1-1-2011

Drills and Exercises as Interventions to Improve Public Health Emergency Response

Donna Beth Knutson
Walden University

Walden University

College of Social and Behavioral Sciences

This is to certify that the doctoral dissertation by

Donna Knutson

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. James Mosko, Committee Chairperson,
Public Policy and Administration Faculty

Dr. George Larkin, Committee Member,
Public Policy and Administration Faculty

Dr. Tanya Settles, University Reviewer,
Public Policy and Administration Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2013

Abstract

Drills and Exercises as Interventions to Improve Public Health Emergency Response

by

Donna Beth Knutson

MSEd, University of Akron, 1985

BSPE, Ohio University, 1983

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Policy and Administration

Walden University

March 2013

Abstract

The 2001 destruction of the World Trade Center and the subsequent anthrax attacks highlighted the inability of an antiquated public health system in the United States to respond effectively to emergencies. Little documentation exists to define how public health agencies can improve performance. The overarching research question was the extent to which drills and exercises improve performance in public health emergencies. Adult learning theory and deliberate practice theory were explored in this context. The research data were from 50 state public health departments, which were required to report performance information to the U.S. Centers for Disease Control and Prevention. The data were examined using Poisson analysis and logistic regression. Results indicated that drills and exercises had no statistically significant impact on public health performance for the 3 performance measures examined; of all predictors, what explained the most variance in reaching performance targets was the number of real emergencies to which a health department had responded in the past. Performing drills and exercises did not predict the likelihood of reaching performance targets. These findings have implications for positive social change for Congressional leaders and other government representatives. Such public servants could use this information to guide their efforts to redirect public health emergency preparedness funds away from drills and exercises and toward other fundamental public health activities. These more focused efforts could facilitate the improvement of public health laboratory capacity, the training of field epidemiologists, and the advancements in technology for enhanced reporting and surveillance.

Drills and Exercises as Interventions to Improve Public Health Emergency Response

by

Donna Beth Knutson

MSEd, University of Akron, 1985

BSPE, Ohio University, 1983

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Policy and Administration

Walden University

March 2013

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Introduction

Whether by removing the handle from the village pump to stop a community cholera outbreak or by passing out condoms in gay bathhouses to reduce the spread of HIV, public health workers respond to emergencies (McLeod, 2000). The ability of the public health community to identify and contain outbreaks as quickly as possible can be the difference between life and death (Perkins, Popovic, & Yeskey, 2002).

The use of biological weapons for terrorism exists. In the period from 1960 to 1999, 415 cases were documented (Tucker, 1999). After the attacks on September 11, 2001, the probability of a biological threat to public health in the United States increased (Kerr, 2008). Military and government intelligence agencies warned public health leaders of the possibility that nations might harm the citizens of the United States (Frist, 2002). The information indicated the mass dissemination of biological agents such as smallpox (variola major), anthrax (*B. anthracis*), or tularemia (*F. tularensis*) was not out of the question. Release of these agents in a population can cause chaos, death, and panic as acts of biological terrorism (Foster, 2003).

The 2001 destruction of the World Trade Center highlighted the inability of an antiquated public health system in the United States to respond quickly (Gursky, Inglesby, & O'Toole, 2003). The subsequent anthrax attacks using the United States Postal Service as a delivery method for spores resulted in five deaths and significant psychological impact on the nation (Day, 2003). Realizing the danger that delays in detecting, reporting, and isolating biologic threats posed for citizens, the public health

system began a transformation, the specific goal of which was to become more proficient at responding to emergencies (Harvey & Lister, 2004).

In addition to the traditional public health sciences of epidemiology, laboratory service, and surveillance, the anthrax attacks of 2001–2002 pointed to the need for a more robust early detection and reporting system (Chretien, Tomich, Gaydos, & Kelley, 2009). In 2001 the U.S. government provided $1 billion to initiate a transformation of the federal, state, and local public health systems (Frist, 2002). The investment was made to better position them to detect and respond to the mass dissemination of biological agents as quickly as possible (George H. Avery & Zabriskie-Timmerman, 2009). New technologies, such as special cabinets that analyzed air to search for the deoxyribonucleic acid (DNA) of specific biological agents resulted from such public health investments (U.S. General Accounting Office, 2002). Investments in human resources began with intense training and education regarding the most likely agents to be used as biological weapons of terrorism (Fitch, Raber, & Imbro, 2003).

One technique adopted by the public health workforce was the use of drills and exercises to simulate an emergency public health response. Drills and exercises were thought to be steps required for a timely and complete response to protect the largest number of people possible (Dausey, Buehler, & Lurie, 2007; High et al., 2008). Although drills and exercises are common in other emergency response fields, such as emergency medical response and firefighting (Perry, 2004), the use of drills and exercises in public health emergencies is a relatively new phenomenon. The military often uses drills and exercises as a way to look at the possible options enemies have in front of them and to

plan counter actions (Crichton, Flin, & Rattray, 2000). The public health community modified the idea of exercises and drills, using them as an opportunity.

In 2006 the U.S. Congress passed the Pandemic and All-Hazards Preparedness Act (PAHPA). The act mandated drills and exercises as activities required of public health departments receiving funds from the Department of Health and Human Services (DHHS; ("The Pandemic and All-Hazards Preparedness Act," 2006). Congress assumed drills and exercises improved performance in public health emergencies and assisted in determining the level of preparedness (Lurie, Wasserman, & Nelson, 2006).

Drills and exercises became a requirement for public health preparedness programs funded by the Centers for Disease Control and Prevention (CDC). At the time of the requirement, there was little evidence that the completion of such drills and exercises actually improved performance of public health agencies during emergencies. Experts in the field believed the exercises were not evaluative in nature (Biddinger, Savoia, Massin-Short, Preston, & Stoto, 2010). A growing body of literature in the medical domain suggested that using simulation improved clinical care (Lamb, 2007). Documentation is rare that drills and exercises improve performance of public health services. The United States invested millions of dollars to execute drills and exercises without analyzing performance measures before and after such actions to confirm changes in function).

Statement of the Problem

Since 2004, the CDC has awarded state health departments close to $6 billion in federal funding for public health emergency preparedness (Office of Public Health Preparedness and Response, 2011). The cost of executing drills and exercises is a subset

of the total funding, and guidance and the requirement to perform drills and exercises have been included in the instructions to grantees since 2004. There are no methods to document the actual amount spent on drills and exercises and no research indicating they improved performance of public health systems in emergencies.

The Trust for America's Health released its eighth annual report on states' abilities to protect the public's health during disasters, terrorism, and naturally occurring disease outbreaks. After billions in funding to increase laboratory capacity and performance, nearly half of states could not demonstrate the ability to identify *E. coli* O157:H7 and report results to the federal government within 4 days (J. Levi, Vinter, Segal, & St. Laurent, 2010). After 7 years of funding, the nation does not know if drills and exercises contribute to changes in performance.

Nature of the Study

The role of exercises and drills in the performance of public health preparedness systems during a public health emergency was the focus of this quantitative study. The research questions addressed the efficiency of timed performance of tasks undertaken by public health departments that have participated in drills or exercises and those that have not. The data analysis plan included descriptive data for all variables collected, the results of the timed measures, and categorical data collected about the entity from reports. The dependent (outcome) variable was change in the unit of time recorded for the specific task during a real event. The change could be measured in real units (hours, minutes, or seconds) or could be measured as increases, decreases, or no change. The independent (predictor) variable was if there was practice or an intervention, the type of practice (e.g., drill or exercise), and the number of practice sessions prior to timed measurement of a

real event. This study was an attempt to explain or predict relationships between the variables (Radhakrishna, Yoder, & Ewing, 2007). Any of the variables could have a strong or weak predictive value on the outcome, with the outcome of interest being what produces a change in the timed response.

Research Question and Hypotheses

1. Is participation in drills and exercises a factor for improving times for staff notification, staff assembly, and completion of an after-action report?

$H_0$1: The amount of time needed to complete staff notification, staff assembly, and an after-action report during a real event is not different among public health departments that have participated in drills and exercises and public health departments that have not participated in drills and exercises.

H_a1: The amount of time needed to complete staff notification, staff assembly, and an after-action report during a real event is different among public health departments that have participated in drills and exercises and public health departments that have not participated in drills and exercises.

2. Is the periodicity of participation in drills and exercises a factor for improving times for staff notification, staff assembly, and completion of an after-action report?

$H_0$2: The number of drills and exercises completed within 6 months does not explain differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

H_a2: The number of drills and exercises completed within 6 months explains the differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

3. Is the length of time between participation in drills and exercises a factor for improving times for staff notification, staff assembly, and completion of an after-action report?

$H_0$3: The amount of time between drills and exercises, and real events does not explain differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

H_a3: The amount of time between drills and exercises, and real events explains differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

4. Does participation in drills and exercises increase the number of states that meet the target times for staff notification, staff assembly, and completion an after-action report?

$H_0$4: The number of drills and exercises completed within 12 months does not explain differences in the number of states that meet the targeted times to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

H_a4: The number of drills and exercises completed within 12 months explains differences in the number of states that meet the targeted times to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

Purpose of the Study

The purpose of this quasi-experimental study was to determine if drills and exercises improve performance of activities common to public health emergencies. The

research design controlled for factors such as the size of the jurisdiction served, financial support, and structure of state public health departments. The outcome of interest, the dependent variable, was the amount of time it took to complete a generalized task common to all public health responses. The independent, or predictor, variable was the completion of drills, exercises, or real events. Statistical analysis provided information on factors that influence the outcome of interest, such as the time between drills and real events, the number of drills or real events, the size of the health department staff, financial support for preparedness, and the structure of the state preparedness program.

Theoretical Framework

The idea of performance measurement for public health programs has been codified in several programs developed by the executive branch of government, such as the Performance Assessment and Rating Tool (PART) or the Government Performance and Results Act (GPRA; (DeGroff, Schooley, Chapel, & Poister, 2010). However, there has been little literature published on the theories behind performance measurement (DeGroff et al., 2010). In the absence of performance measurement theory, individual learning theories have served as a basis for explaining improved performance. ELT and deliberate practice theory relate to gaining experience through learning or practice. The intersection is important to this research. Some think the best way to test changes in performance is to simulate an emergency through drills or exercises (Brandeau, McCoy, Hupert, Holty, & Bravata, 2009; Gebbie, Valas, Merrill, & Morse, 2006; Perry, 2004). Observing the response is a method for gauging improvement in the performance of agencies. It also allows for the testing of preparedness plans (Beaton et al., 2004; Fowkes,

Blossom, Sandrock, Mitchell, & Brandstein, 2010). Each of the theories is relatively new, developed within the last 40 years.

ELT is based on work by scholars such as Dewey, Lewin, Piaget, James, Jung, Friere, and Rogers (A. Y. Kolb & Kolb, 2005). Kolb introduced ELT in 1984. According to ELT, learning occurs in a cycle based on four different phases of learning: concrete experiences, reflective observation, abstract conceptualization, and active experimentation (Kayes, Kayes, & Kolb, 2005). The learner enters the cycle at any one of the points and continues through the cycle as learning occurs. The model suggests that learners face polar-opposite choices, and concrete experiences rarely allow for abstract conceptualization at the same time, as evidenced by individuals who try to walk while reading a book. The concrete experience of walking will not necessarily allow for the conceptualization of the ideas presented in the book.

ELT is also helpful in determining learning styles. Life experiences, exposing the learner to new environmental situations, and building the capacity to learn from one's learning style (D. A. Kolb, Boyatzis, & Mainemelis, 2000). For instance, the accommodating learning style describes people who prefer concrete experiences and active experimentation. The assimilating learning style describes those who prefer reflective observation and abstract conceptualization (Baker, Jensen, & Kolb, 2005).

Kolb (1984) based his work on ideas formulated by Dewey in the early 1900s. Dewey (1938) theorized that the outcome of learning was to obtain freedom of thought. He argued the best way to obtain freedom of thought was through experience that had both continuity and interaction (Dewey, 1938). Because the theory explored the relationship between phases of learning through experience, Kolb's ideas had an effect on

educators. Educators began to think about teaching learners through concrete experiences followed by reflection. Reflection allows the learner to develop new concepts and abstract ideas, and leads to applying the new ideas in different situations (Lisko & O'Dell, 2010). The cycle continues, and it is theorized that this was one way to assure each learner could reach their maximum potential (Boyatzis, Cowen, & Kolb, 1995). Kolb continued to investigate the tenets of the theory with individuals with similar characteristics, and others began to apply the theory to groups (Pistole, Kinyon, & Keith, 2008). Kolb published dozens of articles on ELT and examined the theory in different contexts, using it to predict learning styles of high school students and help create better executives in a management context and in the accounting profession. Kolb, however, did not link the theory to improved, timed performance of tasks.

Deliberate practice theory also addresses gaining mastery in certain skills through experience and explores the methods used in training to achieve extraordinary results. Ericsson et al. (1993) explored the characteristics of training to become an expert performer. Ericsson's theory was also born from work by scientists who came before him and developed theories of skill acquisition, which he analyzed to discover the niche for his work. Ericsson published articles on deliberate practice and explored his theory within many professional and leisure domains, such as with typists, dart throwers, chess players, tennis players, medical professionals, and musicians (Ericsson & Charness, 1994).

The development of deliberate practice theory was a challenge to the notion that expertise developed through a combination of factors such as the length of time one had been practicing the skill and repetition. Ericsson et al. (1993) discovered that the factors

previously thought to influence the potential for mastery held only a weak relationship with observed performance. Factors identified as having a significant role in performance improvement included clear task definition, motivation on the part of the individual to improve, immediate feedback from someone who was trained to recognize the performance needed to be considered an expert, and frequent opportunities for refinement and practice (Ericsson, 2008).

Ericsson (2007) explored using deliberate practice as a way to enhance performance for medical professionals. Law professors have used deliberate practice methods to teach students how to interview, counsel, and negotiate (Farmer & Williams, 2005). The theories of experiential learning and deliberate practice have attracted those interested in adult learning and performance measurement. In the case of the ELT, Kolb (1971) tied the four stages of the learning cycle directly to learning styles. As an individual matures and has more experiences, he or she is able to settle on a learning style that is a combination of two of four phases.

Ericsson (2006) indicated in his research that automaticity hampers performance, and additional experience alone will not improve performance. Gravitating to what feels comfortable or automatic might not allow an individual to overcome performance plateaus and achieve extraordinary results (Ericsson, 2006). Each of the theories might contribute to a change in performance of the system in public health emergencies, but it has been difficult to determine if professionals have considered the theories when developing drills and exercises. There are few references to learning theories in the literature about the construction of drills and exercises. A majority of the literature reviewed regarding theory focused on deliberate practice theory, as it often measures

improvement through actions within a time frame, and it is mentioned as the theory behind drill development (Lee, Trim, Upton, & Upton, 2009).

The literature review includes overviews of the learning theories and an in-depth review of drill and exercise construction and results. The majority of the literature, specifically that regarding exercises and drills, is descriptive or qualitative in nature, and there are few quantitative studies specific to the changes in performance experienced by state health departments. There is little reference to educational theory in the literature describing the purpose or construction of drills and exercises. Research also indicated that team behavioral quality improves after teamwork training through classroom training and practicum activities (Morey et al., 2002).

Definitions

*Public health system: "*The collaborative efforts of a complex network of people and organizations in the public and private sectors, as well as an alignment of policy and practice of governmental public health agencies at the national, state, and local levels" (Medicine, 2002, p. 28).

Public health emergency preparedness: "The capability of the public health and health care systems, communities, and individuals to prevent, protect against, quickly respond to, and recover from health emergencies, particularly those whose scale, timing, or unpredictability threatens to overwhelm routine capabilities" (Nelson, Lurie, Wasserman, Zakowski, & Leuschner, 2008, p. 6).

Drills: A drill is a type of exercise that

> Is a coordinated, supervised activity usually employed to validate a single, specific operation or function in a single agency or organizational entity. Drills

are commonly used to provide training on new equipment, develop or validate new policies or procedures, or practice and maintain current skills. (Department of Homeland Security, 2007b, p. 11)

Exercise: An exercise is an instrument to train for, assess, practice, and improve performance in prevention, protection, response, and recovery capabilities in a risk-free environment. Exercises can be used for testing and validating policies, plans, procedures, training, equipment, and interagency agreements; clarifying and training personnel in roles and responsibilities; improving interagency coordination and communications; identifying gaps in resources; improving individual performance; and identifying opportunities for improvement (Department of Homeland Security, 2007a, p. 1)

Performance measurement: A process of assessing progress toward achieving predetermined goals, including information on the efficiency with which resources are transformed into goods and services (outputs), the quality of those outputs (how well they are delivered to clients and the extent to which clients are satisfied), outcomes (the results of a program activity completed to its intended purpose), and the effectiveness of government operations in terms of their specific contributions to program objectives (Gore, 1997, p. 6)

Public health workforce: The public health workforce, broadly defined, includes all those significantly engaged in work that creates the conditions within which people can be healthy. More specifically, the workforce is composed of those who work for official public health agencies at all levels of government, community-based, and voluntary organizations with a health promotion focus; the public

> health–related staff of hospitals and health care systems; and a range of others in private industry, government, and the voluntary sector. This workforce includes nurses, sanitarians, educators, administrators, physicians, nutritionists, social workers, engineers, and many other professionals; a large group of persons working in the field as aides, extenders, and community health workers; and of course vital administrative support and clerical staff and a remarkable complement of volunteers (Tilson & Gebbie, 2004, p. 341).

All-hazards public health preparedness: The concept of an all-hazards approach to public health and emergency preparedness refers to the functional integration of emergency management activities at all levels of government, with plans designed for a broad range of emergency situations. A hazard is the source or potential source of danger and may be due to natural, social, or technological phenomenon. The all-hazards approach divides emergency management functions into the following four areas: mitigation, preparedness, response, and recovery (Moore, Mawji, Shiell, & Noseworthy, 2007, p. 282).

Limitations of the Study

Inherent in the time series design are some known issues that could affect the results. The first is called practice effects, or intertrial repetition effect, which means the participants might change over time due to repeated testing (Frankfort-Nachmias & Nachmias, 2008, p. 125). The participants can become more proficient because they know the drill versus actually gaining the skill, or they might be experiencing regressing toward the mean. Regression to the mean is a phenomena found in time series studies

where variation in measures that occur naturally is mistaken for actual change (Barnett, van der Pols, & Dobson, 2005).

