablierten Punkt – dem Übungskrankenhaus – an die Evaluationsleitung übergeben werden. Hierbei wurde als Hypothese angenommen, dass durch die Einbindung von Spontanhelfenden in einer Phase, in der nicht genügend Einsatzkräfte vor Ort sind, die Prozessgeschwindigkeit zunehmen würde. Einsatzkräften würde es dann möglich, sich besser um die schwerverletzten Patienten zu kümmern.

Übungsablauf und Ergebnisse

Am Vorabend der Übung wurden die beteiligten Beobachtenden, die Mitglieder der Übungsleitung und die Datenerfassenden in die Übungsevaluation, die Ziele und die jeweilige Aufgabe eingewiesen. Die gesammelten Daten wurden während der Übung bzw. im Nachgang elektronisch erfasst und ausgewertet. Der folgende Abschnitt be-

[33] Diese steht u.a. als Arbeitshilfe auch hier zur Verfügung: www.manv-uebungen.iao.fraunhofer.de

schreibt den Übungsablauf und dessen Auswirkungen auf die Datenerfassung bzw. Ergebnisse. Hierbei werden ausgewählte Ergebnisse präsentiert, um Zusammenhänge besser verständlich zu machen.

Stabs-Rahmen-Übung

Zu Beginn wurden die beiden Stäbe von der Übungsleitung in die bestehende Lage eingewiesen. Danach erfolgten weitere Lageinformationen in Form von Einspielungen. Die Kommunikation zwischen den beiden Stäben erfolgte selbstständig – wie in der Realität zu erwarten. Insgesamt übten die beiden Stäbe sechs Stunden.

Die **Beobachtungsbögen** wurden von allen Beobachtenden ausgefüllt zurückgegeben. Insgesamt wurde die Struktur der Formblätter eingehalten und Beobachtungen von Interpretation getrennt ausgeführt. Die Auswertung erfolgte in Anlehnung an Dresing und Pehl (2015) durch Kategorisierung in die Bereiche des erwähnten Stabsmodells. Die Beobachtenden berichteten über Fehler in den Standardverfahren, aber auch im Informationsmanagement und im Führungsprozess, die dann Auswirkungen auf die weitere Arbeit im Stab zeigten.

Die Antworten der **Befragung mit standardisierten Fragebögen** wurden getrennt von Einsatzleitung und Krisenstab als eigenständige Datensätze ausgewertet, um etwaige Vergleiche zwischen beiden Stäben zu ermöglichen. Insgesamt nahmen an der Einsatzkräftebefragung des Krisenstabs 13 Personen (n_K=13) und an der Befragung der Einsatzleitung neun Personen (n_E=9) teil. Die Rücklaufquote aus dem Krisenstab liegt deutlich unter der der Einsatzleitung. Während aus der Einsatzleitung nahezu alle Mitglieder geantwortet haben, sind es aus dem Krisenstab etwa 40 Prozent. Vor allem liegen nicht aus allen Stabsbereichen Antworten vor. Aus der Befragung der Beobachtenden liegen nur Daten der Gruppe der Beobachtenden im Krisenstab vor. Die Abfrage soziodemographischer Variablen konnte Erfahrungsunterschiede zwischen beiden Stäben zu Tage fördern: In der Einsatzleitung waren vor allem sehr erfahrene Einsatzkräfte, die neben der Stabserfahrung auf eine entsprechend lange Erfahrung als operative Einsatzkraft zurückblicken konnten. Im Gegensatz hierzu bestand der Krisenstab aus eher Krisenmanagement-unerfahrenen Mitgliedern. Die Selbsteinschätzung der Leistung des unerfahrenen Stabes ist bei der Befragung deutlich höher, als die des erfahrenen Stabs[34].