Internal threats to validity were that the data were self-reported and that no researcher was onsite to validate the information (Podsakoff & Organ, 1986). Another internal threats was the maturation of the field, which could have affected results negatively or positively (Cooper & Hedges, 1994). Technology also played a part in internal validity, particularly around communication issues. Social media and other forms of public communication have proliferated, and health departments may be using the advances in technology differently (E. Avery et al., 2010). In a disaster scenario, traditional forms of communication with the public might be impossible. Using alternative methods, such as Twitter or Facebook, might cause changes in timed responses (Aten et al., 2011).

External validity threats also existed in this study. For instance, the characteristics of the public health community are narrow, and there may not have been enough variation in responses to distinguish an impact of the independent variable. The results of the study cannot be generalized beyond the public health community due to the unique nature of activities that only occur through the governmental public health agency (Reis & Judd, 2000).

There are limitations of the measurement instrument. The instrument is weak in empirical validity because there are no current criteria available to reference when assessing the predictive validity values of drills and exercises for a real event. The idea of drills and exercises as a way to build expertise is not well supported by empirical evidence (Ericsson, 2004). The instrument used for the study was construct referenced,

meaning the organization reporting results did so to improve its individual position against a standard developed by the community of practice, as opposed a reference group (William, 1996). The benchmark for the timed performance was the competition for the participants—not a competition among the participants. Participating health departments compete against their own best score, not the scores of others.

Significance of the Study

Researchers have reported the value of drills and exercises as a method to solidify relationships, build competencies, test plans, or define the roles of responding agencies in emergencies (Ablah, Nickels, Hodle, & Wolfe, 2008; Biddinger et al., 2008; Brand, Kerby, Elledge, Johnson, & Magas, 2006; Dausey et al., 2007; Gebbie et al., 2006; Savoia, Biddinger, et al., 2009; Savoia, Testa, et al., 2009), but few studies have been designed to evaluate performance of the public health system in relation to participation in drills (Biddinger et al., 2010). I looked at the performance of public health departments on select measures before and after participation in drills and exercises to determine if participation drills and exercises could predict improvement in performance. Funds are expended each year at the local, state, and federal levels to conduct drills and exercises without evidence that they improve performance. Findings could influence the decisions of lawmakers as well as those managing public health departments at all levels.

Social Change

Public health as a field has been tied to social change for the last 2 centuries, linking the health of the public to social and economic determinants (Szreter, 2002). In many instances, the role of governmental public health has been to protect the public's health through the creation of policy that puts barriers between people and the things that

harm them and the community, such as requirements for certain immunizations for healthcare workers (Lindley, Horlick, Shefer, Shaw, & Gorji, 2007), or the enactment of tobacco-free policies and increases in excise taxes to prevent exposure to secondhand smoke (Centers for Disease Control and Prevention, 2011). Laws created for the good of the public's health have withstood challenges to the Supreme Court and have impacted society in a positive manner (Gostin, 2005). Policy legislation should be based on evidence derived from a comprehensive review of the literature to determine what works and what does not work, a process often referred to as systematic reviews (Fox, 2005). Without such knowledge, policies could go into effect that do not use government funds efficiently, particularly in this era of declining revenue and budget shortfalls for federal, state, and local governments.

If drills make a difference, this study would contribute to the body of knowledge with which to conduct a systematic review of the practice and would support the use of discretionary funds for this purpose. If not, policy makers could use the information to redirect dollars toward evidence-based policies and practices that would continue to promote social change on a broader scale than just emergency preparedness. For example, access to $6 billion would allow 500,000 to 1,000,000 people per year to have publically funded health insurance (Frick, 2009). The Women, Infants, and Children's program (WIC) had an estimated $6 billion budget for 2009 and managed to feed more than 9 million women, infants, and children each month (Carr, 2010). This study could help policy makers to re-examine the role of drills and exercises as contributors to public health preparedness and make fiscal adjustments if necessary.

Summary

Completing drills and exercises is a requirement for state public health departments receiving federal funds to build preparedness capacity. The purpose of my study was to review scientific evidence captured in literature to determine if drills and exercises change the performance of public health systems in emergencies. I examined ELT and the theory of deliberate practice as accepted methods of adult learning and performance improvement. I examined the role of drills and exercises as activities that might influence performance in public health emergencies. The quasiexperimental study examined timed responses as the outcome of interest and created models to determine if the predictor variable—drills or exercises—influenced the timed response. The chapters that follow review current literature about public health exercises, educational theories, general public health capacities, and performance measurement. I also include a description of the data collection and analysis plan, as well as full data analysis, results, and recommendations.

Chapter 2: Literature Review

Introduction

Deliberate practice theory is a framework through which to understand how learning and improved performance might take place in the context of drills, exercises, and simulation-based training (SBT). The value of SBT is its ability to create a learning environment that mimics situations found in the real world (Salas, Rosen, Held, & Weissmuller, 2009). Guided practice through SBT assures that certain target competencies are acquired by participants (Salas et al., 2009). Deliberate practice theory should underpin the development of SBT activities because it allows for the measurement of performance. Salas et al. (2009) stated that deliberate practice theory is suited for SBT. The competencies desired are describable, practice occurs and measurements are taken in a realistic environment, and the participants are subject to direct feedback. Each of these components is necessary to develop expertise. The following is a review of the current literature describing deliberate practice theory, public health systems and preparedness, exercises, drills, and SBT.

The literature search was conducted using several databases available online through two resources: the Walden University Library (e.g., Thoreau and its associated databases) and the CDC Library, which provides access to over 100 databases (e.g., PubMed, MEDLINE, EBSCOhost, ERSH-DB, Google Scholar, Cochrane Library, and JSTOR). Each of the resources allowed electronic access to several journals, and the CDC library retrieved articles from hard copy if they were not available electronically. Searches were conducted on several key words, including *public health emergency response, emergency response, public health preparedness, preparedness, performance*

measurement, exercises, drills, simulation-based training, disaster, bioterrorism, hospital preparedness, training, deliberate practice, experiential learning, expert practice, and *public health policy.* Articles for review were selected based on their relevance to the topic and if they specifically mentioned the role of drills and exercises in developing expertise. As the literature review showed, there have been more descriptive and qualitative articles published than quantitative articles.

Deliberate Practice Theory

Expert performance was thought to be a combination of innate talent, perception, memory, and problem-solving skills (Fernand, 1998). Early studies were conducted with exceptional chess players in an attempt to gain a clearer understanding of how they remembered moves and gained their expertise (Chase & Simon, 1973). Chase and Simon (1973) found that grandmaster chess players did not arrive at that standing with less than a decade of exposure to the game. An average of 10 years was the amount of time needed to become an expert in domains other than chess.

While data confirmed expertise took a long time to develop, Ericsson (Ericsson et al., 1993) proposed that length of exposure to the activity was not the only factor in building expertise. Ericsson, Krampe, and Tesch-Römer (1993) developed a theoretical framework to explain expert performance as the result of gradual attainment of expertise enhanced with focused practice over a long period, calling the idea deliberate practice theory. The theory expands on findings that indicated that experience alone related weakly to expertise and that innate talent could bring people to a certain point, but moving beyond was difficult without structured, deliberate practice (Ericsson et al., 1993).

Deliberate practice improves current level of performance, overcomes weaknesses, and uses experts to identify ways to improve performance (Plant, Ericsson, Hill, & Asberg, 2005). They are primarily solitary in nature and focused on improvement in a particular domain (Plant et al., 2005). A basic assumption inherent in the theory is that the amount of time spent in activities that meet the definition of deliberate practice will improve performance monotonically (Ericsson et al., 1993). There are three characteristics of the practice itself. First, exposure to the activity at a young age will allow the expertise to build continuously and gradually for 10 years or more. Second, practice is not enjoyable, and the only motivation is to achieve better performance. Third, practice is intense and unsustainable over long periods throughout the day. To achieve expertise, an individual must overcome three constraints: a resource constraint, a motivational constraint, and an effort constraint.

Ericsson et al. (1993) tested the theory in the laboratory by predicting that current performance relates to the amount of past deliberate practice and that deliberate practice starts slowly and builds over time. They also predicted that the highest-level performers had the largest weekly doses of deliberate practice, with each daily practice limited so as not to tire out the participant. The researchers used the diaries of elite musicians to determine that practicing alone was the most important activity for improving performance and overcoming the resource constraint, as there is no need for teachers or other supplies to practice alone. They also confirmed that musicians' practice required effort as evidenced by the frequent napping after practice and that practice was of short duration, no longer than 1 to 1.5 hours per session (Ericsson et al., 1993).

Salas et al. (2009) included deliberate practice as one of the theories that should inform best practices of performance measurement in SBT. The authors stated that measuring and observing expertise during SBT are two different types of performance measurement, yet they can be developed to have an impact on SBT, such as emergency preparedness exercises. The authors stated that the components of the expert behavior must be identified through qualitative evaluation and then brought into SBT through the use of corrective feedback to change the behavior of the developing expert (Salas et al., 2009). In addition, Salas et al. recommended event-based measurement as one of the tools appropriate to measure performance in SBT. The scenarios should be based in theory, with clearly defined outcomes that can be measured in order to improve competency (Salas et al., 2009).

Ericsson (2000) proposed three components identified in a framework for testing the theory of deliberate practice in new domains (Ericsson, 2000). First, the researcher must identify a collection of reproducible tasks that represent the expert performance desired. The second component is mapping and analyzing the process used by the expert to determine what techniques make the performance superior. The last component is explaining how the mechanisms identified were acquired by practice activities (Ericsson, 2000). What follows is the application of deliberate practice theory in several domains, although there is limited direct research relating deliberate practice theory to public health emergency response SBT.

Sports

Hodges, Kerr, Starkes, Weir, and Nananidou (2004) conducted two studies with triathletes and swimmers to understand the theory of deliberate practice and its

relationship to performance. The researchers conducted a third study, designed to replicate Ericsson's (1993) research that indicated many of those who used deliberate practice did not find the activities enjoyable. In the first two studies, Hodges et al. proposed a negative relationship between practice hours and performance, meaning fewer practice hours would lead to slower performance. In addition, they proposed that the number of hours spent in non-sport-specific fitness activities (e.g., swimmers who spent time playing tennis) would have no effect on performance scores. The researchers were also interested in the relationship between accumulated years of practice and most recent practice, and the amount each accounted for variance in performance scores (Hodges et al., 2004).

Using hierarchical regression methods, Hodges et al. (2004) confirmed that sport-specific practice, accumulated and recent, was the best predictor of performance for the triathletes. Time spent in non-sport-specific activities was not a predictor of performance times. The regression model for swimmers indicated that the amount of time spent practicing was a predictor for long endurance events but not for short, sprint-type events. Gender played a role in predicting performance scores in the short events, with males posting faster times than females. The hypothesis that more practice time would result in faster performance scores did not hold true for short-distance swim events, possibly indicating that practice is a factor in only longer, endurance events.

In the third study, Hodges et al. (2004) tested the hypothesis that activities performed by athletes specifically designed to improve their performance are not enjoyable. The researchers used exercise diaries and questionnaires with the athletes and found that individuals rated activities less enjoyable in their diaries than in the

questionnaires. The conclusion by the author was that the perception of enjoyment by the athletes was not useful when describing the execution of activities deemed to fit the definition of deliberate practice.

Hodges et al. (2004) contributed to the deliberate practice body of knowledge by the three studies presented with another glimpse into the role of practice and performance. Hodges et al., however, did not feel the results of the studies confirmed the role of deliberate practice in attaining expert performance. The authors called for more longitudinal research over the lifetime of the performer's career.

Ward, Hodges, Starkes, and Williams (2007) studied elite and subelite soccer players to determine the role of domain and nondomain practice activities in the development of expertise. The study reviewed the practice times and micropractice routines of soccer players from 8 to 18 years of age. Ward et al. tested the hypothesis that the most recent accumulation of team practice was the best predictor of expertise for the group of soccer players under study (Ward et al., 2007).

Results supported the idea that more accumulated practice resulted in higher expertise. Elite players started playing matches at a younger age than subelite players. Team practices in the most recent year were a strong predictor of skill level (Ward et al., 2007). The data suggested that early participation in teams, some early match play, individual practice, and playful soccer- related activities provided the environment for appropriate skill progression from subelite to elite (Ward et al., 2007). The authors concluded that engaging in practice was likely the best predictor when compared with team and match-play participation, as well as playful soccer activities (Ward et al., 2007).

There was no expertise predictor effect of the participant's nonsoccer activities when analyzed in light of subelite or elite status.

De Bruin, Smits, Rikers, and Schmidt (2008) compared the practice habits of current nationally rated chess players against formerly rated players who dropped out of national training. The authors predicted that deliberate practice hours contributed to the national ratings of chess players throughout their careers. The researchers also predicted those players who remained in the national training program would obtain higher ratings due to more deliberate practice. Lastly, they predicted that the player's sex would make no difference in their performance (de Bruin et al., 2008).

The participants noted the number of hours spent on chess study and chess play per week at different ages of their lives (de Bruin et al., 2008). After obtaining national ratings, the researchers developed several regression models to predict the variables most likely to relate to expertise. Serious chess study alone and play with others were strong predictors of success regardless of when they occurred in the career of the chess players. Differences in levels of deliberate practice did not explain the performance differences between those who stayed in the national training program and those who dropped out. Men accumulated more deliberate practice hours than women, and the result was lower performance for women in the study. The authors determined that deliberate practice played an important role early in the life of chess players and did not diminish over time. The results of the study supported the role of deliberate practice as a tool to facilitate the development of expertise in chess.

Military

McKinney and Davis (2003) tested the theory of deliberate practice with U.S. Air Force pilots. In the study, McKinney and Davis examined if the benefits of deliberate practice were extended to activities that were not practiced by the pilots and tried to determine if deliberate practice had different effects on the phases of decision making, such as the assessment phase and the action phase. The Air Force uses flying simulations as an integral part of training. McKinney and Davis structured the research to examine the behavior of pilots in simulated crises to see if the practice helped them when making decisions in a real-life crisis of the same nature.

Pilots who undertook deliberate practice for complete scenarios exhibited positive crisis decision making (McKinney & Davis, (2003). The effectiveness of decision making for the pilots related significantly to the amount of time spent in deliberate practice. However, if the scenario was not fully practiced, deliberate practice made no difference in the decision-making ability of pilots. The deliberate practice activities were thought to increase the potential for good decision making in any scenario but that was not the case (McKinney & Davis, 2003). Further research showed that deliberate practice was a good method to use to prepare decision makers to choose correctly in fully practiced, once-in-a-lifetime crises but showed no impact on decision making in unpracticed scenarios.

Beaubien et al. (2006) used deliberate practice as a method to train U.S. Army National Guard officers. Officers were selected to play the role of support and stability operations, an often complex and ill-defined role that typically is needed after man-made or natural disasters. The army chose a simulation, Red Cape, as the scenario for educating

the officers instead of costly and time-consuming full-mission exercises. The officers provided answers to a series of challenge questions following training on 15 scenarios associated with the simulation. Trainers were on hand to discuss the answers with the officers and to provide immediate individual and group feedback. The researcher provided no further empirical information about the success of the training but did provide lessons learned from executing the approach (Beaubien et al., 2006).

Williams, Ericsson, Ward, and Eccles (2008) examined the theory of deliberate practice and provided an assessment of the practical application of the theory for the military. The summary did not present data from specific studies undertaken by the military but called for additional research to define tasks amenable to deliberate practice in the military domain.

Medicine and Health Professionals

Ericsson (2007) discussed the application of deliberate practice theory to the domain of medicine, seeking clinicians with objectively superior performance based on better diagnostic abilities and better treatment outcomes. With the increased use of simulations, Ericsson proposed that more opportunities for clinicians to practice in sessions that met deliberate practice criteria would likely improve clinician performance. Research has also indicated that increases in practice have been associated with increases in performance (Ericsson, 2007). Ericsson, Whyte, and Ward (2007) also believed the same opportunities for increased performance as predicted with clinicians might be possible in the field of nursing (Ericsson et al., 2007). There was no research presented in the articles; however, Ericsson suggested more research and application of deliberate practice was needed for the medical and health professions.

Academic Performance

Moulaert, Verwijnen, Rikers, and Scherpbier (2004) designed a study to determine if there was a relationship between deliberate practice and the achievements of undergraduate medical students residing in the Netherlands. The authors used a questionnaire probing study habits and personal behaviors to measure aspects of deliberate practice. Activities deemed as deliberate practice included self-study, study resources, planning, study style, and motivation. Moulaert et al. linked achievement test scores and the answers to the questionnaires to analyze correlations between deliberate practice and study achievements. The results indicated students who reported studying more achieved higher test scores. In addition, high achievers owned more books and read more articles than the lower achievers. Lastly, the high achievers scored higher on planning and study-style factors (Moulaert et al., 2004). The researchers concluded the aspects they identified were representative of deliberate practice activities, and those who used deliberate practice activities as part of their study habits contributed to the success of the medical students.

Plant, Ericsson, Hill, and Asberg (2005) explored the role of deliberate practice and academic performance based on factors such as cumulative grade point average (GPA), quality, and quantity of study, and previous performance. Previous researchers indicated there is a weak relationship between amount of time spent studying and academic performance (Schuman, Walsh, Olson, & Etheridge, 1985). Plant et al. studied observable indicators of effective learning by extending the theory of deliberate practice and the theory of self-regulated learning based on several parallels between the two theories. Plant et al. chose cumulative GPA as a marker of academic expertise, as well as

standardized test scores, and asked participants about study habits and time spent studying. The participants completed daily logs recording their study habits and environments. In this case, deliberate practice equated to activities such as solitary studying in an environment free from distractions. The results of the study indicated deliberate practice played a role in higher GPAs.

Occupational Performance

Dunn and Shriner (1999) examined the role of deliberate practice in the field of education, mapping out the activities of teachers that fit the definitions set forth by Ericsson. Dunn and Shriner surveyed teachers who ranked selected activities as relevant to their work, requiring significant effort, and occurring frequently. From the list, the authors summarized activities that could lend themselves to the definition of deliberate practice and collapsed the list into one category of preparation, planning, and evaluation. The authors developed a second study to determine if studying the daily activities relative to formal and informal evaluations of students using daily logs and survey questions could identify expertise in teachers.

Dunn and Shriner (1999) found several factors that made it difficult to portray activities of teachers ready for improvement with deliberate practice. For instance, student improvement, not self-improvement, motivates teachers, which violates a tenet of deliberate practice theory and does not allow removal of the motivation constraint. Immediate feedback from both the students and other professionals is difficult in the classroom. The field of teaching is ill structured, and the novice teachers receive feedback after only a few months of studying under a mentor or coach before stepping

into the classroom on their own. Lastly, the teaching environment might not be conducive to rewarding or recognizing the development of expertise (Dunn & Shriner, 1999).