Die Erhebung von **Prozesskennzahlen** gestaltete sich sehr schwierig. Das Reagieren auf Maßnahmen des Stabes bzw. die Vorbereitung von Einspielungen nahm sehr viel Raum in Anspruch. Die Dokumentation von Rückmelde-, Abarbeitungszeit sowie Erle-

[34] Was wiederum auch in im Falle der Risikoeinschätzung bei finanziellen Entscheidungen zutrifft (Bachmann & Hens 2011.).

digungsgrad erfolgte nur in wenigen Fällen, so dass bei den zugrundeliegenden Messwerten die Streuung zu hoch ist, um Aussagen treffen zu können. Um dennoch Aussagen über die Arbeit im Stab treffen zu können, wurde die Dokumentation und Kommunikation der Stäbe angefordert. Die Analyse dieser Ergebnisse erfolgte anhand von Zeitstempeln im Einsatztagebuch bzw. von gesendeten E-Mails. Aus diesen Daten ließ sich allerdings nicht mehr rekonstruieren, auf welche Einspielung reagiert wurde. Diese Daten mussten daher aus der weiteren Betrachtung ausgeschlossen werden.

Im Rahmen der Triangulation dienten die im Stabsmodel genannten Bereiche Führungsprozess, Koordinationsbereiche, Stabsstruktur und Standardprozeduren als Rahmen für die Integration der verschiedenen Ergebnisse. Dabei wurden die Ergebnisse der standardisierten Befragung und der qualitativen Beobachterdokumentation auf Übereinstimmungen bzw. Divergenzen geprüft. Durch die Verwendung der gleichen Items für die Beobachtenden und Teilnehmenden in der Befragung wurden offensichtliche Diskrepanzen deutlich. Jedoch konnten durch die qualitative Betrachtung individuelle Gründe, wie z. B. das Verhalten einzelner Stabsmitglieder oder die zu eng gewählte Taktung von Arbeits- und Besprechungsphasen als Ursachen ausgemacht werden.

Feldübung

Die Alarmierung der Einsatzkräfte erfolgte, wie in einem realen Einsatz auch, seitens der Leitstelle. Allerdings wurde den Einheiten das Übungsdatum und der Startzeitpunkt bekannt gegeben, so dass diese in den Unterkünften auf die Alarmierung warteten. Um realistische Eintreffzeiten zu erzeugen, wurden Fahrtzeiten im Vorfeld berechnet und die tatsächliche Ansprache der Einheiten erfolgte nach einem Einsatzmittelablaufplan[35]. Die im Vorfeld der Übung angesprochenen Spontanhelfenden wurden an einem zentralen Ort gesammelt und von dort aus auf die Einsatzorte verteilt.

Im Teilszenario Deichverteidigung musste die Übungsleitung sehr stark in das Übungsgeschehen eingreifen, um das Übungsziel zu erreichen. Eine der zwei übenden Gruppen bestand aus einer gemischten Gruppe aus Spontanhelfenden und Einsatzkräften, die andere bestand ausschließlich aus Einsatzkräften. Sie wurden erst nach Intervention der Übungsleitung mit der Aufgabe betraut, Sandsäcke zu füllen. Im Teilszenario Bereitstellungsraum konnte ein solcher Eingriff nicht vorgenommen werden, da Unklarheiten über den eigentlichen Einsatzauftrag seitens der Einsatzkräfte bestanden. Durch die Übungsleitung wurde das vorrangige Ziel, nämlich Lunchpakete für die eingesetzten Kräfte in adäquater Zeit an die Einsatzstellen zu liefern, schließlich priorisiert. Entgegen der Planung zur Etablierung zweier Pack-Linien, wurde nur eine Linie durch Einsatzkräfte aufgebaut und dann direkt mit den anwesenden Spontanhelfenden verstärkt. Die Einweisung in die Evaluation der Verletztendarstellenden für das Teilszenario MANV

[35] Für eine Erläuterung siehe DRK 2016.