Sonnentag and Kleine (2000) designed a study to examine whether supervisory performance ratings improved if insurance agents engaged in deliberate practice. To investigate the hypotheses, Sonnentag and Kleine differentiated deliberate practice at work from specific trainings or conferences by defining it using Ericsson's (Ericsson et al., 1993; Ericsson & Lehmann, 1996) principles as a continuous effort to improve competence through regularly performed activities. Sonnentag and Kleine also acknowledged that deliberate practice activities in a work setting might not be as obvious as rehearsing tasks associated with the occupation. The researchers defined deliberate practice as the continuous performance of supporting activities, such as consuming reading material relevant to the task or seeking expert advice to help solve a problem.

Sonnentag and Kleine (2000) interviewed 100 insurance agents and provided them with diaries to chart their work-related activities during the week. The researchers also obtained performance ratings for each of the agents from their supervisors. The data points included current and cumulative time spent in support activities and current and cumulative time spent in activities deemed deliberate practice. The activities determined to fit the deliberate practice model had goal-oriented outcomes as reported by the agents. Of the 10 activities, 1.03 qualified as deliberate practice activities (Sonnentag & Kleine, 2000). Years of experience by the agents were not a significant predictor of performance ratings, but the number of cases handled predicted performance as captured by supervisory ratings. The current amount of time spent in deliberate practice accounted for 6% of the variance in performance. Measures of cumulative supporting activities and

cumulative deliberate practice did not predict performance, although current time spent on deliberate practice was a predictor of supervisory performance ratings. Sonnentag and Kleine concluded deliberate practice is different for participants, and perhaps the key attributes of repeatedly practicing and practicing for improvement versus the actual activity practiced were more important for performance improvement in this domain.

A study of South African small businesses tested a potential relationship between cognitive abilities, domain-specific knowledge, and deliberate practice on business success. Business owners from Cape Town participated in interviews and completed a survey assessing their practical business knowledge. Unger, Keith, Hilling, Gielnik, and Frese (2009) used the work of Sonnentag and Kleine (2000), using six activities identified by them as deliberate practice activities for insurance agents. They also added four new activities based on structured interviews, providing 10 deliberate practice activities. The participants indicated whether they performed the activities at all, how frequently, and if the goal was to improve competence, skills, or knowledge. To complete the data set, the authors provided a measure of the quality of the deliberate practice based on evidence of learning (Unger et al., 2009). After analysis, deliberate practice showed a strong positive effect on entrepreneurial knowledge. Education showed a significant positive relationship to deliberate practice, and cognitive ability showed a positive relationship with deliberate practice. In addition, deliberate practice was a mediator between entrepreneurial knowledge and cognitive ability, explaining why there is a relationship between the two variables.

Canine Activities

Helton (2007) took a different approach to testing the relationship between deliberate practice and performance that removed any confounding notion of early life experiences and talent by testing the theory using dogs that practice the sport of agility. Dogs in training fit Ericsson's definitions of deliberate practice: Dogs are motivated to improve, they perform repetitive tasks, they receive feedback, and tasks are well defined. In this study, Helton worked with 37 dogs and their handlers. The dogs were of various breeds, of either sex, and of varying heights and weights. The participants were novices to national level competition in the sport of agility. Dogs were assessed at live events and measures of performance were simply the judging criteria used to rate the dogs after agility runs. Handlers filled out a questionnaire that assessed the amount of training hours accumulated. As predicted, hours of training were associated with performance of the dogs. Controlling for other variables such as age, height, sex, and breed did not change the relationship or effect. The author discusses the interaction between the handler and the dogs, noting that a study is underway to determine if the dog's improvement is an effect of the handler improving his or her skills (Helton, 2007).

Leisure Activities

Tuffiash, Roring, and Ericsson (2007) examined the potential role of deliberate practice within the domain of tournament-rated Scrabble players. In the study, the researchers used the players' scores as quantifiable, measureable outcomes over time. Tuffiash et al. designed a study that confirmed those with higher Scrabble scores or ratings spent more time in Scrabble-related activities, thus meeting the definition of deliberate practice. The contribution to the science of deliberate practice of this study was

that the authors moved through the three components of defining the representative skills, reproducing them in a laboratory setting and delineating the specific practice mechanisms to improve performance in a single study and not in a series of studies (Tuffiash et al., 2007).

Oulasvirta, Wahlström, and Ericsson (2011) applied deliberate practice theory to study the acquisition of skills of intermediate users of smart phones. The study was designed to determine if expert users could demonstrate superior skills and better results than novices, and if the practice attributed to differences in performance between the two groups (Oulasvirta et al., 2011). The groups participated in a series of directed tasks, and each participant provided the researchers with learning histories and interviews. Video footage recorded "think-aloud" processes, as well as solutions to tasks presented and actual task completion. Novice users were last in task completion times and performed more nonuseful tasks on their way to solutions than the casual or expert users. Casual users and experts were much closer in their performances than novices and casual users or experts. Overall the experts did not significantly outperform the casual users.

The authors concluded that the three steps to identifying superior performance in this domain were not clear and the specific skills that allowed consistent, superior performance needed to be identified (Oulasvirta et al., 2011). The authors could not identify a collection of reproducible tasks that represent the expert performance desired, map or analyze the process used by the expert to determine what techniques make the performance superior, or explain how the mechanisms identified were acquired by practice activities (Ericsson, 2000).

Overview of Literature in Public Health Emergency Preparedness

While there are no specific examples that outline the use of deliberate practice in emergency response settings, it is fair to say the theory of deliberate practice could be used to plan, evaluate, and measure performance in emergency response settings, as suggested by Salas (2009). Although literature is scarce on the theory behind the use of drills, exercises, and SBT specifically for public health emergency response, one must understand the public health environment to grasp the importance that drills played in preparing the nation for an emergency. The next sections describe the status of literature describing research in public health emergency preparedness.

Funds were set aside for public health systems research specifically targeting preparedness (Richmond, Hostler, Leeman, & King, 2010), but until 2009 there had not been a full literature search in the area to inform a research agenda (Savoia, Massin-Short, et al., 2009). Abramson, Morse, Garrett, and Redliner (2007) outlined the current state of disaster medicine and public health preparedness literature. In the work, the authors aimed to report on the research methodology found in 303 articles with disaster medicine or public health preparedness in the title published between 2002 and 2007. Of the articles, 32.3% were reviews or commentaries, 24.4% were case studies, and 12.2% used survey methodology. Few of the articles reviewed used quasiexperimental or experimental design (8.9%), and 5.3% used hypothesis-driven methods. The authors concluded there are multiple challenges facing the field due to the infrequent and complex nature of disaster research (Abramson et al., 2007).

Biddinger et al. (2008) identified many studies that confirmed exercises might improve performance in theory. The researcher recognized there was little quantitative

evidence supporting the use of drills or exercises for performance improvement. Biddinger et al. indicated operational drills were far more difficult to use to determine the quality of responses in a public health emergency but realized that moving from operations-based to discussion-based exercises made it difficult to measure the performance of participants who could not actually take action as they would during an operational event. The literature showed little progress describing the development of adequate measures of performance during or after drills and exercises.

Savoia et al. (2009) completed a literature search using search criteria and articles commonly found bibliographic databases. The search produced 1,376 studies available for analysis, of which 547 were eligible for review after the application of selection criteria. The search revealed a lack of empirical research, with most of the candidate articles identified as commentaries or reviews. Most of the earlier work was qualitative in nature, with a small increase in studies using quantitative methods from 2004 to 2008. Even with a slight increase in the later years, only 20% of the articles reviewed relied on quantitative methods (Savoia, Massin-Short, et al., 2009). The researchers concluded that the least-developed area in preparedness literature was that of the development of metrics and standards from which to evaluate preparedness. The literature also lacked new contributions to the field, with no true experiments documented and a small number of quasiexperimental or mathematical model studies presented.

In another analysis, Yeager, Menachemi, McCormick, and Ginter (2010) characterized the nature and size of public health preparedness literature. In this study, the researchers examined public health, disaster medicine, and non-health-related disaster journals to determine if there were article titles that indicated the content related to public

health emergency preparedness. Following exclusion due to criteria established by the authors, 823 articles were chosen for abstraction and review (Yeager et al., 2010). The results emphasized the knowledge gap present in the field. Less than one-third of the articles reported on empirical research, and most of those consisted of survey methodology or secondary data analysis. Primary data collection was present in 11.3% of the articles reviewed. Further analysis indicated that most of the literature focused on one of four phases of the lifecycle of disasters—the preparedness phase—with gaps in the response, recovery, and mitigation phases. Although exercises and drills were the most frequent topic of the articles, these represented only 12.3% of the entire sample. Yeager concluded that the body of literature was overwhelmingly subjective.

Reviewing the literature in the field of public health emergency preparedness indicated a body of knowledge stemming from descriptive and qualitative methodology, and lacking empirical and quantitative design. The secondary use of data to provide snapshots of certain aspects of preparedness, such as the use of drills and exercises, or organizational design has been a theme. The lack of the use of metrics or standards to measure emergency preparedness is a gap in the knowledge base. The next section explores literature that highlights the role of public health emergency preparedness within the larger context of the public health system.

Public Health Systems and Emergency Preparedness

Several authors contributed to the body of knowledge describing the status of public health systems in the era of preparedness funding (Beitsch et al., 2006; Frank, 2005; Fraser & Brown, 2000; Kinner & Pellegrini, 2009). While the contributions were descriptive or qualitative, the information served as a baseline from which to understand

the progress made over the 10 years of preparedness funding, with emphasis on both state and local health departments.

Dual-Use and Public Health Infrastructure

Fraser and Brown (2000) introduced the idea of preparedness capacities being built on existing core activities of public health agencies to avoid creating a "preparedness silo" of activities, personnel, and procedures, calling the idea "dual use." Fraser and Brown cited the work of local health agencies when dealing with the emergence of the West Nile virus as an example of the utility of preparing resources for dual use. Fraser and Brown discussed results from surveys that indicated many county and city health departments had not invested in preparing response plans and the contribution of federal guidance on standards or templates had not helped promote these efforts (Fraser & Brown, 2000).

Morse (2002) attempted to fill the information gap cited by Fraser and Brown by producing a guide for local health agencies that outlined the processes needed for collaboration and response planning at the local level. The guide discussed activities (such as conducting a local assessment to look for system strengths and weaknesses, and learning the incident command system) and relationships (such as those with other local health departments, the state health department, the CDC, and the Emergency Management System). It also identified key players to help organize the plan and play an integral role in its execution should an event befall the community. The guide served as a starting point for many local public health agencies and emphasized the primary role of the local health department while accentuating the partnerships needed to be fully

prepared (Morse, 2002). The information in the guide was subjective and not based on qualitative or quantitative data analysis.

The Journal of Law, Medicine & Ethics reprinted a speech in 2002, former senator Sam Nunn addressed the strain on local, state, and federal public health agencies and made the case for dual use of bioterrorism funding at the state and local levels. He advocated for an integrated public health, research, and medical care approach to prevent and respond to terrorism, specifically building the infrastructure of the systems that were in place at that time. He proposed that new terrorism funding should benefit the entire public health infrastructure, thus moving a target directly on prevention activities, which would not only aid in future biological terrorism attacks but also protect the nation from other preventable diseases such as West Nile virus. He concluded that the dollars appropriated for bioterrorism would aid in improving the health of all Americans and protecting the nation from biological weapons (Nunn, 2002).

Baker et al. (2005) contributed an article describing the status of the public health infrastructure, in which they stated the influx of funding for public health preparedness would improve some aspects of the public health infrastructure. The article also stated that it was likely that as federal funding increased, state and local health departments would reduce funds for general public health functions (E. L. Baker et al., 2005). Baker et al. discussed performance of public health entities, using the 10 essential components of public health as a yardstick. Prior to 2002, local public health officials reported they were unable to carry out the roles necessary to respond to a public health emergency. Baker et al. suggested a strong program of accreditation as a method to assure local health departments built the capacity to fulfill duties related to public health emergencies.

Morrow (2007) called for a balance in both funding and efforts to assure that the public health infrastructure was robust enough to address all preventable disease burdens and would not create silos for specific issues, such as bioterrorism. The reported global human impact of disasters in 2004 was over 241,000 deaths and over 1.45 million affected by natural disasters. The reported global economic impact the same year was over $100 billion in damages. However, the same year, over 1.2 million individuals in the United States lost their lives to cancer and heart disease. Morrow argued for strong balance to address current threats as well as theoretical ones (Morrow, 2007).

While most were applauding the stream of funding moving into public health for preparedness, some experts were concerned about how funds were distributed to states for preparedness activities (Buehler & Holtgrave, 2007). Frank (2005) was concerned that the high levels of funding earmarked for preparedness were missing the mark for true change in population-based morbidity and mortality. Hyde (2006) described the costs to general public health functions as an unexpected consequence of emergency preparedness planning.

Buehler and Holtgrave (2007) pointed out the difficulty in providing formula funding for terrorism preparedness without taking into consideration specific threats based on the jurisdiction considered for funding. They also pointed out that the frequent "set-asides"—portions of funds earmarked for specific activities and thus not subject to the same rules when distributed by formula—rendered formulas ineffective. Frank (2005), on the other hand, expressed concern over the disproportionate amount of funding moving into preparedness activities when funds to combat the leading causes of death in states received much smaller amounts of federal funding. For example, New York

received just $1.3 million to reduce heart disease, which killed nearly 40% of New Yorkers in 2002; yet the state received over $34 million the same year to combat terrorism (Frank, 2005). Hyde, Kim, Martinez, Clark, and Hacker (2006) conducted interviews with local health departments in Massachusetts and noted benefits to public health such as staff development opportunities and multisector collaboration, as well as several challenges, including lack of sustainable resources for staff, confusion about shared use of equipment, and dual use functionality. The respondents felt those in charge had little appreciation of the effort and time invested in preparedness planning in light of the few available resources. They also felt funds were shifted to preparedness activities, reducing efforts in general public health programming, which created frustration among staff. The authors concluded with a statement that called for increased research to determine if the benefits of federal funding were sufficient to outweigh collateral damage caused by focus on a single issue instead of implementation of traditional public health programs.

Beitsch et al. (2006) described the placement of preparedness activities in state-level public health departments in an effort to describe the variety of organizational designs employed following the influx of preparedness funding. The authors made a statement about the accomplishments in public health preparedness but cited no formal evaluation of performance. A questionnaire distributed to state agencies contained 27 questions regarding activities related to federal funding. Over half the states chose to decentralize their activities; many subdivided into preparedness regions (Beitsch et al., 2006).

States chose to create a stand-alone preparedness office as often as they created matrix or mixed offices (incorporating preparedness activities within an existing office). The leaders of preparedness programs typically reported to the state health official or deputy (Beitsch et al., 2006). Most had responsibility for both the CDC's grant and the companion grant for hospitals administered by the Health Resources Services Administration. Regarding performance measures or metrics, less than half of the programs chose to select additional measures to chart performance in addition to CDC measures and requirements. One-third of states supplemented federal dollars with state funds, and federal/state funds allowed for an average increase of staff at the state level of 42 positions dedicated to preparedness (Beitsch et al., 2006). Table top exercises were the most frequent form of drills. Analysis by geography, population size, organizational structure, or region showed no trend differences regarding exercises or drills. The authors encouraged further research to examine the relationship between additional funding and actual performance in events or drills (Beitsch et al., 2006).

Following up on the report by Beitsch et al. (2006), Duncan, Ginter, Rucks, Wingate, and McCormick (2007) also reported on the variety of organizational designs associated with preparedness programs in the United States, although their goal was to examine organizational design in a framework of equifinality rather than public health and its essential functions. Viewed through this lens, the authors concluded that successful leaders in emergency preparedness relied more on expert power than positional power. They also exhibited a strong ability to influence and persuade others to accomplish their goals, and they relied on diplomacy and negotiation skills. The analysis

also indicated that successful emergency preparedness leaders returned favors for those from whom they asked favors (Duncan et al., 2007).

Lovelace, Bibeau, Gansneder, Hernandez, and Cline (2007) evaluated the ability of local health departments in North Carolina to prepare fully for all-hazards events, a shift in the CDC guidance to states that took place in 2005. Practitioners began to see an all-hazards approach as further preparation for low-probability, high-consequence events that might divert needed funding and human resources from typical public health practice (Lovelace et al., 2007). Lovelace surveyed local health department leaders to determine the self-reported level of preparedness for different events: bioterrorism, chemical, natural disasters, radiation emergencies, outbreaks, and mass trauma. Participants identified preparedness levels on a scale of 1 to 4, with 1 being the lowest score and indicating lack of preparedness for that type of event. The results showed that more funding was associated with better preparedness for bioterrorism, radiation, outbreaks, and mass trauma. Those health departments that experienced natural disasters (e.g., hurricanes and flooding) considered themselves more prepared for those events. Those that had a need to prepare for radiation emergencies due to the proximity of nuclear power plants scored themselves high in that category. The results supported Buehler and Holtgrave's (2007) suggestions that formula funding would not lead to all-hazards preparedness but could lead to preparedness for the events most likely to happen in a particular jurisdiction. Lovelace did not examine actual performance differences between or among health departments. The self-reported scores were the participants' perception of the potential of response and did not reflect on actual performance.

Koh, Elqura, Judge, and Stoto (2008) confirmed the notion of regionalization presented by Bietsch et al. (2006) by describing the history, lessons learned, and case studies of municipal services conducting business on a regional scale. Regionalization of some activities in public health had already begun, such as some laboratory services and mutual aid agreements. The case studies included information from Massachusetts and the national capital region. The case studies indicated that participants in the regions felt networking improved and standardization of equipment became more frequent, and they liked sharing resources for training. Even with regionalization, however, problems still existed within the partners. Legal governance was not addressed, and there may have been regions established within the state for different purposes that were not congruent with those established for public health emergency response (Koh et al., 2008). Koh et al. readily advocated for further research, including developing clear outputs that measure preparedness. They also confirmed concerns raised by Hyde et al. (2006) and Frank (2005) that practitioners felt preparedness dollars might be detracting from general public health activities.

Hargan (2008) outlined the federal role in emergency response framed within the context of federalism and public health law. The article explored the legal boundaries of the federal government when responding to a disaster and described the objectives determined to be federal in nature: disease monitoring, stock-piling countermeasures, developing vaccines, establishing communications plans, and setting up local plans (Hargan, 2008). The article concluded with a statement that a balance should exist between federal and state partners, but the onus of responsibility for planning and responding was a state function. The federal government is limited in its ability to act

based on the Constitution, and the limitations have been supported by legal decisions as decided by the Supreme Court and legislative and appropriations decisions made by Congress.

Avery and Wright (2010) analyzed the effect of federal funds provided by the CDC on the preparedness level of local health departments. The authors stated the federal government was in a good position to provide technical assistance and resources to state health departments, but was less equipped to provide such resources to local health departments (G.H. Avery & Wright, 2010). The study examined the provision of resources to determine if correlations existed among funds from the CDC and increased preparedness activities at the local level. The study also attempted to determine if the performance of public health activities had a relationship with local leadership. The study did not look at local public health preparedness performance measures. The researchers analyzed data from the 2005 National Association of City and County Health Officials national profile dataset. The authors created a model with variables representing activities that might influence public health activities and scaled responses from health departments to the responses of interest. The results indicated no direct relationship between CDC funding and level of preparedness activities at the local level. Health departments that hired an individual with responsibility to coordinate preparedness activities showed a strong relationship between the coordinator and the level of preparedness activities. There was also a strong relationship between the CDC funding and epidemiology activities but not for any other activities required (e.g., communications or laboratory capacity). Based on the analysis, the authors concluded that the idea that federal funding passed to states

and then to local levels improved the state and local public health systems was incorrect; therefore they suggested further research.

Klaiman and Ibrahim (2010) studied how health departments were structured and if structure had an influence on pandemic planning. The authors recognized the lack of research about effectiveness in areas such as preparedness and response in regards to health department structure and its impact on performance. The dependent variables were data obtained from the CDC's state and local pandemic planning checklist, and state pandemic plans were compared specifically against the checklist. Researchers coded the data on a three-point scale based on written documentation in the plans. Independent variables were chosen based on previous work by Ford, Duncan, and Ginter (2003), using five clusters of variables identified through qualitative research to define key organizational variables inherent to public health agencies. Ford et al.'s variables were used to test the impact of each on preparedness scores drawn from the written plans. Regression models showed that professional, trained staff had an impact on overall preparedness. Longer tenure had a negative effect on preparedness, with results leading the researchers to conclude that those who have been at the health department for a long time favored the traditional epidemiology and laboratory skills more than the new skill sets needed for preparedness—such as collaboration skills, integrated command, and quick action (Klaiman & Ibrahim, 2010). The researchers focused on planning and not on specific performance measurement during a response.

The dominant literature contains descriptive methodology, secondary data analysis, and commentaries. Public health systems have had to assimilate specific emergency preparedness activities into the larger system and have advocated for dual-use

permission when planning for and expending funds. To attain preparedness goals, researchers have explored creative methods such regionalization and other strategies. A knowledge gap identified was the standardization of metrics to measure preparedness at the systems level. The next section examines the literature as it relates specifically to the use of exercises and drills as a method to test and improve preparedness.

Exercises and Drills

The notion of drills and exercises and other SBT activities as concrete methods to practice various aspects of response was institutionalized by the Department of Homeland Security (DHS) and confirmed in the National Strategy for Homeland Security (National Security Council, 2007). DHS established the Homeland Security Exercise and Evaluation Program (HSEEP) to provide standardized methods and techniques when developing and executing drills and exercises. DHS stated that the advantages of an exercise program included assessing and improving performance (Homeland Security Exercise and Evaluation Program, 2007a).

Lee, Trim, Upton, and Upton (2009) authored a descriptive paper on large emergency response exercises and analyzed the multiple differences between large and small exercises. In the paper, Lee defined a successful exercise as one in which the participants learn (Lee et al., 2009). The authors also commented on the actions desired or observed during large exercises in the context of learning, modern art, group psychotherapy, and management theories, hypothesizing those unsuccessful exercises might not be rooted in theory. The diversity of players allows the incorporation of many theories into exercise design. Lee et al. provided one of the few references to theories appropriate for developing exercises. The authors also emphasized the need for

interdisciplinary input when designing scenarios. Public health systems have participated in national, state, and local drills and exercises. The next section describes outcomes of the experiences as reported in the literature.

Public Health Participation in Exercises

There are four types of discussion-based exercises defined through HSEEP: seminars, workshops, tabletops, and games. There are three types of operational exercises promoted through HSEEP: drills, functional exercises, and full-scale exercises (Homeland Security Exercise and Evaluation Program, 2007b). This literature review focused on studies involving public health, hospital, or medical training incidents that involved the use of either discussion-based or operational exercises. In most cases, the literature described discussion-based events, such as table top exercises or workshops. There is a growing interest in using computer simulation as a method of exercising emergency response, and SBT warrants review. Articles describing the intent of exercises often cited performance improvement as a desired outcome. However, a majority of the data reported did not include performance metrics.

Outcomes Experienced

Beaton et al. (2004) evaluated the impressions and opinions of dispensing site workers who participated in an exercise designed to test the state's plan for dispensing medical material following an emergency. The exercise was more complex than a discussion-based drill because participants moved and handled material as part of the scenario. Results indicated that the participants had increased confidence levels following the exercise as demonstrated by an analysis of survey data. The article did not provide

any information about the performance aspect of the participants, but the subjective opinion of how they felt after the exercise was completed (Beaton et al., 2004).

Gebbie, Valas, Merrill, and Morse (2006) undertook research to develop a standardized set of criteria to provide local, state, and federal public health authorities with a method to appropriately match exercises, purpose, and level of preparedness for the agency. Gebbie et al. used a Delphi survey method to develop an instrument that could be used to measure actions taken by public health agencies during a response. The work resulted in a format that allowed exercise observers to record the responses of agencies participating in drills or exercises to minimize subjective interpretation of the actions. However, there were not subsequent articles published indicating there had been widespread use or adoption of the criteria after its development.

High et al. (2008) outlined a niche for health educators in writing and organizing table top exercises to simulate responses to disasters. The authors put forth the premise that exercises can be beneficial to participants to improve the relationships needed to respond to emergencies. The intent of the exercise was to assess and improve the performance of the agencies participating. However, when describing the research questions, all responses were qualitative in nature and did not report specific quantitatively measured performance criteria. As with other studies, High et al. observed that over half of the participants indicated through surveys that they were more confident that the sectors involved in the response could work well together. The participants also felt the exercise helped identify gaps in the response plan.

Ablah, Nickles, Hodle, and Wolfe (2008) designed and tested a 4-week exercise to simulate an outbreak of infectious disease. The results of the study followed the

analysis of focus group results. The participants self-reported that they felt better prepared after the drill, and the authors interpreted the data to mean the greater collaboration shown during the exercise had a positive effect on the participants' confidence levels.

Johnson, Herrmann, Wallace, Troutt, and Myint (2009) described the development and implementation of a functional exercise that tested the public health response to a simulated foodborne outbreak attributed to terrorism. The authors make a strong case for operations-based exercises and a fundamental component of training public health staff. The exercise design allowed agencies to demonstrate the effectiveness of their training. The exercise evaluation was judgment based and relied on the opinions of the participants, not on actual performance. The participants responded to survey questions following the exercise. The participants did not feel strongly the exercise allowed participating agencies to improve their responsiveness. However, the authors felt the exercise was a success in that it allowed for improved relationships among responders, and they felt the efficacy of the plans and capability of the personnel to implement the plans were adequately tested.

Biddinger, Savoia, Massin-Short, Preston, and Stoto (2010) authored a synopsis of the experiences of the Harvard School of Public Health's Center for Public Health Preparedness exercise program between the years of 2005 and 2009. In the article, Biddinger et al. stated that exercises evaluate the performance of systems, individuals, or agencies. To meet the four aims of the evaluation, Biddinger et al. used three surveys completed by participants and did a content analysis of the after-action reports developed by the teams following the exercise. While the work completed over the years included

both discussion- and operations-based exercises, the data discussed were limited to those participating in table top discussions. Biddinger et al. also concluded from the analysis that the participants indicated they felt more confident about the subject matter after the table top discussion than before. The researchers concluded that table top exercises could be useful in evaluating specific capabilities of public health systems, especially when the right measurement tool is available. However, the data reported did not include any capabilities measures.

The research by Fowkes, Blossom, Sandrock, Mitchell, and Brandstein (2010) concentrated on 90 table top exercises conducted through the Area Health Education Center programs in California. The aim of the study was to determine if there were adequate plans in place for the communities using exercises. The results, captured through surveys, interviews, content analysis, and reports, indicated that 91% of the clinics participating changed the planning document following the participation in the drill. The researchers felt the participation in the exercises strengthened long-term relationships with community partners. There were no data to indicate whether the participants could implement the plans, and there were no data indicating whether the experience improved performance of the participants.

Olson, Scheller, Larson, Lindeke, and Edwardson (2010) used a simulated public health emergency as a vehicle to determine if general training resulted in better performance in a simulation. The researchers placed participants in one of three groups based on self-reported bioterrorism training and educational experiences. The results indicated those who had significant training or education in bioterrorism answered more questions correctly than the nontrained control group, but the most highly trained

individuals were the least efficient, taking the most time to complete the chapters. The time needed by the most experienced to complete the chapters was a surprising finding for the researchers, although they felt the information should not be interpreted in a negative light (Olson et al., 2010).

Hospital and Health Care Participation in Exercises

Mann, MacKenzie, and Anderson (2004) analyzed a state-by-state trauma needs assessment for mass-casualty events appended with a state disaster preparedness component after the events of September 11, 2001. The report discussed results of the state disaster preparedness component, correlating information from states with information gleaned from the funded Metropolitan Medical Response System (MMRS). The report also described a scale depicting the maturity level of the state trauma systems. The result was a point-in-time picture of the state's preparedness efforts to determine if previous MMRS investments had an influence on preparedness scores. Emergency preparedness activities addressing both public health and medical responses appeared in the plans of nearly all states. Fifty-eight percent of states had conducted drills or activated the plan to test the plan's completeness. Most states did not require mandatory training in disaster management, biological terrorism, or chemical terrorism for hospital or emergency management system (EMS) personnel. Resources such as personal protective equipment were in short supply for EMS personnel, and only five states indicated that all hospitals in the state had decontamination facilities (Mann et al., 2004). Only eight states had systems linking hospital bed capacity, staffing, and supplies to need in real time. Finally, the number of MMRS projects within a state had no effect on the total preparedness scores in this study. There were significant associations between the

maturity of the state trauma system and preparedness scores, with states that had legislated trauma systems in place having the highest preparedness scores, which may have been the result of local level integration of public health and medical practice. The authors recognized the weaknesses in self-reported data and reminded readers that the reports did not substitute for a demonstration that proves implementation is possible.

The research by Hsu et al. (2004) performed a systematic review of the roles of drills and exercises and their effectiveness in training hospital staff for mass-casualty events. The researchers identified 243 citations, of which 21 were eligible for further review based on criteria. Hsu et al. found most of the studies examined had poor study and evaluation design, and thus limitations. Although most research done with public health departments described table top exercises (Yeager et al., 2010), none of the studies in Hsu et al.'s review were table top exercises. The results indicated hospital employees who participated in drills became more familiar with procedures and response components, but there was little evidence to determine whether drills were the most effective teaching method for this population.

Hsu et al. also found the authors of the studies did not typically specify the methods used for evaluation of the drills, and when mentioned, they involved the opinions of expert observers or surveys of participants. Simulations that used technology, such as computers, provided evidence of increased practical skill obtainment, such as increased knowledge of treatment options, identification of issues that could slow down a response, and increased information retention. Because there were no table top exercises that met the research criteria, the authors made no statement about the effectiveness of the discussion-based exercise as an educational tool in a hospital setting. Hsu et al.

concluded there was a need for well-designed studies and evaluation tools before determining the effectiveness of mass-casualty drills for educating hospital staff.

Katz, Staiti, and McKenzie (2006) conducted surveys and interviews to determine the progress made by communities after 2 years of preparedness funding. The public health respondents indicated they felt community preparedness levels had increased, communication among those designated as first responders had increased, laboratory capacities had improved, and they felt better prepared to respond to emergencies. However, provider communities could not make the same statements. Hospitals and health care providers felt funding was insufficient to meet the new demands in information systems and training because of staff turnover. Large hospitals felt unprepared to handle a surge in patients, and being prepared for the range of issues posed by the variety of threats (e.g., nuclear, biologic, or chemical) was unreachable. The article concluded with a statement that public health was more prepared, but health care was lagging behind (Katz et al., 2006). The study data were qualitative and self-reported.

Lewis, Berg, and Gates (2006) conducted research on a full-scale exercise designed to improve hospital preparedness to add to the body of knowledge supporting the value of frequent exercises (twice per year), as hospitals are required to perform to maintain accreditation. Trained participant–victims were health care providers who recognized appropriate hospital response. The participant–victims completed a 14-question survey after the exercise. The results indicated there were no significant differences in hospital performance if the staff participated in the required two exercises compared to one exercise annually, and there was no difference between hospitals that did not participate in the exercise the prior year and those that participated annually.

Lewis et al. proposed staff turnover, failure to train more frequently, and failure to correct prior deficiencies as possible reasons for the lack of long-term improvement by the facilities.

Bartley, Stella, and Walsh (2006) described efforts to corroborate Hsu et al's (2004) systematic literature review assessing the utility of disaster simulations as a method of testing hospital disaster plans. The article presented findings of surveys conducted with 50 participants likely to be in positions of responsibility during a disaster. The authors aimed to test if participation through an audiovisual presentation followed by a simulated disaster exercise and debriefing provided the same gains in knowledge and confidence as a full-scale exercise. The authors suggested achievement of the aims of the test, which was to increase knowledge of the staff and to assess self-reported preparedness data from participants. However, the low ambitions of the study highlighted the need for multiple strategies to continue to assure the hospital staff is prepared, including employing a disaster manager to maintain proficiency of the staff and mandatory participation in the modified full-scale exercise for Australian hospital staff.

Morrison and Catanzaro (2010) described the use of simulations as teaching tools in undergraduate nursing courses. The authors described a series of public health emergencies in which the student nurses assumed roles (e.g., health care professional, casualties) and made decisions about infection control, isolation, quarantine, and general patient safety. The qualitative evaluation indicated the simulation engaged the students, the students thought the simulation was helpful and appropriate for the class, and they felt they met the learning objectives. There was no performance evaluation associated with the simulation (Morrison & Catanzaro, 2010).

Time Series Analysis

Time series experiments have been advocated for the evaluation of community-level interventions, specifically in the role of evaluating policy as intervention (Biglan, Ary, & Wagenaar, 2000). The classic time series analysis was published in 1968, in which Campbell and Ross introduced the time series analysis as a mode of analysis to be considered when a researcher has little or no control over the assignment of cases to control or experimental groups (Campbell & Ross, 1968). Campbell and Ross (1968) presented how the validity of experiments can be violated unless a robust time series design is used.

Eccles, Grimshaw, Campbell, and Ramsay (2003) advocated using a framework for choosing the appropriate analysis of complex interventions; however they also acknowledge that the sequence recommended may not always be possible, especially when evaluating policy interventions with little or no evidence supporting their success. Time series designs are discussed as one of three options for nonrandomized designs (Eccles et al., 2003). The time series design attempts to reveal intervention effects that are greater than expected, or secular changes. The authors suggested statistical techniques such as autoregressive integrated moving averages (ARIMA) and time series regression modeling as adequate methods to detect changes due to intervention (Eccles et al., 2003).

There is evidence in the literature that time series designs are appropriate to evaluate terrorist incidents (Barros, 2003; Cauley & Im, 1988; Enders, Parise, & Sandler, 1992; Enders & Sandler, 2002). Public health policy and cost effectiveness data have also been analyzed using time series analysis (Holder & Blose, 1987; Keeler, Hu, Barnett, &

Manning, 1993; Lawrence et al., 2003; Soumerai, McLaughlin, Ross-Degnan, Casteris, & Bollini, 1994).

Summary

Deliberate practice theory details how expertise develops over time with clear, concise, deliberate practice sessions. Deliberate practice increases performance in many domains but has not been explicitly linked to the use of drills, exercises, and SBT in the domain of public health emergency response. While public health agencies have been participating in exercises, very few have been subject to specific performance measurement. The results of drills and exercises that have contributed to the body of knowledge are generally qualitative in nature and reflect feelings of improved confidence among the participants.

I integrated aspects of deliberate practice theory with actual performance results from public health departments that have participated in drills and exercises to determine if practice influences performance. Chapter 3 describes the research design and approach; the setting and samples; the use of drills, exercises, and real events as the basis for timed performance measures; and the data analysis methods proposed. Chapter 4 outlines the data analysis and results, and Chapter 5 includes results and recommendations.

Chapter 3: Methodology

Introduction

The study design was to determine if drills and exercises positively influenced performance of public health responders in emergencies. Although there are many studies describing exercises and drills, as well as qualitative outcomes, there is a gap in the information linking drills and exercises to improved performance. Descriptions of the research design and approach, the setting and sample sizes, the structure of the performance measures and other instruments used to test performance, data collection, and analysis methods are presented, as well as protections used to assure the protection of participants during the conduct of the study are presented in this chapter.

Research Problem

The research question addressed whether drills and exercises positively influenced public health responders in public health emergencies. Due to the complexity of developing a causal relationship between activities and outcomes, performance measurement in public health emergencies is difficult (DeGroff et al., 2010). A quantitative study controlling factors that might influence performance, such as size, financial support, and structure for state governmental public health departments, could help determine if drills and exercises influence public health performance in emergencies. There have been several qualitative studies regarding the perception of the strengths and weaknesses of the interactions among responders during exercises (Dausey et al., 2007; Gebbie et al., 2006). It has been shown that levels of confidence among responders improved following drills and exercises (Savoia, Biddinger, et al., 2009). Researchers called for the development of standard measures of acceptable performance (Biddinger et

al., 2010; Lurie et al., 2006). With a lack of standard measures of acceptable performance, there is little or no published research comparing performance of public health entities before and after exposure to drills or exercises. Since 2002 the CDC created several performance measures for specific public health functions in an emergency (U.S. Government Accountability Office, 2007). PAHPA (2006) codified the use of drills to improve public health response in an emergency. The CDC began requiring the use of established performance measures as a method of measuring preparedness in 2005 (U.S. Department of Health and Human Services & Centers for Disease Control and Prevention, 2005). Regular, systematic reporting of consistent performance measures during drills, exercises, and real events was established in 2007 (U.S. Department of Health and Human Services & Centers for Disease Control and Prevention, 2007). Analysis of the performance measures before and after drills and exercises in relation to real events could shed light on their utility.