der Feldübung konnten aufgrund organisatorischer Gegebenheiten erst am Übungsmorgen erfolgen. Das Teilszenario MANV bot durch die Verzögerungen in der Alarmierung ebenfalls die Herausforderung, dass Spontanhelfende bereits auf die Einsatzkräfte warteten. Diese wurden daher sehr früh in das Szenario integriert. Der Leitende Notarzt, die zentrale Führungskraft der MANV-Einsatzbewältigung, kam entgegen der im Kreis erwarteten Realität sehr früh im Szenario zum Einsatz. Dieser musste für reale Einsätze bereitstehen und fuhr daher nicht über die Sperrstelle, sondern auf direktem Weg zur Einsatzstelle. Als weitere Verzögerung musste zusätzlicher Verkehr durch Nachbarn hingenommen werden.

Die **Beobachtenden** wurden vor Übungsbeginn in den einzelnen Teilszenarien platziert. Diese übernahmen dann gemeinsam mit der Übungsleitung vor Ort auch gewisse Steuerungsaufgaben, z. B. die Betreuung der anwesenden Spontanhelfenden. Spontanhelfende wurden insgesamt zögerlich angesprochen. Sobald diese eingesetzt wurden, war der Umgang respektvoll, hilfsbereit und zielgerichtet im Hinblick auf die Einsatzaufgabe. In Bezug auf die Einsatzbewältigung, also taktische Herangehensweise und Prozesstreue wurden im Szenario Deichverteidigung die unterschiedliche personelle Ausstattung der Fülllinien genannt und der gleiche Ablagepunkt für gefüllte Sandsäcke beider Linien. Hierdurch entstanden Verzögerungen und Wartezeiten auf den vorgelagerten Prozessschritten. Im Szenario MANV berichteten die Beobachtenden, dass Patienten mehrfach gesichtet wurden, da diese nicht – wie vorgesehen – direkt mit einer Priorität gekennzeichnet wurden. An der Sperrstelle wurde beobachtet, dass das erste Rettungsmittel sich zunächst auf die dort auf sie wartenden Patientin konzentrierte. Aus dem Teilszenario Bereitstellungsraum berichteten die Beobachtenden, dass der Einsatzauftrag durch die Einsatzkräfte anders interpretiert wurde. Diese gingen von einer kurzfristigen Bereitstellung aus.

Für die **Befragung** mit standardisierten Fragebögen wurden Einsatzkräfte und Spontanhelfende im Nachgang der Übung namentlich angeschrieben und um Teilnahme der Befragung gebeten. Eine Woche später erfolgte eine Erinnerung. Insgesamt war die Umfrage 18 Tage im Feld. Die Auswertung erfolgte im Anschluss getrennt für jedes Teilszenario. Insgesamt nahmen 25 Einsatzkräfte teil, wovon ein Datensatz aufgrund fehlender Dateneingaben ausgeschlossen wurde (n_E=24). Die Rücklaufquote beträgt zehn Prozent. Bezogen auf das Szenario konnten für das Teilszenario Deichverteidigung 6, für MANV 17 und für das Teilszenario Bereitstellungsraum ein Datensatz verzeichnet werden. An der Befragung der Spontanhelfenden nahmen acht Personen (n_S=8; Rücklaufquote 33 %) teil. Diese verteilten sich auf zu jeweils drei Datensätzen im Teilszenario Bereitstellungsraum und MANV sowie ein Datensatz im Szenario Deichverteidigung. Ein Teilnehmender gab hier „keine Antwort“ an.

In Bezug auf die Kommunikation gaben beide Gruppen, Einsatzkräfte wie Spontanhelfende an, dass Sie die Kommunikation als zielführend empfanden. Die Zustimmungswerte zur Verständlichkeit der Anweisungen und zur Kommunikation sind in beiden Gruppen hoch [36].

Die Auswertung der **Prozesskennzahlen** erfolgte automatisch in einer entsprechend vorbereiteten Tabellenkalkulation. Aus dem Teilszenario Deichverteidigung lagen Werte für zwei Fülllinien vor. Aus dem Teilszenario Bereitstellungsraum nur von einer. Im Teilszenario MANV lagen nach Übungsende alle Versorgungsbewertungskarten in ausreichender Qualität vor.

Die Zusammenführung im Rahmen der **Triangulation** bot folgendes – die Übung insgesamt beschreibendes – Bild: Die Rückmeldungen der Beobachtenden aus dem Szenario Deichverteidigung zeigten, dass sich die beiden Gruppen unterschiedlich groß waren und sich bei der Verladung der Sandsäcke gegenseitig[37] beeinflussten. Dies hatte zur Folge, dass zur Berechnung der Prozessleistung die Personenanzahl berücksichtigt wurde. Die Kennzahl Prozessleistung wurde so angepasst, dass die Leistung in Anzahl pro Person und Zeiteinheit definiert wurde. Der Effekt der gegenseitigen Beeinflussung konnte rechnerisch nicht ermittelt werden, aus diesem Grund sind diese Kennzahlen niedriger gewichtet. Für die Fragestellung, ob die Angaben im Tätigkeitenkatalog zutreffen, können sie nicht herangezogen werden. Im Teilszenario MANV zeigten Beobachtungen, dass sich die gemessenen langen Zeiträume in der Patientenversorgung vor allem durch die Sperrstelle und die fehlende Dokumentation durch die Einsatzkräfte ergaben. Der Ablauf im Schwerpunkt des Bereitstellungsraums war durch eine durch die Beobachtung festgestellte Fehlinterpretation des Einsatzauftrages sehr stark beeinträchtigt. Dadurch entstanden deutliche Verzögerungen im Ablauf. Die geplante getrennte Erhebung der Leistung beim Packen von Lunchpaketen mit und ohne Spontanhelfende konnte deshalb nicht gemessen werden. Die Einbindung von Spontanhelfenden als übergeordnetes Ziel wurde seitens der Beobachter und auch die Befragung von Einsatzkräften und Spontanhelfenden positiv empfunden.

Diskussion

In den REBEKA-Übungen wurde ein triangulativer Methodenmix zur Übungsevaluation eingesetzt. Dieser hatte zum Ziel, möglichst viele Perspektiven auf das Übungsgeschehen zu erhalten. Nachfolgend soll ein kritischer Blick auf dieses Vorgehen geworfen werden.

[36] Verständlichkeit $x\square_S = 4{,}43$ und $x\square_E = 4{,}86$; Kommunikation $x\square_S = 4{,}5$ und $x\square_E = 4{,}13$

[37] Es entstanden Wartezeiten beim Verpacken der Sandsäcke auf Paletten.

In beiden Übungen wurde deutlich, dass der freie Verlauf, die Realitätsnähe für Einsatzkräfte und auch Übungsdarstellende eine große Rolle spielen muss, um Einsatzkräften zu ermöglichen, auf ihre individuelle Art das Übungsziel zu erreichen. Die Evaluation sollte diesen Freiraum möglichst nicht einengen oder gar unterbinden. Auf der anderen Seite sind aus Sicht der Methoden gewisse Rahmenbedingungen zwingend erforderlich um überhaupt Messwerte und damit Ergebnisse zu produzieren, die wiederum erst eine Bewertung der Zielerreichung ermöglichen. In Tabelle 13 werden die gewählten Methoden mit den jeweiligen Voraussetzungen aufgelistet. Die eingesetzten Methoden lieferten vor allem dann auswertbare Daten, wenn sie den Übungsteilnehmenden maximalen Freiraum ließen. Vor allem die Prozesskennzahlen aus dem Teilszenario Deichverteidigung und Betreuungsplatz sowie die Prozesskennzahlen der Stabs-Rahmen-Übung sind hierbei zu erwähnen. Diese Kennzahlen erforderten eigentlich ein experimentelles Setting (vgl. Döring & Bortz 2016, S. 193), in dem die Einflussfaktoren kontrolliert werden. Dies hätte in beiden Übungen allerdings erfordert, Spontanhelfende und Einsatzkräfte zum Evaluationsschema und den dadurch eingeschränkten Möglichkeiten der Einsatzbewältigung vorher einzuweisen. Hier hätte also in der Übungsanlage der freie Verlauf eingeschränkt werden müssen.