The idea of practice makes perfect has been the focus of research in the medical field through the use of simulation (Friedrich, 2002). Only one researcher linked learning theory to the purpose of drills and exercises (Lee et al., 2009), but theories of adult learning and performance based on experience or repetition are well researched. ELT (David A. Kolb, 1984) outlines specific stages adults experience when learning through experience. The theory of deliberate practice posits the concept that many individuals can achieve exceptional performance when exposed to specific coaching methods (Ericsson, 1996; Ericsson & Lehmann, 1996). An appropriate approach to modeling and analyzing the behavior of public health systems before and after the introduction of an intervention is discussed.

Research Design and Approach

When conceptualizing this research, a quasiexperimental, time series design was conceived to determine if drills and exercises had an effect on performance. Specifically, it was thought the research design was a "multiple group–single intervention" design (Glass, Wilson, & Gottman, 2008, p. 23). The design was thought to be appropriate because the organizations participating were state health departments that received public health emergency response funding from the CDC and shared characteristics. The goal of the quasiexperimental design was to separate the effects of treatments from the other noncomparable characteristics of the experimental units.

Subsequent considerations determined the traditional time series design was not appropriate for this study for several reasons. Factors leading to the determination follow. The data were collected on entities (e.g., state health departments) and events (e.g., drills, exercises, and real events relating to public health emergency response). For each entity, the data contained information about location, money received, staff size, and plans to spend grant money. These data did not lend themselves to a time series analysis.

Data for each event included the number of drills, exercises, and real events captured during the reporting period and the duration of the event. The research question focused on the drill or exercise as the treatment at the entity level, yet data at the entity level were sparse, with some entities having less than 10 observations per measurement. With sparse data at the entity level, a time series analysis would not have the strength necessary for the evaluation of legitimate relationships.

The entities had the freedom to choose which events to report during a specified period, which meant the measurements were irregularly spaced and differed in total

number from entity to entity. As the research question was designed to treat the time between measurements as an independent variable, it would be improper and would introduce bias to use statistical methods to transform unevenly spaced observations into evenly spaced observations with the use of interpolation (Eckner, 2012).

Aggregating data to create larger data sets, such as looking at information in quarterly or monthly periods also introduced a dependence that would prevent examining the time between measurements as a variable, which also made a time series analysis improper for the data set. The measurements chosen were all counts of units of time, and the conclusion was that a "time-to-event" time series model was more appropriate than a traditional time series such as ARIMA.

An alternative model developed to analyze count data is the Poisson regression model. Poisson regression can be described as one of many generalized linear models, and its foundation is the Poisson distribution, a depiction of the distribution of errors (Coxe, West, & Aiken, 2009). The Poisson distribution for count data was developed because of its ease of interpretation, and its similarity to linear regression models (Hutchinson & Holtman, 2005; Karazsia & van Dulmen, 2008). Count data are composed of variables that are positive integers of discrete value that reflect the number of occurrences of an event in a fixed period of time (Coxe et al., 2009). To be of good fit, the data used in Poisson models must meet two assumptions. The mean and variance of the count data are assumed to be equal, and the events observed are assumed to be independent of each other (Karazsia & van Dulmen, 2008). The Poisson regression model is depicted as an equation:

$$\ln(\hat{\mu}) = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p,$$

where ln is the natural logarithm, $\hat{\mu}$ is the predicted count on the outcome variable, b_0 is the intercept, b_1 is the regression coefficient, and X_1 is the first predictor (Coxe et al., 2009).

The participants were required to submit reports to the CDC indicating the number of drills, exercises, or real events they participated in during a specified period and to submit timed response for three performance measures. The plotted timed measures allow readers to recognize when the intervention occurred. The statistical analysis methods used determined if the intervention had an effect on the time to complete the task during a real event. Because the intervention of drills and exercises were staggered across specified periods of time (at the will of the experimental units), the research design strengthens the ability of the researcher to demonstrate if an intervention is effective over time and not due to coincidence (Glass et al., 2008, p. 25). Based on the type of data collected, a stratified, multiple-group, single-intervention design defined the experimental units. Examining several other factors, such as size, budget, and number of personnel determined if there were characteristics of states that played a role in different responses to the intervention.

Since 2007 all states are required to submit data on certain performance measures (U.S. Department of Health and Human Services & Centers for Disease Control and Prevention, 2007). The measures are construct referenced; participating health departments compete against their own best score, not the scores of others (William, 1996). The benchmark for the timed performance will be competition for the participants, not competition among the participants. Each participating state received the protocols for data collection, developed by the CDC. The protocols included the measurement

components, such as the intent of the measure, the data collection and reporting processes, definition, and other guidance (U.S. Department of Health and Human Services & Centers for Disease Control and Prevention, 2007). Each entity was required to submit up to six measures for two 6-month periods in the budget year (potentially 12 measures per entity). The measures were associated with a drill (e.g., table top, full-scale exercise, or functional exercise) or a real incident.

The review included three measures. Table 1 contains the measures and capabilities for the study.

Table 1

Public Health Capability and Performance Measures

Capability	Measure
Communication	Time to notify all primary staff (secondary or tertiary staff as needed) with public health agency incident command system functional responsibilities that the public health agency's emergency operations center is being activated.
Response	Time for primary staff (and secondary or tertiary staff as needed) with public health agency incident command system functional responsibilities to report for duty at the public health agency's emergency operations center.
Implementation	Time to complete a draft of an after-action report/improvement plan.

Sample Size and Setting

The sample size was 50 governmental public health departments or agencies that resided at the governor's cabinet level in their respective states. The state health departments had flexibility in creating the organizational structure necessary to fulfill the requirements of the public health preparedness activities. Beitsch et al. (2006) described

the placement of preparedness activities in state-level public health departments in an effort to portray the variety of organizational designs employed following the influx of preparedness funding. A questionnaire that was distributed to state agencies contained 27 questions regarding activities related to the federal funding. Over half the states chose to decentralize their activities, with many subdividing themselves into preparedness regions (Beitsch et al., 2006).

Regardless of the structure chosen, each entity was responsible for a single report submitted to the CDC detailing the performance measures. The Outcome Monitoring and Evaluation Branch (OMEB), Division of State and Local Readiness (DSLR), and Coordinating Office of Terrorism and Preparedness and Emergency Response (COTPER) received electronically submitted data. Members of the division verified the data through site visits or conference calls.

Two periods for each of the fiscal years 2006 and 2007 data were examined. Fiscal years 2006 and 2007 were selected because the questions were worded in a similar manner. The corresponding questions for 2008 and 2009 were changed enough to be noncomparable. In fiscal year 2006, grantees were provided with 23 measures to report to the CDC. In 2007 grantees were provided with six measures to report to the CDC.

The answer to Question 6B, “Time to notify all primary staff (secondary or tertiary staff as needed) with public health Agency Incident Command System functional responsibilities that the public health agency’s Emergency Operations Center is being activated,” was measured in minutes. It was a self-reported measure submitted to the CDC via electronic reporting. The grantee could have submitted at least two and up to six events (real or exercises/drills) for each reporting period, with a potential of eight to 24

total events for the 2-year period. Professionals in the field agreed to a target of 60 minutes to complete the notification of staff, and receive acknowledgement from the staff contacted.

The answer to Question 6C, "Time for primary staff (secondary or tertiary staff as needed) with public health agency Incident Command System functional responsibilities to report for duty at the public health agency's Emergency Operations Center," was measured in minutes. It was a self-reported measure submitted to the CDC via electronic reporting. The grantee could have submitted at least two and up to six events (real or exercises/drills) for each reporting period, with a potential of eight to 24 total events for the 2-year period. Professionals in the field agreed to a target of 150 minutes to complete the assembly of staff.

The answer to Question 9A, "Time to complete a draft of an After-Action Report/Improvement Plan," was measured in days. It was a self-reported measure submitted to the CDC via electronic reporting. The grantee could have submitted at least two and up to six events (real or exercises/drills) for each reporting period, with a potential of eight to 24 total events for the 2-year period. Professionals in the field agreed to a target of 60 calendar dates to complete the after-action report.

In total grantees could have submitted 72 events during the two fiscal years for each of the capabilities shown in Table 1. The first reporting period was from August 31 to February 29, and the second was from March 1 to August 8 for each fiscal year. (Raw data from the 50 participants is available upon request from the researcher.) Of interest were the measurements of time for real incidents following a drill or an exercise (the intervention).

Structure of Performance Measures

The protocol for the definition, collection, calculation, and reporting of the performance measures was distributed to grantees with program guidance. Tables 2–4 contain definitions of the requirements for the measures.

Table 2

Performance Measure 6B

Term	Definition
Performance measure	Time to notify all primary staff (and secondary or tertiary staff as needed) with public health agency incident command system (ICS) functional responsibilities that the public health agency's emergency operations center (EOC) is being activated.
Intent	To ensure timely and effective coordination within the public health agencies and with key response partners in a complex incident, grantees must demonstrate the capability to rapidly notify staff to report for public health EOC duty and track responses to ensure that all eight core ICS functional roles can be staffed with one staff person per position. This capability is critical to maintain even though not every incident requires full staffing of the ICS. Rapid notification of staff depends on maintaining accurate contact information for preidentified public health agency staff to fill each incident command system functional role.
Target	Mean[a] = 60 minutes from time that the public health director or designated official begins notifying preidentified primary staff (and secondary or tertiary staff as needed).
Measurement specifications	Start time: Date and time that public health director or designated official began notifying preidentified primary staff (and secondary or tertiary staff as needed). Stop time: Date and time that the last primary staff person (and secondary or tertiary staff as needed) to fill all eight ICS functional roles acknowledged notification.

[a]Mean based on all reported drills, exercises, and real incidents for the budget period.

Table 3

Performance Measure 6C

Term	Definition
Performance measure	Time required for primary staff with public health agency incident command system functional responsibilities (and secondary or tertiary staff as needed) to report for duty at the public health agency's EOC.
Intent	To ensure timely and effective coordination within the public health agencies and with key response partners in a complex incident, grantees must demonstrate the capability to rapidly staff all eighth core ICS functional roles in the public health EOC with one staff person per position. This capability is critical to maintain even though not every incident requires full staffing of the ICS. Rapid staffing the public health agency's EOC requires that the agency preidentify staff to fill these roles and that staff can receive and acknowledge notifications and report immediately to the EOC.
Target	Mean[a] = 2.5 hours from time that public health director or designated official notifying preidentified primary staff (and secondary or tertiary staff as needed) that the public health agency's EOC will be activated.
Measurement specifications	Start time: Date and time that public health director or designated official began notifying preidentified primary staff (and secondary or tertiary staff as needed) that the public health agency's EOC was being activated. Stop time: Date and time that the last primary staff person (and secondary or tertiary staff as needed) to fill all eight ICS functional roles signed in at the EOC.

[a]Mean based on all reported drills, exercises, and real incidents for the budget period.

Table 4

Performance Measure 9A

Term	Definition
Performance measure	Time to complete a draft of an after-action report/improvement plan.
Intent	The systematic observation and assessment of response capabilities is critical to developing and maintaining emergency response readiness. Grantees must demonstrate through the use of after-action reporting the capability to analyze each response action (real or simulated), describe needed improvement, and prepare a plan for making improvements in a minimal amount of time. The report should include how response operations did and did not meet objectives, recommendations for correcting gaps or weaknesses, and a plan for improving response operations.
Target	Mean[a] = 60 calendar days from the end of the exercise or the end of public health emergency response operations as determined by the incident commander.
Measurement specifications	Start time: Date of the day following the end of the exercise or the end of public health emergency response operations as determined by the incident commander. Stop time: Date draft after-action reports/improvement plans was submitted for clearance within the public health agency.

[a]Mean based on all reported after-action reports/improvement plans.

Data Collection and Analysis

A formal data use agreement between the principal investigator and the Outcomes Monitoring and Evaluation Branch of the Division of State and Local Readiness allowed data sharing. The data included assessment and budget information for all grantees (n = 62) for 2005–2009. The assessment data included 52 Microsoft Excel files for the period of 2005–2009 and included entries from all grantees, as well as a codebook. The budget information arrived in five Excel files containing budget figures for all class/object categories for each grantee. Exploring the question structure for consistency and the assessment data set of interest resulted in three measures reported in the same method for two years, 2007 and 2008 (see Table 1). The data reported were essential to answering the research questions:

1. Is participation in drills and exercises a factor for improving times for staff notification, staff assembly, and completion of an after-action report?

$H_0$1: The amount of time needed to complete staff notification, staff assembly, and an after-action report during a real event is not different among public health departments that have participated in drills and exercises and public health departments that have not participated in drills and exercises.

H_a1: The amount of time needed to complete staff notification, staff assembly, and an after-action report during a real event is different among public health departments that have participated in drills and exercises and public health departments that have not participated in drills and exercises.

2. Is the periodicity of participation in drills and exercises a factor for improving times for staff notification, staff assembly, and completion of an after-action report?

$H_0$2: The number of drills and exercises completed within 6 months does not explain differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

H_a2: The number of drills and exercises completed within 6 months explains differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

3. Is the length of time between participation in drills and exercises a factor for improving times for staff notification, staff assembly, and completion of an after-action report?

$H_0$3: The amount of time between drills and exercises, and real events does not explain differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

H_a3: The amount of time between drills and exercises, and real events explains differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

4. Does participation in drills and exercises increase the number of states that meet the target times for staff notification, staff assembly, and completion an after-action report?

$H_0$4: The number of drills and exercises completed within 12 months does not explain differences in the number of states that meet the targeted times to complete staff

assembly, staff notification, and an after-action report during a real event for public health departments.

H_a4: The number of drills and exercises completed within 12 months explains differences in the number of states that meet the targeted times to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

The unit of measure regarding notification and assembly was minutes. The unit of measure for the after-action report was days. Grantees self-reported the data following the protocols outlined in the performance measurement guidance (see Tables 2–4). The idea that practicing would make an improvement in an actual timed response was logical, and it would be expected to see the improvement last over time. The methods and models chosen to analyze the data could help determine the support of the hypotheses.

Protection of Participants

The data provided for this study were the result of timed activities performed by governmental units. Analyzing or attributing specific information to individuals was impossible. The study involved the use of existing data and records, and it evaluated the benefit of a public health practice. Data about individual subjects were not contained within the data set; only data representing the institution were included. Because of the characteristics of the data set, the study was exempt from the basic Health and Human Services Policy for Protection of Human Research Subjects (45 CFR 46.101(b)(4) and 45 CFR 46.101(b)(5). No research was undertaken until Walden University's Institutional Review Board (IRB) approved the research proposal (Walden University IRB approval #10-26-12-0158535).

Summary

This proposal describes a legitimate research gap and supports documentation that affirms a lack of evidence for the problem and appropriate methods to examine data in order to determine if drills and exercises as interventions contribute to improvements in public health responses to emergencies. The final two chapters describe the full data analysis and provide conclusions and recommendations for further research.

Chapter 4: Results

Introduction

The study design was to determine if drills and exercises positively influenced performance of public health responders in emergencies. Although there are many studies describing exercises and drills, as well as the qualitative outcomes of such activities, there is a gap in the information linking drills and exercises to improved performance. Previous chapters included an introduction to the study, a comprehensive literature review, and information pertaining to the research problem and hypotheses.

Purpose and Research Questions

The purpose of the research was to compare the performance of the public health preparedness systems before and after exposure to drills or exercises. The data collected were essential to answering the research questions:

1. Does participation in drills and exercises act as a factor for improving times for staff notification, staff assembly, and completion of an after-action report?
2. Does the periodicity of participation in drills and exercises act as a factor for improving times for staff notification, staff assembly, and completion of an after-action report?
3. Does the length of time between participation in drills and exercises act as a factor for improving times for staff notification, staff assembly, and completion of an after-action report?
4. Does participation in drills and exercises increase the number of states that meet the target times for staff notification, staff assembly, and completion of an after-action report?

PAHPA (2006) codified the use of drills to improve public health response in an emergency. The CDC began requiring the use of established performance measures as a method of measuring preparedness in 2005 (U.S. Department of Health and Human Services & Centers for Disease Control and Prevention, 2005). Analysis of the performance measures before and after drills and exercises in relation to real events could shed light on their utility. This chapter documents research results, including the research setting, demographics of the institutions participating, data collection process, data analysis, trustworthiness of the data, and results.

Research Setting

Public health professionals responsible for a specific cooperative agreement collected the data obtained for this study. The data were associated with the PHEP cooperative agreement administered by the CDC. Sixty-two grantees received funds to prepare for public health emergencies; the data examined included the 50 state public health departments.

State health agencies may be viewed as institutions that support and link people and personal health services (Association of State and Territorial Health Officials, 2011). State and territorial health agencies employ over 100,000 individuals, and the largest source of funding for state health agencies comes from the federal government. A small percent (6%) of health departments mentioned emergency preparedness as a top priority, and 94% of the state health department was responsible for preparedness cooperative agreements from CDC. Preparedness was the topic that received the most attention regarding training that was provided from the state health departments to local health departments. Common services provided by state health departments included vaccine

ordering and management, maternal and child health services, epidemiology and surveillance, regulation, licensing and inspections, and interventions for the prevention or treatment of infectious and chronic diseases (Association of State and Territorial Health Officials, 2011).

All-hazards preparedness and response made up 5% of state health agencies' expenditures in 2009, with a median of $17 million, a minimum of $2.59 million and a maximum of $115 million (Association of State and Territorial Health Officials, 2011). The data examined in this study were from federal fiscal years 2006 and 2007. The funds awarded to states for those years and the preceding and following years appear in Table 5:

Table 5

Funds Available for State, Territorial, and Local Health Departments by Fiscal Year

Year	Amount
2005	$862,777,000
2006	$766,440,000
2007	$896,736,525
2008	$704,867,418

Note. Eligible entities include the 50 United States, Puerto Rico, Virgin Islands, Northern Mariana Islands, Palau, Micronesia, Marshall Islands, Guam, American Samoa, New York City, Chicago, and Los Angeles County

Demographics

Data collected from the participants included information relating to costs and number of personnel involved in emergency preparedness activities per state. There were no identifiers in the data that indicated which staff members actually participated in the

drill, exercise, or real event recorded. Table 6 identifies the amount spent on personnel and the number of full-time equivalent (FTEs) employed for each of the years examined and the prior and following years.