Dies ist aber aus praktischer Sicht kritisch zu bewerten: Sobald die Evaluation zu sehr in den Einsatzverlauf eingreift, kann die Übung als nicht realistisch wahrgenommen werden. Für Einsatzkräfte, denen nur Übungen zur Übertragung des Erlernten in die Praxis bleiben, ist eine realistische Umgebung und Ablauf aber essentiell. Wenn Patientendarstellende z. B. keine Veränderungen auf getroffene Maßnahmen zeigen, irritiert das die übenden Einsatzkräfte. Werden taktische oder operative Entscheidungen durch die Übungsleitung widerrufen, stört dies den Einsatzablauf erheblich.

Gleiches gilt auch für die Befragung der Teilnehmenden und der Beobachtenden im Nachgang einer Übung. Grundsätzlich scheint dieses Instrument sinnvoll. Mit der standardisierten Befragung konnten die Evaluationsfragen gut operationalisiert werden. Allerdings ist in beiden Übungen die Beteiligung nicht zufriedenstellend. Es stellt sich die Frage nach den Gründen: Die ehrenamtlichen Helferinnen und Helfer sind nicht hauptberuflich Einsatzkräfte, sondern nur in ihrer Freizeit verfügbar. Damit könnte der Faktor Zeit eine Rolle spielen. Ebenfalls möglich wäre eine geringe Wertschätzung dieses Instruments für die Evaluation bzw. deren Nutzen aus der Evaluation.

Tabelle 13: Gewählte Methoden für die Triangulation in den REBEKA-Übungen, deren Voraussetzungen und Möglichkeit in der Triangulation herangezogen zu werden.

Methode	Voraussetzungen				Verwertbares Ergebnis aus den Übungen vorhanden?
	Aktivitäten in Übungsanlage	**Anspruch an Evaluationspersonal im Feld**	**Einschränkung des freien Verlaufs notwendig**	**Reife Evaluationskonzept**	
Qualitative Betrachtungsebene					
Übungs-beobachtung mit erfahrenen Einsatzkräften	a) Schulung der Beobachtenden b) Strukturierte bzw. mit Leitfragen versehene Erfassungsbögen c) Konzentration auf einen Aspekt der Übung oder auf ein Übungsziel	Gering	Nein	Hoch, wird in der Praxis angewendet	Ja
Quantitative Betrachtungsebene					
Befragung mit Fragebogen	Teilnahme-bereitschaft	Entfällt	Nein, wenn im Nachgang der Übung	Hoch	Teils/teils
Prozesskennzahlen					
Feldübung MANV	a) Schulung der Übungs-darstellenden b) Klare Prozess- und Leistungsdefinition	Hoch, tw. gleichzeitiges Erfassen und Spielen der Verletzungen	Geringfügig, ggf. viele Evaluations-personen im Feld	Hoch, wird in der Praxis angewendet	Ja
Feldübung Deich und Bereitstellungsraum	a) Sequentielle Abarbeitung (ohne und mit Spontanhelfenden) b) Eintreffen der Einsatzkräfte vor den Spontanhelfenden c) Einweisung in Evaluationskonzept und Vorgehen.	Gering	Hoch, vordefinierter Ablauf muss eingehalten werden	Gering	Wenige, nicht aussagekräftig
Stabs-Rahmen-Übung	Schulung der Übungsleitung	Hoch, Übungsleitung muss dokumentieren und leiten. Ggf. muss ad hoc reagiert werden.	Teilweise, Einspielungen sollten abgearbeitet werden. Zuordnung zur Einlage muss gegeben sein.	Gering	Wenige, keine Aussagekraft