Table 6

Funds Available for Staffing and Number of FTEs per Fiscal Year

Year	Personnel Class/Object	Number
2005	$209,032,155	3,601
2006	$207,294,871	3,668
2007	$182,558,335	3,480
2008	$201,961,243	3,731

Categorizing state health agencies by size (small, medium, and large) is common based on the population served. Table 7 identifies the number of states of each size, the parameters of each grouping, and the name of the state as of 2009 (Association of State and Territorial Health Officials, 2011).

Table 7

Size Categories of State Health Departments

Size	Number	Parameter	Names
Small	17	Under 2,750,000	Alaska, Delaware, Hawaii, Idaho, Maine, Montana, Nebraska, Nevada, New Hampshire, New Mexico, North Dakota, Rhode Island, South Dakota, Vermont, Washington DC, West Virginia, and Wyoming
Medium	16	2,750,001 to 6,250,000	Alabama, Arkansas, Colorado, Connecticut, Iowa, Kansas, Kentucky, Louisiana, Maryland, Minnesota, Mississippi, Oklahoma, Oregon, South Carolina, Utah, and Wisconsin
Large	17	6,250,001 and over	Arizona, California, Florida, Georgia, Illinois, Indiana, Massachusetts, Michigan, New York, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Texas, Virginia, and Washington

As an important point of reference, about half of the states have decentralized governance structures; that is, they do not have regulatory or legal authority over local or city governmental public health agencies (Association of State and Territorial Health Officials, 2011). Since 2003 the Trust for America's Health has filed public reports grading a state's ability to protect the public's health in the event of a bioterrorism event. The Trust for America's Health's grading schemata included a variety of indicators related to the preparedness level in each state. Each state earned a score based on the total number of indicators met, with a possible high score of 10. Regional groupings assist the

federal government when providing technical assistance to public health departments

Table 8 outlines the regional structure.

Table 8

Public Health Regions as Designated by the DHHS

Region	States
1. Boston	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont
2. New York	New Jersey, New York, Puerto Rico, and Virgin Islands
3. Philadelphia	Delaware, District of Columbia, Maryland, Pennsylvania, Virginia, and West Virginia
4. Atlanta	Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, and Tennessee
5. Chicago	Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin
6. Dallas	Arkansas, Louisiana, New Mexico, Oklahoma, and Texas
7. Kansas City	Iowa, Kansas, Missouri, and Nebraska
8. Denver	Colorado, Montana, North Dakota, South Dakota, Utah, and Wyoming
9. San Francisco	Arizona, California, Hawaii, Nevada, American Samoa, Commonwealth of the Northern Mariana Islands, Federated States of Micronesia, Guam, Marshall Islands, and Republic of Palau
10. Seattle	Alaska, Idaho, Oregon, and Washington

As the indicators changed over time, it is unfair to compare scores from year to year. Figure 1 shows the scoring distribution for 2005, 2006, 2007, and 2008.

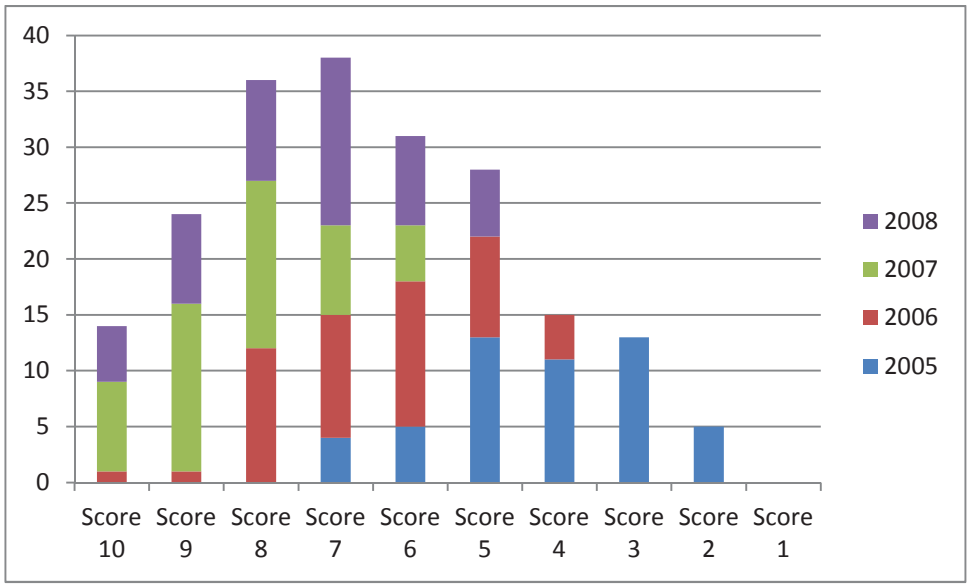

Figure 1. Preparedness indicator.

Note how the number of indicators met increases from 2005 to 2008, implying that states are becoming more prepared (Hearne, Segal, & Earls, 2005; Jeffrey Levi, Segal, Gadola, Juliano, & Speulda, 2006; Jeffrey Levi, Vinter, & Segal, 2007; Jeffrey Levi, Vinter, St. Laurent, & Segal, 2008).

Data Collection

The data under study were archival and were collected as part of ongoing technical assistance and monitoring of the PHEP cooperative agreement administered from 2006 to 2008. Each year provided new opportunities to improve on the requirements and data collection instructions for grantees. The data specific to the performance measures were submitted midyear and at the end of the budget year for each year reviewed. In 2006 data were submitted related to questions regarding communication,

crisis communication, detection and surveillance, information and technology, and laboratory services. The grantees were instructed to designate one person with the responsibility to submit required data elements electronically (Division of State and Local Readiness, 2006). The guidance received by states also indicated that the CDC would provide more information regarding evaluation based on drills and exercises and real events.

The data were collected and stored at the CDC. After approval of a data-sharing agreement, the 5 years of data were delivered electronically. The entire data set contained 56 separate files captured in spreadsheets labeled to indicate the budget period from which the data were reported and the type of measure the data characterized. Five of the files captured the funding allocation data for each of the budget periods under study and classified the funding information into class and/or object categories such as personnel, contracts, travel, and supplies. Each file consisted of three tabs: Two were associated with the responses and coding of the questions, and one served as a data codebook. The data included the grantee's name, a unique identifier for each grantee, and columns representing the actual data turned in as a response to the performance measure. The data on the second tab represented the actual data turned in by the grantee. The third tab contained the codebook. Table 9 displays the number of files available for each budget year.

Table 9

Number of Data Files per Budget Year

Budget period	Number of files
BP 6	13
BP 7	11
BP 8	11
BP 9	9
BP 10	12

Data Analysis

To compile data for budget periods 7 and 8, and the first submission of budget period 9, spreadsheets were created for each of the performance measures of note. The requirement to submit data was universal for all grantees, resulting in 64 respondents in each file. Edited data files eventually contained only the data from the 50 states and Washington, D.C. Elements from each data set referring to the research hypotheses were isolated and extracted to construct a chart for comparison. Charting the common elements allowed for the removal of questions not addressed consistently over multiple budget periods. Data from budget period 9 were inconsistent with those from previous years; thus they were rejected for the analysis. Table 10 displays the common elements relating to staff notification. Table 11 represents the common data elements related to staff assembly. Table 12 represents the common data elements related to the after-action report.

Table 10

Common Data Elements for Staff Notification by Budget Period

	BP 8 (EY)	BP 8 (MY)	BP 7 (EY)	BP 7 (MY)
How many times did the agency notify staff to fill all eight ICS core functional roles due to a drill, exercise, or real incident from 03/01/2008 to 08/09/2008?	BP8_361_5847_6B	BP8_274_5357_6B	BP7_248_4963_6B	BP7_213_4538_6B
Click the button to calculate and display the time to notify all primary staff (and secondary or tertiary staff as needed) with public health agency ICS functional responsibilities that the public health agency's EOC is being activated.	BP8_361_5854_6B	BP8_274_5364_6B	BP7_248_4970_6B	BP7_213_4543_6B

Table 11

Common Data Elements for Staff Assembly by Budget Period

	BP 8 (EY)	BP 8 (MY)	BP 7 (EY)	BP 7 (MY)
How many times was the public health agency's EOC activated due to a drill, exercise, or real incident from 03/01/2007 to 08/30/2007?	BP8_362_5859_6C	BP8_275_5369_6C	BP7_249_4975_6C	BP7_214_4548_6C
Click the button to calculate and display the time for primary staff (and secondary or tertiary staff as needed) with public health agency ICS functional responsibilities to report for duty at the public health agency s EOC.	BP8_362_5866_6C	BP8_275_5376_6C	BP7_249_4981_6C	BP7_214_4554_6C

Table 12

Common Data Elements for After-Action Reports by Budget Period

	BP 8 (EY)	BP 8 (MY)	BP 7 (EY)	BP 7 (MY)
How many after-action report/improvement plans for exercises or real incidents did the public health agency submit for clearance from 03/01/2007 to 08/30/2007?	BP8_363_5868_9A	BP8_276_5380_9A	BP7_250_4985_9A	BP7_215_4558_9A
Click the button to calculate and display the time to complete a draft of an after-action report/improvement plan.	BP8_363_5879_9A	BP8_276_5389_9A	BP7_250_4994_9A	BP7_215_4567_9A

The data were sorted into tables for each participant providing information about the event (real or drill) and the amount of time it took to complete the action represented by the measure. The spreadsheet data pertaining to the three questions examined (6B, 6C, and 9A) represented thousands of data points. For this analysis, there were three questions of particular importance for analysis: Was the event a drill, exercise, or real event? What time did the event start? What time did the event end? Tables for the three issues displayed summaries by question and year. Tables 13, 14, 15, and 16 show the number of events reviewed for questions 6B, 6C, and 9A, and a summary of all questions.

Figure 2 summarizes the number of events for the four periods and three performance measures reviewed. Tables 17, 18, and 19 show the times recorded for Questions 6B, 6C, and 9A. Table 20 shows the fields used in new data matrices.

Table 13

Summary Statistics for Staff Notification (6B) by Budget Period—All Participants

	Real events	Drills	Total
Midyear, BP 7	24	88	112
Year end, BP 7	28	103	131
Midyear, BP 8	24	81	105
Year end, BP 8	27	87	103
Total events	103	359	462

Table 14

Summary Statistics for Staff Assembly (6C) by Budget Period—All Participants

	Real events	Drills	Total
Midyear, BP 7	29	48	77
Year end, BP 7	25	65	90
Midyear, BP 8	22	41	63
Year end, BP 8	25	61	86
Total events	101	215	316

Table 15

Summary Statistics for After-Action Reports (9A) by Budget Period—All Participants

	Real events	Drills	Total
Midyear, BP 7	25	157	182
Year end, BP 7	12	159	171
Midyear, BP 8	23	80	103
Year end, BP 8	15	96	111
Total events	75	492	567

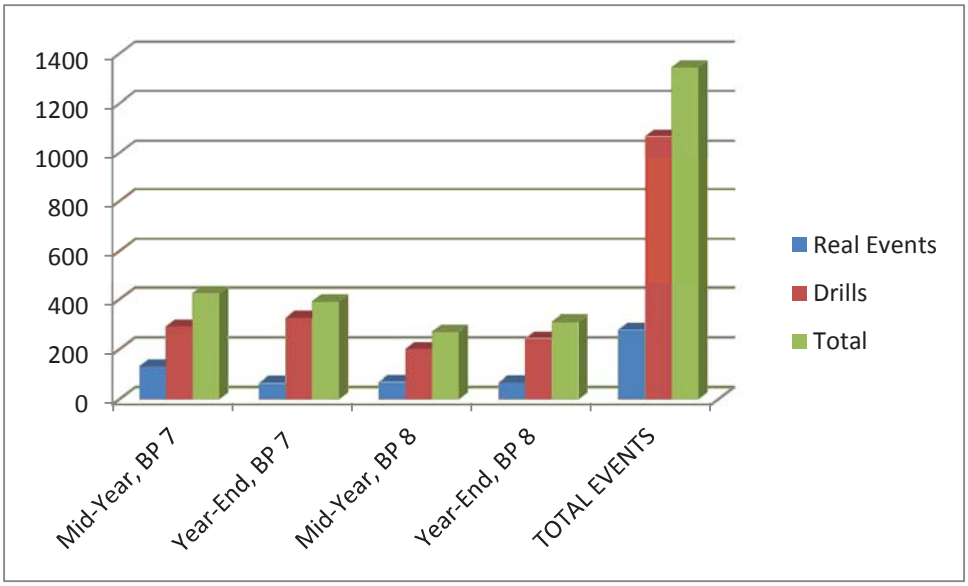

Figure 2. Number of events.

Table 16

Summary Statistics for 6B, 6C, and 9A by Budget Period—All Participants

	Real events	Drills	Total
Midyear, BP 7	78	293	424
Year end, BP 7	65	327	392
Midyear, BP 8	69	202	271
Year end, BP 8	67	244	311
Total events	279	1066	1345

Table 17

Summary Statistics for Time Needed for Staff Notification, in Minutes

	Events	Mean	Standard deviation
Real events	103	194.7282	1390.8738
Drills	359	77.3398	795.7523
Total events	462	103.5108	931.1050

Table 18

Summary Statistics for Time Needed for Staff Assembly, in Minutes

	Events	Mean	Standard deviation
Real events	101	71.7426	222.9332
Drills	215	82.1442	237.0054
Total events	316	78.8196	232.2975

Table 19

Summary Statistics for Time Needed for After-Action Reports, in Days

	Events	Mean	Standard deviation
Real events	75	31.2533	33.6520
Drills	492	36.3740	28.2493
Total events	567	35.6966	29.0405

Table 20

Data Elements Used for Each Performance Measurement

6B	6C	9A
Grantee	Grantee	Grantee
GranteeID	GranteeID	GranteeID
Metric	Metric	Metric
Instance	Instance	Instance
EventType	EventType	EventType
IsBusinessHours	Start	JurisdictionCount
IsUnannounced	End	Start
Start		End
End		

The results of the calculations looking at the means and variances to assure they meet the assumptions for a Poisson regression appear in Table 21. They affirm that a Poisson model would be the appropriate model of choice.

Table 21

Means and Variance of the Distribution of Hours or Months with Outliers Removed

Measure	Mean	Variance
6B	0.61	0.58
6C	0.77	0.93
9A	1.26	1.00

Results

Discussion of the hypotheses under study is based on the data analysis performed for each.

$H_0$1: The amount of time needed to complete staff notification, staff assembly, and an after-action report during a real event is not different among public health departments that have participated in drills and exercises and public health departments that have not participated in drills and exercises.

Analysis of hypothesis $H_0$1 cannot be performed. There was not enough data to evaluate the performance times of agencies that did not participate in a drill or exercise against those that did participate in a drill or exercise. In fact, there was no single state that reported only drills or only real events in the entire data set.

$H_0$2: The number of drills and exercises completed within 6 months does not explain differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

A Poisson regression conducted for measures 6B, 6C, and 9A determined that the null hypothesis for question 2 would be rejected. The predictors for the times were the trailing drill and trailing event. Table 22 indicates the results of the regression for Measure 6B—time to notify.

Table 22

Poisson Regression for Measure 6B—Time to Notify Staff, Hypothesis 2

Variable	Estimate	Std. error	*z* score	*p* value
Intercept	0.79	0.12	6.5	~0.0
6-month trailing drill count	0.097	0.056	1.7	0.084
6-month trailing event count	-0.087	0.030	-2.91	0.0036[a]
Outside of business hours	-0.15	0.18	-0.84	0.40
During business hours	-0.079	0.18	-0.43	0.66

[a]significant

The results indicated the time needed to complete notification was negatively related to the number of real events in the last 6 months but positively related to the number of drills and exercises over the last 6 months. Health department staff was quicker to react to the notification if there were more real emergencies in the recent past and slower to react to the notification if there were more drills and exercises in the recent

past. Every drill or exercise that followed a drill or exercise over a 6-month period resulted in an increase in the time to complete staff notification by nearly 6 minutes. A relationship between the number of drills completed in a 6-month period and the time it takes to notify the staff likely appears to exist in the data.

For every real event completed over a 6-month period, the time to complete notification for the next real event decreased by about five minutes. The results were significant ($p < 0.01$) and suggested one way to reduce the time needed to notify staff of an event was to avoid overdrilling. Whether the call to notify staff came during or outside of business hours did not have a predictive relationship to the time it took to complete the notifications.

The Poisson regression for Measure 6C—time to assemble staff—yielded different results. In this case, the information relating to the drill or event occurring during business hours was not included. Table 23 displays the result of the regression.

Table 23

Poisson Regression for Measure 6C—Time to Assemble Staff, Hypothesis 2

Variable	Estimate	Std. error	*z* score	*p* value
Intercept	0.94	0.11	8.9	~0.0
6-month trailing drill count	-0.043	0.067	-0.63	0.53
6-month trailing event count	-0.068	0.037	-1.8	0.067

The results indicated the time needed to assemble staff related negatively to the number of real events in the last 6 months. Health department staff was quicker to assemble if there were more real emergencies in the recent past. For every real event completed over a 6-month period, the time required for staff to assemble for the next real event decreased about 4 minutes.

The same regression was performed for Measure 9A—time to complete an after-action report. In this case, the information relating to the drill or event held during business hours was not included. Table 24 displays the result of the regression.

Table 24

Poisson Regression for Measure 9A: Time to Complete an After-Action Report, Hypothesis 2

Variable	Estimate	Std. error	*z* score	*p* value
Intercept	1.23	0.20	6.2	~0.0
6-month trailing drill count	-0.019	0.049	-0.38	0.71
6-month trailing event count	-0.28	0.076	-3.7	~0.00[a]

[a]significant

The results indicate the time needed to complete the after-action report related negatively to the number of real events in the last 6 months. Health department staffs were quicker to clear an after-action report if there were more real emergencies in the recent past and slower to clear an after-action report if there were more drills and exercises in the recent past. For every real event completed over a 6-month period, the

time to complete an after-action report for the next real event decreased by about 1 week. The results were significant ($p < 0.01$). The results suggested that one way to reduce the time needed to complete an after-action report for an event was to avoid overdrilling.

For the second hypothesis, there were two significant findings regarding trailing events. The hypothesis tested if the number of drills, exercises, and/or real events in a 6-month period made a difference in the performance of health departments. One interpretation could be that drills and exercises reduce gains in performance and that real events increase gains in performance. Real events following real events (and not drills) showed gains in performance. Drills and exercises following drills and exercises (and not real events) showed lapses in performance.

The third hypothesis was related to the relationship between performance in events, drills, and exercises and the amount of time that elapsed between each.

> $H_0$3: The amount of time between drills and exercises and real events does not explain differences in time needed to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

A Poisson regression conducted for measures 6B, 6C, and 9A determined the validity of the hypothesis. The predictor used was time since the last event. Tables 25, 26, and 27 display the results for Measure 6B—time to notify, 6C—time to assemble, and 9A—time to complete an after-action report.

In this analysis, the data failed to reject the null hypothesis as none of the variables of interest had any relationship to the time to notify staff, time to assemble staff, or time to complete an after-action report. In order for data to be included in the analysis,

entities had to have had at least two real events, which eliminated many of the entities from the analysis. A larger dataset might be the key

Table 25

Poisson Regression for Measure 6B—Time to Notify Staff, Hypothesis 3

Variable	Estimate	Std. error	*z* score	*p* value
Intercept	.70	0.17	4.2	~0.0
Time since last event	-0.0024	0.0020	-1.2	0.22
Outside business hours	0.080	0.20	.40	0.69
During business hours	0.20	0.21	0.94	0.35

Table 26

Poisson Regression for Measure 6C—Time to Assemble Staff, Hypothesis 3

Variable	Estimate	Std. error	*z* score	*p* value
Intercept	.81	0.13	6.0	~0.0
Time since last event	-0.00045	0.0018	-0.25	0.81

Table 27

Poisson Regression for Measure 9A—Time Complete an After-Action Report, Hypothesis 3

Variable	Estimate	Std. error	*z* score	*p* value
Intercept	.68	0.17	4.0	~0.0
Time since last event	0.0059	0.0033	1.8	0.27

to identifying relationships. Interpretations of the results are not advisable due to the small dataset size and the small number of entities that had more than one real event.

The fourth hypothesis examined the relationship between meeting the performance target for events, drills, and exercises, and the total number of drills and events in a 12-month period.

> $H_0$4: The number of drills and exercises completed within 12 months does not explain differences in the number of states that meet the targeted times to complete staff assembly, staff notification, and an after-action report during a real event for public health departments.

To determine the validity of this hypothesis, a different model was used—a logistic regression. A logistic regression best predicts the value of a variable (Y) for each value of another variable (X). This model was a good fit because the hypothesis looked at the relationship between drills, real events, and their recentness, and whether those factors influenced the ability of an entity to meet the target performance measure. Entities could only have one of two responses to meeting the target—yes or no. The data (times) for performance measures 6B, 6C, and 9A were transformed into binary success variables, with only two choices for the outcome—did they meet the target (yes) or did they miss the target (no).The equation representing a logistic regression is

$$\ln [Y/ (1-Y)] = b_0 + b_1 X_i$$

where $[Y/ (1-Y)]$ represents the log of the odds ratio of the dependent variable, and $b_0 + b_1 X_i$ is the linear relationship of X to the predicted logit (Association of State and Territorial Health Officials, 2011).

Logistic regression was used to determine if there was a likelihood that drills made it more probable for state health departments to meet the targeted times for notification and assembly of staff, as well as the development and clearance of after-action reports. Tables 28, 29, and 30 display the results.

Table 28

Logistic Regression for Measure 6B—Time to Notify Staff is 60 Minutes or Less

Variable	Log odds ratio	Std. error	*z* score	*p* value
Intercept	0.25	0.54	0.46	0.64
Outside of business hours	1.2	0.56	2.19	0.029[a]
During business hours	1.5	0.67	2.3	0.022 [a]
12-month trailing event count	0.28	0.14	1.9	0.051
12-month trailing drill count	-0.0078	0.12	-0.07	0.95
Time since last event	-0.0034	0.0024	-1.4	0.16

[a]significant

The results indicated that an institution was more likely to notify staff within 1 hour ($p < .05$) and more likely to meet the target if the notification occurred during business hours ($p < .05$). The results related positively to the number of real events during the last 12 months. There were no other significant variables in the analysis, including the number of drills performed in the last 12 months.

Table 29 displays the results of a logistic regression to measure the likelihood that staff could assemble in 2.5 hours or less.

Table 29

Logistic Regression for Measure 6C—Time to Assemble Staff is 150 Minutes or Less

Variable	Log odds ratio	Std. error	*z* score	*p* value
Intercept	1.89	0.88	2.13	0.03
12-month trailing event count	0.13	0.23	0.6	0.56
12-month trailing drill count	0.43	0.38	1.1	1.14
Time since last event	0.002	0.0004	0.54	0.59

This regression indicated that the likelihood of an institution assembling staff within 150 minutes or less did not relate to the variables used in the analysis, and there was no significance found in the statistics, and the data failed to reject the null hypothesis. Drills or real events have no predictive value to meeting the established targets.

Table 30 displays the results of a logistic regression performed using data from the last measure—the time needed to submit and clear an after-action report. The target for the report was 2 months after the completion of the event or drill.

Table 30

Logistic Regression for Measure 9A—Time to Complete After-Action Report Is 60 Days or Less

Variable	Log odds ratio	Std. error	*z* score	*p* value
Intercept	1.1	0.35	3.1	0.0019
12-month trailing event count	0.41	0.25	1.6	0.10
12-month trailing drill count	0.015	0.071	0.21	0.84

The likelihood that an institution would file an after-action report within 2 months was positively related to the number of real emergencies over the last 12 months. Other variables were not significant in the analysis.

Summary

Public health departments have been responding to emergencies as part of normal business for years. After terrorists destroyed the World Trade Center and mailed biological agents to businesses and individuals, the federal government invested billions of dollars to improve the ability of state health departments to respond to emergencies (Nelson, Willis, Chan, Shelton, & Parker, 2011). With the passage of PAHPA, Congress made mandatory drills and exercises at the state level to improve responsiveness ("The Pandemic and All-Hazards Preparedness Act," 2006). Performance measures were developed and state health departments were required to report their progress to the DHHS twice a year.

State, local, and territorial public health departments submitted self-reported performance data over a 5-year period if they received emergency preparedness funding from the CDC. The data included information about typical activities performed during

public health emergencies, such as notifying staff of an emergency, assembling staff at the proper location, and creating and clearing an after-action report. Data examined from 50 states and the District of Columbia produced performance data from years 7 and 8. Examination found three consistent questions in the 2 years examined, and those data became the data set to determine if drills and exercises contributed to increases in performance for health departments.

The research questions generated four hypotheses. The first hypothesis was not tested, as there were no groups that were without a drill or exercise during budget years 7 and 8. Testing the second and third hypotheses by Poisson analysis indicated drills and exercises did not influence the performance of health departments, but there was a relationship between the number of real events and performance. The last hypothesis, analyzed using a logistic regression, examined if drills and exercises predicted the likelihood that health departments were able to meet the time targets established. Again, there was no relationship between drills and exercises and the likelihood of meeting the targets, but there was a positive relationship between real events and the likelihood of meeting the targets. In general, real events had more influence on performance than drills and exercises. Data showed that response times did not decrease with repeated practice, but they did decrease with repeated real events.

Chapter 5: Summary and Conclusions

Introduction

The study was designed to determine if drills and exercises positively influenced performance of public health responders in emergencies. Although there are many studies describing exercises and drills, as well as the qualitative outcomes of such activities, there is a gap in the information linking drills and exercises to improved performance. Previous chapters included an introduction to the study, a comprehensive literature review, information pertaining to the research problem and hypotheses, and the results.

The research questions generated four hypotheses. The first hypothesis was untested, as there were no groups participating in the study that were without a drill or exercise during the years under study. Tested by Poisson analysis, the second and third hypotheses indicated that the number of drills and exercises did not influence the performance, but there was a relationship between the number of real events and performance of health departments. The last hypothesis, analyzed using logistic regression, sought to determine if drills and exercises predicted the likelihood that health departments were able to meet the time targets established. Again, no predictive relationship existed between the performance of drills and exercises and the likelihood of meeting established targets, and there was a positive relationship between real events and the likelihood of meeting the targets.

This chapter provides an interpretation of the findings, a discussion of the limitations of the study, recommendations for future study, and an outline of how the results have implications for social change.

Interpretation of the Findings

The literature pertaining to public health and the use of drills and exercises to improve performance has been theoretical in nature, and true quantitative studies are lacking (Biddinger et al., 2008, p. 8). There was little quantitative evidence supporting the use of drills or exercises for performance improvement. The literature showed little progress describing the development of adequate measures of performance during or after drills and exercises (Savoia, Massin-Short, et al., 2009). The literature also lacked new contributions to the field, with no true experiments documented and a small number of quasiexperimental or mathematical model studies presented. In another study, primary data collection was present in only 11.3% of the 823 articles reviewed (Yeager et al., 2010). Although exercises and drills were the most frequent topic of the articles, these represented only 12.3% of the entire sample. Yeager concluded that the body of literature was overwhelmingly subjective.

This study extended the knowledge base by applying actual performance data from state health departments against a known set of performance targets. The use of quantitative data has been scarce in the literature and the use of an original design and analysis regarding the use of drills and exercises to enhance performance lacks in the literature. Additionally, using drills and exercises to improve preparedness had not been tested in a quantitative manner prior to the Congressional mandate to do so. The results of this study could begin to create additional questions that will spur further research about the use of drills and exercises to improve performance, as well as the use of science and research to influence decision making at the Congressional level.

Based on the Poisson analysis and logistic regressions performed on the self-reported data collected during budget years 7 and 8 of the Public Health Emergency Preparedness grants, the use of drills and exercises did not directly influence the performance of public health departments in emergencies. The quantitative study confirmed that the best predictor of performance in a real emergency was the number of real emergencies a health department had responded to in the past. The qualitative literature confirmed that when people practice together, they feel more prepared to handle emergencies, feel stronger as a team, and become more familiar with the routines and procedures needed to respond (Ablah et al., 2008; Beaton et al., 2004; Biddinger et al., 2010; Fowkes et al., 2010; High et al., 2008; Johnson et al., 2009). The quantitative literature has been silent on the actual gains in performance during real events that might be due to drilling and exercising. This was the first study to indicate the quantitative gains experienced by using drills and exercises to improve performance were minimal and not confirmed statistically.

Although there were no specific examples that outlined the use of deliberate practice in emergency response settings, Salas et al. (2009) suggested the theory of deliberate practice could be used to plan, evaluate, and measure performance in emergency response settings. Deliberate practice theory details how expertise develops over time with clear, concise, deliberate practice sessions. Deliberate practice increases performance in many domains but has not been explicitly linked to the use of drills, exercises, and SBT in the domain of public health emergency response. Although public health agencies have been participating in exercises with the notion that practice will improve performance, very few have been subject to specific performance measurement

during or after the drills. The results of this study could bring attention to the need to extend the body of knowledge regarding the theoretical underpinnings of the construction and purpose of drills and exercises and if they are a way to improve public health performance.

Limitations of the Study

There were limitations to the study. First, a repeated activity may cause the learner or organization to change over time due to repetition (Frankfort-Nachmias & Nachmias, 2008, p. 125). The acts of notifying staff, assembling staff, and providing a cleared after-action report could become easier over time as the health department becomes more familiar with the routines involved in performing these tasks. In the data collected for this study, it must be assumed that the organizations did change over time (e.g., made personnel changes); however, no statistical tests were performed on this particular data set to determine if the organizations simply improved with multiple practice sessions. One conclusion, however, was that drilling or exercising was not predictive of the ability of the organization to meet the established performance goals.

Another limitation identified was regression toward the mean. The phenomena of regression to the mean is likely targeted to individuals rather than groups, and when examined in a group setting, the data are typically from individuals who have been selected and placed in groups (Barnett et al., 2005). There is very little data supporting the theory that organizations can regress toward the mean.

A threat to internal validity existed in this study because data were self-reported by the health department staff. This limitation was valid for the study; however, steps taken to eliminate the threat to the extent possible were documented. For instance, data

were rejected if there were inconsistencies across the budget year fields and the intrayear reporting dates. Data were also rejected if an organization had missing or incomplete reporting. Lastly, data were rejected if there were miscalculations of the interval between start and end dates and times. The result of data cleaning was that the data set was much smaller, which presented a different problem during analysis: Grouping the small amount of data remaining for analysis might have introduced further dependencies that would prohibit looking at time between measurements as variable.

Maturation of the field introduced another limitation. This was still a valid concern, although because the CDC requires the entities to conduct drills and exercises, it would be hard to prevent this from happening. The research questions were not designed to determine if the field improved overall but to determine if the action of drilling and exercising improved performance in a real event. More research is needed to measure and provide insight into maturation of the field due to the aggregate performance of drills, exercises, and real events.

On original intent, a goal was to examine the individual state health departments to determine if stratification by region, size, or budget of the health department could determine if any other variables influenced the role of drills and exercises in relation to real events. Due to the data cleaning needed to overcome the inconsistencies of the self-reported data, the results were reported for all 50 entities as a whole; no individual state's data could be singled out. Generalization of results beyond the public health community could not be done or discussed at the micro (e.g., specific site) level because the analysis was done looking at all entities as a single group. This remains a limitation of the study as articulated in earlier chapters.

The last limitation identified remains—the validity of the measurement instrument. The targets identified for each of the measures were construct referenced. The tool and the targets were not scientifically validated, and as most researchers agreed, more work is needed in this area (Nelson et al., 2008). In this study, data collected conformed to agreed-upon processes and protocols and the collective agreement from state health department officials regarding the targets. There were no studies that validated that the targets set were truly achievable or that there might be other thresholds that should be considered when seeking improvement in public health performance during emergencies.

Recommendations

The study lent itself to several recommendations for further study. One of the strengths of this study was the use of self-reported performance data to determine the impact of drills and exercises on performance in real events. This was one of the few qualitative studies available for review that extended the body of knowledge of public health performance in emergencies. It is suggested that additional quantitative studies be designed and conducted using actual performance data from the state health departments. Data exist for further research. Qualitative studies, combined with additional quantitative studies, might determine if it is appropriate to describe the role of drills and exercises as a method to provide a certain level of comfort among first responders, rather than as a method to improve performance.

A limitation of the study was the measurement instrument. The measurement instrument has changed over time. The questions were inconsistent and varied from year to year, making it difficult to compare responses across funding years. In addition, the

data pertaining to the three performance measures in the study were a fraction of the data collected and made available to the researcher. Further refinement of the information needed to answer specific research questions (e.g., if drills and exercises contributed to gains in performance) should be undertaken. Scholars in the field may want to determine the research questions that are most important to help define successful performance and design a data collection instrument with the research methodologies in mind.

The targets assigned to the performance measures need validation. Conducting further research determining if the targets as established are key to timely and accurate responses is a priority. For instance, research could determine if the outcome for state A is better than state B in an emergency because state A met a time threshold for assembling staff.

Another limitation involved the study design. The initial study design included a traditional time series analysis. Based on the instructions provided to the state health departments, the entities were free to choose the interval between drills and exercises, and the real events were not scheduled. Because of the lack of regularly spaced interventions (e.g., drills and exercises), as well as differences in the specific type of event (e.g., table top exercise or functional drill), a traditional time series analysis would have been inappropriate. If the CDC determined there was value in exploring the impact of drills and exercises on performance in real events, it could require regularly spaced drills and exercises, as well as mandate the types of exercises to create a standardized schedule of events for all grantees. Data collected in that manner would lend themselves to a more traditional time series analysis.

Lastly, a limitation emerged regarding the theoretical framework for drills and exercises. Continuing research is suggested to determine the best theory from which to design drills and exercises. It was not clear that drills and exercises were designed with any theoretical underpinnings, such as deliberate practice theory or ELT. The field needs more work to uncover the appropriate theory behind the use of drills and exercises for improved performance for governmental public health agencies.

Recommendations based on the findings are appropriate. The results indicated that the number of real events in a 6-month period was a better indicator of how a health department might perform during a real event. Drills and exercises had no impact on performance in a real event. Research to examine the differences in behavior, action, and results between drills and real events might inform the field. Performance actually deteriorated from drill to drill, but from real event to real event, performance gains were significant. Determining the underlying factors to gains in performance, and beginning to replicate those factors in drills and exercises would inform the field.

The results also indicated that the time between drills and exercises made no difference to the performance during a real event. Developing additional research to determine if there is a practice threshold, and modify requirements of state health departments based on the results, is appropriate. During the period under study, states were to report on at least six drills, exercises, or real events in each 6-month period. Based on the data analysis, the periodicity between drills and exercises did not influence the actual performance in a real event. Perhaps there is a threshold that can be determined that may be more or less than the current requirement.

The last recommendation focuses on the target times. In this study, the time to notify staff and to complete an after-action report within the specified targets related positively to the number of real events that had occurred. In this study, drills and exercises did not predict the ability of the entities to meet the targets. More research is needed to determine if drills and exercises are appropriate mechanisms to improve performance.

Implications for Social Change

As stated previously, although drills and exercises have been mandated by law as a method to improve performance during public health emergencies, the nation does not know if drills and exercises contribute to changes in performance. Based on the three performance measures examined for budget years 7 and 8, drills and exercises do not produce gains in performance for staff notification and assembly and the production of an after-action report. The only significant factor in performance improvement for public health departments in these three measures was the number of real events they responded to over time. The body of knowledge indicated that drills and exercises produced higher levels of confidence within first responders, and allowed first responders to be more familiar with procedures and polices (Ablah et al., 2008; Beaton et al., 2004; Biddinger et al., 2010; Fowkes et al., 2010; High et al., 2008; Johnson et al., 2009). Policy makers could use the information to reframe the role of drills and exercises in overall emergency preparedness and remove the emphasis from improved performance to improved coordination. This might allow leaders to direct some level of preparedness dollars allocated for drills and exercises toward evidence-based policies and practices that would focus on researching activities that improve performance rather than improve

coordination. In addition, policy makers could remove the legislated mandate to drill and exercise, thus freeing up funds for generalized public health needs, such as laboratory improvements, electronic medical records, and reducing health care–associated infections. Health departments would continue to respond to emergencies and practice with other response agencies but would also be able to improve overall services to their constituents by making investments in standard public health infrastructure.

Public health departments, as well as the DHHS and the Department of Homeland Security, have placed emphasis on exercises as a way to improve performance. The results of this study did not support that conclusion. There continues to be a research gap to discover what really does improve performance of public health agencies other than real events. The governmental bodies whose position is that drills and exercises improve public health performance may need to reframe the intent of drills and consider other methods with strong theoretical underpinnings to begin to assess public health performance. Refocusing resources directed to drills and exercises to yet-unidentified activities that truly improve performance, or to larger needs of the community, might be considered. The CDC may want to re-examine its requirements for state health departments to conduct drills and exercises as a way to improve performance. There is value in drilling and exercising, but based on the three performance measures under study, the value does not translate into gains in performance.