Das Zusammenspiel zwischen Evaluation und Übungsanlage ist eine Forderung beispielsweise der ISO 22398: 2013. Die Formulierung von Übungszielen und immanent den Evaluationszielen ist zentral. Die Übungsanlage muss diesen Fragen folgen und möglichst viele gedachte Übungsverläufe planerisch antizipieren. Die Komplexität des

Szenarios und dessen Abarbeitung darf dabei nur den großen Bogen spannen. Die beschriebenen Übungen zeigen, dass entweder die Übungsziele und die damit verknüpften Evaluationsfragen zu komplex entworfen wurde oder das Evaluationsschema nicht ausreichend Einzug in die Planung der Übungsanalage gehalten hat. Evaluation muss Komplexität reduzieren, was in den Übungen durch die Triangulation auch größtenteils gelang. Evaluationsstrukturen müssen bei der Übungsanlage bereits mit entworfen und auch – wenn nur gedanklich – getestet werden. Reduktion von Komplexität bedeutet auch Übungsziele einzuschränken. Die Einschränkung der Übungsziele – ebenso wie des freien Übungsverlaufs – muss jedoch verhältnismäßig sein.

In den zurückliegenden REBEKA-Übungen waren zwar Übungsziele formuliert, allerdings waren diese nicht eindeutig und mit einer einzigen Methode zu erheben. Die Überschneidung und der gegenseitige Einfluss beispielsweise verschiedener Einlagen in der Stabs-Rahmenübung oder die Sandsackfüllstation im Teilszenario Deich machen deutlich, dass überschneidende Übungsziele bzw. davon abhängige Evaluationsfragen und –methoden eine exakte unabhängige Auswertung unmöglich machen. In diesem Zusammenhang spielt auch die Fokussierung der verschiedenen Übungsbeteiligten (Patientendarstellende, Übungsleitung, professionelle und spontane Einsatzkräfte sowie Evaluierende) auf die eigene spezifische Aufgabenstellung eine Rolle.

Fazit

Im Rahmen der Evaluation der beiden Übungen zeigte sich, dass durch die Kombination verschiedener Methoden Ausfälle einzelner Methoden kompensiert werden konnten. Einzeln betrachtet bot keiner der gewählten Methoden einen umfassenden Blick auf das Übungsgeschehen. In diesem Sinne bietet die Triangulation aus den Erfahrungen der REBEKA-Übungen einen geeigneten Lösungsansatz zu Übungsevaluation im Bevölkerungsschutz. Sie birgt jedoch die Gefahr der Intermethodenabhängigkeit und das Risiko gesteigerter Planungskomplexität. Durch Triangulation kann der Forderung aus der Praxis nach einem freien Verlauf einer Übung Rechnung getragen werden. Gleichzeitig erfordert sie aber entsprechende Evaluationsressourcen sowie ausreichende Vorbereitung. Für die Praxis sollten geeignete Methodensets für verschiedene Übungsvarianten und Ziele zur Verfügung stehen. Hier besteht noch weiterer Forschungs- und Entwicklungsbedarf.

Literaturverzeichnis

AK V der Innenministerkonferenz. *Feuerwehr Dienstvorschrift 100 - Führung und Leitung im Einsatz* (idF v. 1999).

Bachmann, K. , Hens, T. (2011). *Risikowahrnehmung bei finanziellen Entscheidungen in der Schweiz.* Zürich, Universität Zürich, Institut für Banking und Finance. Studie: 2011. Verfügbar unter https://www.zora.uzh.ch/id/eprint/95105/1/Bachmann_Hens_2011_Risikowahrnehmung.pdf [28.01.2019].

(BMI) Bundesministerium des Innern (2017). *Unterstützung des Ehrenamts im Bevölkerungsschutz.* Verfügbar unter https://www.bmi.bund.de/DE/Themen/Bevoelkerungsschutz/Ehrenamt/ehrenamt_node.html [22.05.2017].

Brauner, F., Stiehl, M. ; Lechleuthner, A., Mudimu, O. A.(2014). Evaluation von Übungen des Massenanfalls von Verletzten (MANV). In *Notfall + Rettungsmedizin (2014), 17 (2),* S. 147–152.

Djalali, A., Carenzo, L., Ragazzoni, L., Azzaretto, M., Petrino, R., Della Corte, F., Ingrassia, P. L. (2014). Does Hospital Disaster Preparedness Predict Response Performance During a Full-scale Exercise? A Pilot Study. In *Prehospital and disaster medicine (2014), 29 (5),* S. 441–447.

Döring, N., Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften.* Berlin, Heidelberg : Springer Berlin Heidelberg.

Dresing, T., Pehl, T. (2015). *Praxisbuch Interview, Transkription & Analyse : Anleitungen und Regelsysteme für qualitativ Forschende.* 6. Aufl. Marburg,.

Drews, P., Sautter, J., Ofterdinger, M., Rieger, A. (2018). Using Quality Indicators to Evaluate MCI Exercises: Experiences from an Airport Exercise. ePoster-Presentation. In Polyakov, O.; Grotto, I. (Hrsg.) (2018). *iPRed V.*

(DRK) Deutsches Rotes Kreuz e.V. - Generalsekretariat; Sautter, J. (Mitarb.), Schneider, F. (Mitarb.), Kippnich, U. (Mitarb.), Wirth, S. (Mitarb.), Habermann, M. (Mitarb.), Siebel, C. (Mitarb.), Kippnich, M. (Mitarb.), Frings, S. (Mitarb.), Häusler, H. (Mitarb.), Max, M. (Mitarb.) (2016). *Durchführung und Auswertung von MANV-Übungen.* Berlin.

Heumüller, E., Richter, S., Lechner, U. (2014). Ein konzeptionelles Stabsmodell für die Entwicklung von Stabsrahmenübungen. In Kundisch, D., Suhl, L., Beckmann, L. (Hrsg.)(2014). *MKWI 2014 - Multikonferenz Wirtschaftsinformatik : 26. - 28. Februar 2014 in Paderborn.* Paderborn: Univ.

Karsten (2016), A.. *Evaluation von Stabsübungen : Grundlagen zur Verbesserung der Führungsfähigkeit.* Verfügbar unter http://jemps.de/evaluation-von-stabsuebungen/. [02.10.2018].

Mayring, P.(2001). Combination and Integration of Qualitative and Quantitative Analysis. In *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*

(2001), 2(1) Verfügbar unter http://www.qualitative-research.net/index.php/fqs/article/download/967/2111.

McCreight, R. (2011). *An introduction to emergency exercise design and evaluation.* Lanham, Md : Government Institutes.

Mitschke, T., Karutz, H. (2017). Aus-, Fort- und Weiterbildung im Bevölkerungsschutz. In Karutz, H., Geier, W., Mitschke, T. (Hrsg.) (2017). *Bevölkerungsschutz : Notfallvorsorge und Krisenmanagement in Theorie und Praxis* (S. 153–166). Berlin Heidelberg: Springer

Perry, R. W. (2004). Disaster Exercise Outcomes for Professional Emergency Personnel and Citizen Volunteers. In *Journal of Contingencies and Crisis Management (2004), 12(2),* S. 64–75.

DIN EN ISO 9000: 2015. *Qualitätsmanagementsysteme –Grundlagen und Begriffe (ISO 9000:2015); Deutsche und Englische Fassung EN ISO 9000:2015.*

Sautter, J., Habermann, M., Frings, S., Schneider, F., Schneider, B., Bracker, H. (2014). Übungsunterstützung für Einsatztrainings des Massenanfalls von Verletzten (MANV). In: Plödereder, E. (Hrsg.) (2014): *Informatik 2014 - Big Data - Komplexität meistern; Tagung vom 22.-26. September 2014 in Stuttgart.* Bonn: Ges. für Informatik, (GI-Edition : Proceedings, 232).

ISO 22398: 2013. *Societal security - Guidelines for Excercises.*