Public health practitioners need to decide if they want to continue to pursue a rigorous schedule of drilling. The cost to conduct drills might outweigh the coordination benefits. There are also costs when data show that an intervention is ineffective, yet entities are required to continue to conduct that intervention. This is the case with drills

and exercises when expecting performance gains in staff notification and assembly and after-action reporting. The data show the drills and exercises did not lead to performance gains, and actually, performance deteriorated from drill to drill. Performance increased only from real event to real event.

Conclusion

Drills and exercises have been used by the military to evaluate options during battle (Crichton et al., 2000). Traditional first responders have practiced through simulation as a way to keep their skills sharp (Perry, 2004). After the destruction of the World Trade Center and subsequent anthrax attacks, experts in public health felt that the use of drills and exercises would improve the timeliness of responses to public health emergencies (Dausey et al., 2007; High et al., 2008). Congress agreed and passed PAHPA, authorizing the CDC to provide funds to state health departments to improve their timeliness to public health emergencies and mandating the use of drills and exercises as a method to improve performance (Lurie et al., 2006).

The literature contained little information regarding the role of drills and exercises to improve performance. Four research questions were developed to understand if drills and exercises improved public health response in emergencies, focusing on state health departments that received public health emergency preparedness funds from the CDC. States were required to submit data to the CDC twice a year regarding their performance on certain aspects of emergency response. Data were obtained through a data-sharing agreement with the CDC for data encompassing funding available from 2006 through 2010. Data from 2006 to 2008 were examined and three performance measures were extracted from the data for the analysis: time to notify staff, time to assemble staff, and

time to submit and clear an after-action report. Four hypotheses were generated, and three could be tested using a Poisson analysis and a logistic regression.

The results indicated that the number of real events in the preceding 6 months significantly shortened the total time to complete staff notification, and a relationship between the number of drills and exercises and the total time to complete staff notification seems to exist in the data, albeit resulting in longer times to complete staff notification. There was a negative relationship between time to assemble staff related to the number of real events in the preceding 6 months, meaning more real events in the past reduced the time to assemble. The results indicated that the number of real events in the preceding 6 months significantly shortened the total time to complete and clear after-action reports. Drills and exercises completed in the prior 6 months did not affect performance, and results indicated performance actually suffered with drills and exercises instead of improved with real events. The amount of time between events (drills or real events) showed no relationship to performance. There were no indications from this analysis that there are retention issues if longer periods lapse between events.

The number of real events was significantly related to the likelihood of being able to complete staff notification within the target of 60 minutes, and although occurrence both within and outside of business hours was significantly related to prior real events, it was more likely that the notification would be completed within the time frame if the notification came within business hours. The time to complete staff assembly within the targets had no predictive relationship to the number of drills or real events, and the time to submit and clear an after-action report within the established targets was positively related to the number of real events.

Based on self-reported data from 2006 to 2008 submitted by state health departments receiving CDC public health preparedness funding, drills and exercises did not have an effect on performance for time to notify staff, time to assemble staff, and time to submit and clear an after-action report. In fact, overdrilling seems to erode performance as indicated by Poisson analysis and logistic regression of specific data points. Policy makers and leaders in the public health preparedness field should consider this information when requiring drills and exercises as part of performance improvement within grant activities.

References

Ablah, E., Nickels, D., Hodle, A., & Wolfe, D. J. (2008). "Public health investigation": Focus group study of a regional infectious disease exercise. *Public Health Nursing, 25*(6), 546-553. doi: 10.1111/j.1525-1446.2008.00742.x

Abramson, D. M., Morse, S. S., Garrett, A. L., & Redlener, I. (2007). Public health disaster research: Surveying the field, defining its future. *Disaster Medicine and Public Health Preparedness, 1*(1), 57-62. doi: 10.1097/DMP.0b013e318065b7e1

Association of State and Territorial Health Officials. (2011). ASTHO Profile of State Public Health (Vol. 2). Retrieved from http://www.astho.org/uploadedFiles/_Publications/Files/Survey_Research/ASTHO_State_Profiles_Single%5B1%5D%20lo%20res.pdf

Aten, J. D., Leavell, K., Gonzalez, R., Luke, T., Defee, J., & Harrison, K. (2011). Everyday technologies for extraordinary circumstances: Possibilities for enhancing disaster communication. *Psychological Trauma: Theory, Research, Practice, and Policy, 3*(1), 16-20. doi: 10.1037/a0021259

Avery, E., Lariscy, R., Armador, E., Ickowitz, T., Primm, C., & Taylor, A. (2010). Diffusion of social media among public relations practitioners in health departments across various community population sizes. *Journal of Public Relations Research, 22*(3), 336-358. doi: 10.1080/10627261003614427

Avery, G. H., & Wright, T. (2010). Does federal assistance to health departments for bioterrorism preparedness improve local public health activity? An empirical evaluation using the 2005 NACCHO profile of local health departments. *Journal*

of Homeland Security & Emergency Management, 7(1), 1-19. doi: 10.2202/1547-7355.1630

Avery, G. H., & Zabriskie-Timmerman, J. (2009). The impact of federal bioterrorism funding programs on local health department preparedness activities. *Eval Health Prof, 32*(2), 95-127. doi: 10.1177/0163278709333151

Baker, A. C., Jensen, P. J., & Kolb, D. A. (2005). Conversation as experiential learning. *Management Learning, 36*(4), 411-427. doi: 10.1177/1350507605058130

Baker, E. L., Potter, M. A., Jones, D. L., Mercer, S. L., Cioffi, J. P., Green, L. W., . . . Fleming, D. W. (2005). The public health infrastructure and our nation's health. *Annual Review of Public Health, 26*(1), 303-318. doi: 10.1146/annurev.publhealth.26.021304.144647

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemology, 34*(1), 215-220. doi: 10.1093/ije/dyh299

Barros, C. P. (2003). An intervention analysis of terrorism: The spanish eta case. *Defence and Peace Economics, 14*(6), 401-412. doi: 10.1080/1024269032000085170

Bartley, B., Stella, J., & Walsh, L. (2006). What a disaster?! Assessing utility of simulated disaster exercise and educational process for improving hospital preparedness. *Prehospital and Disaster Medicine, 21*(4), 249-255. doi: 10.1017/51049023X00003782

Beaton, R. D., Stevermer, A., Wicklund, J., Owens, D., Boase, J., & Oberle, M. W. (2004). Evaluation of the Washington state national pharmaceutical stockpile * dispensing exercise, part II-dispensary site worker findings. *The Journal of Public*

Health Management and Practice, 10(1), 77-85. http://journals.lww.com/jphmp/pages/default.aspx

Beaubien, J., Paley, M., Badugu, S., Shadrick, S., Ennis, C., & Jacklin, S. (2006). *Using deliberate practice to train military-civilian interagency coordination*. Paper presented at the Interservice/Industry Training, Simulation, and Education Conference. http://ntsa.metapress.com/app/home/contribution.asp?referrer=parent&backto=issue,151,158;journal,7,26;linkingpublicationresults,1:113340,1

Beitsch, L. M., Kodolikar, S., Stephens, T., Shodell, D., Clawson, A., Menachemi, N., & Brooks, R. G. (2006). A state-based analysis of public health preparedness programs in the United States. *Public Health Reports, 121*(6), 737-745. http://www.publichealthreports.org/

Biddinger, P. D., Cadigan, R. O., Auerbach, B. S., Burstein, J., Savoia, E., Stoto, M. A., & Koh, H. K. (2008). On Linkages: Using exercises to identify systems-level preparedness challenges. *Public Health Reports, 123*(1), 96-101. http://www.publichealthreports.org/

Biddinger, P. D., Savoia, E., Massin-Short, S. B., Preston, J., & Stoto, M. A. (2010). Public health emergency preparedness exercises: lessons learned. *Public Health Reports (Washington, D.C.: 1974), 125 Suppl 5*, 100-106. http://www.publichealthreports.org/

Biglan, A., Ary, D., & Wagenaar, A. C. (2000). The value of interrupted time-series experiments for community intervention research. *Prevention Science, 1*(1), 31-49. doi: 10.1023/a:1010024016308

Boyatzis, R. E., Cowen, S. S., & Kolb, D. A. (1995). *Innovation in professional education: Steps on a journey from teaching to learning*. San Francisco: Jossey-Bass.

Brand, M., Kerby, D., Elledge, B., Johnson, D., & Magas, O. (2006). A model for assessing public health emergency preparedness competencies and evaluating training based on the local preparedness plan. *Journal of Homeland Security and Emergency Management, 3*(2). doi: 10.2202/1547-7355.1182

Brandeau, M. L., McCoy, J. H., Hupert, N., Holty, J.-E., & Bravata, D. M. (2009). Recommendations for modeling disaster responses in public health and medicine: A position paper of the Society for Medical Decision Making. *Med Decis Making, 29*(4), 438-460. doi: 10.1177/0272989x09340346

Buehler, J. W., & Holtgrave, D. R. (2007). Who gets how much: Funding formulas in federal public health programs Available from Ovid Technologies Journals@Ovid. (1078-4659). http://ovidsp.ovid.com/ovidweb.cgi?T=JS&NEWS=N&PAGE=fulltext&AN=00124784-200703000-00010&D=ovfti

Campbell, D., & Ross, H. L. (1968). The Connecticut crackdown on speeding: Time-series analysis in quasi-experimental analysis. *Law & Society Review, 3*(1), 33-54. http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1540-5893

Carr, A. (2010). WIC & SNAP 101 - How do these programs work? Retrieved from http://www.foodandcommunityfellows.org/blog/2010/wic-snap-101-%E2%80%93-how-do-these-programs-work

Cauley, J., & Im, E. I. (1988). Intervention policy analysis of skyjackings and other terrorist incidents. *The American Economic Review, 78*(2), 27-31. http://www.aeaweb.org/aer/index.php

Centers for Disease Control and Prevention. (2011). Ten great public health achivements - United States, 2001-2010. *MMWR: Morbidity & Mortality Weekly Report, 60*(19), 619-623. http://www.cdc.gov/mmwr/

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55-81. doi: 10.1016/0010-0285(73)90004-2

Chretien, J.-P., Tomich, N. E., Gaydos, J. C., & Kelley, P. W. (2009). Real-time public health surveillance for emergency preparedness. *American Journal of Public Health, 99*(8), 1360-1363. doi: doi: 10.2105/AJPH.2008.133926

Cooper, H., & Hedges, L. (Eds.). (1994). *The handbook of research synthesis*. New York: The Russell Sage Foundation.

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment, 91*(2), 121-136. doi: 10.1080/00223890802634175

Crichton, M. T., Flin, R., & Rattray, W. A. R. (2000). Training decision makers – Tactical decision games. *Journal of Contingencies and Crisis Management, 8*(4), 208-217. doi: 10.1111/1468-5973.00141

Dausey, D., Buehler, J., & Lurie, N. (2007). Designing and conducting tabletop exercises to assess public health preparedness for manmade and naturally occurring biological threats. *BMC Public Health, 7*(1), 92. doi: 10.1186/1471-2458-7-92

Day, T. G. (2003). The Autumn 2001 anthrax attack on the United States Postal Service: The consequences and response. *Journal of Contingencies and Crisis Management, 11*(3), 110-117. doi: 10.1111/1468-5973.1103004

de Bruin, A. B. H., Smits, N., Rikers, R. M. J. P., & Schmidt, H. G. (2008). Deliberate practice predicts performance over time in adolescent chess players and drop-outs: A linear mixed models analysis. *British Journal Of Psychology (London, England: 1953), 99*(Pt 4), 473-497. doi: 10.1348/000712608X295631

DeGroff, A., Schooley, M., Chapel, T., & Poister, T. H. (2010). Challenges and strategies in applying performance measurement to federal public health programs. *Evaluation and Program Planning, 33*(4), 365-372. doi: 10.1016/j.evalprogplan.2010.02.003

Department of Homeland Security, F. E. M. A., Homeland Security Exercise and Evaluation Program. (2007a). *Exercise planning and conduct* (Vol. II). Washington: Government Printing Office.

Department of Homeland Security, F. E. M. A., Homeland Security Exercise and Evaluation Program. (2007b). *Overview and exercise program management* (Vol. 1). Washington: Government Printing Office.

Dewey, J. (1938). *Experience and education*. New York: Touchstone.

Division of State and Local Readiness. (2006). *Continuation funding opportunity announcement AA154: Public health emergency preparedness*. Atlanta, GA: Retrieved from http://emergency.cdc.gov/cdcpreparedness/coopagreement/pdf/fy06announcement.pdf.

Duncan, W. J., Ginter, P. M., Rucks, A. C., Wingate, M. S., & McCormick, L. C. (2007). Organizing emergency preparedness within United States public health departments. *Public Health, 121*(4), 241-250. doi: 10.1016/j.puhe.2006.10.014

Dunn, T. G., & Shriner, C. (1999). Deliberate practice in teaching: What teachers do for self-improvement. *Teaching and Teacher Education, 15*(6), 631-651. doi: 10.1016/s0742-051x(98)00068-7

Eccles, M., Grimshaw, J., Campbell, M., & Ramsay, C. (2003). Research designs for studies evaluating the effectiveness of change and improvement strategies. *Quality and Safety in Health Care, 12*(1), 47-52. doi: 10.1136/qhc.12.1.47

Eckner, A. (2012). *A note on trend and seasonality estimation for unevenly-spaced time series*. Retrieved from http://www.eckner.com/papers/trend_and_seasonality.pdf

Enders, W., Parise, G. F., & Sandler, T. (1992). A time-series analysis of transnational terrorism: Trends and cycles. *Defence Economics, 3*(4), 305-320. doi: 10.1080/10430719208404739

Enders, W., & Sandler, T. (2002). Patterns of transnational terrorism, 1970–1999: Alternative time-series estimates. *International Studies Quarterly, 46*(2), 145-165. doi: 10.1111/1468-2478.00227

Ericsson, K. A. (2000). Expertise in intepreting: An expert-performance perspective. *Interpreting, 5*(2), 189-222. http://benjamins.com/#catalog/journals/intp.5.2/main

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine, 79*(10), S70-S81. http://journals.lww.com/academicmedicine/toc/2004/10001

Ericsson, K. A. (2006). The Influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 683-703). New York City: Cambridge University Press.

Ericsson, K. A. (2007). An expert-performance perspective of research on medical expertise: the study of clinical performance. *Medical Education, 41*(12), 1124-1130. doi: 10.1111/j.1365-2923.2007.02946.x

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine, 15*(11), 988-994. doi: 10.1111/j.1553-2712.2008.00227.x

Ericsson, K. A. (Ed.). (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games*. Mahwah, NJ: Lawrence Erlbaum.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*(8), 725-747. doi: 10.1037/0003-066x.49.8.725

Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance* (1st ed.). New York City: Cambridge University Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363-406. doi: 10.1037/0033-295x.100.3.363

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review Of Psychology, 47*(1), 273-305. doi: 10.1146/annurev.psych.47.1.273

Ericsson, K. A., Whyte, J., & Ward, P. (2007). Expert performance in nursing: Reviewing research on expertise in nursing withing the framework of the expert-performance approach. *Advances in Nursing Science, 30*(1), E58-E71. http://journals.lww.com/advancesinnursingscience/Fulltext/2007/01000/Expert_Performance_in_Nursing__Reviewing_Research.14.aspx

Farmer, L. C., & Williams, G. R. (2005). *The rigorous application of deliberate practice methods in skills courses*. Paper presented at the UCLA/IALS Sixth International Clinical Conference on Enriching Clinical Education, Los Angeles, CA. http://cdn.law.ucla.edu/SiteCollectionDocuments/workshops%20and%20colloquia/clinical%20programs/larry%20farmer%20and%20gerald%20williams.pdf

Fernand, G. (1998). Expert memory: a comparison of four theories. *Cognition, 66*(2), 115-152. doi: 10.1016/s0010-0277(98)00020-1

Fitch, J. P., Raber, E., & Imbro, D. R. (2003). Technology challenges in responding to biological or chemical attacks in the civilian sector. *Science, 302*(5649), 1350-1354. doi: 10.1126/science.1085922

Ford, E. W., Duncan, J., & Ginter, P. M. (2003). The structure of state health departments: A strategic analysis. *Medical Care Research and Review, 60*(1), 37-51. doi: 10.1177/1077558702250231

Foster, D. (2003). Smallpox as a biological weapon: Implications for the critical care clinician. *Dimensions of Critical Care Nursing, 22*(1), 2-7.

http://journals.lww.com/dccnjournal/Fulltext/2003/01000/Smallpox_as_a_Biological_Weapon__Implications_for.1.aspx

Fowkes, V., Blossom, H., Sandrock, C., Mitchell, B., & Brandstein, K. (2010). Exercises in emergency preparedness for health professionals in community clinics. *Journal of Community Health, 35*(5), 512-518. doi: 10.1007/s10900-010-9221-1

Fox, D. M. (2005). Evidence of evidence-based health policy: The politics of systematic reviews In coverage decisions. *Health Affairs, 24*(1), 114-122. doi: 10.1377/hlthaff.24.1.114

Frank, E. (2005). Funding the public health response to terrorism. *British Medical Journal, 331*(7516), 526-527. doi: 10.1136/bmj.331.7516.526

Frankfort-Nachmias, C., & Nachmias, D. (2008). *Research methods in the social sciences* (7th ed.). New York: Worth Publishers.

Fraser, M. R., & Brown, D. L. (2000). Bioterrorism preparedness and local public health agencies: Building response capacity. *Public Health Reports, 115*(4), 326-330. http://www.publichealthreports.org/

Frick, K. (2009). Spending $6 billion more on cancer research. Retrieved from http://economistlens.blogspot.com/

Friedrich, M. J. (2002). Practice makes perfect. *JAMA: The Journal of the American Medical Association, 288*(22), 2808-2812. doi: 10.1001/jama.288.22.2808

Frist, B. (2002). Public Health And National Security: The Critical Role Of Increased Federal Support. *Health Affairs, 21*(6), 117-130. doi: 10.1377/hlthaff.21.6.117

Gebbie, K. M., Valas, J., Merrill, J., & Morse, S. (2006). Role of exercises and drills in the evaluation of public health in emergency response. *Prehospital Disaster Medicine, 21*(3), 173-182. doi: 10.1017/S1049023X00003642

Glass, G. V., Wilson, V. L., & Gottman, J. M. (2008). *Design and analysis of time-series experiments*. Charlotte: Information Age Publishing, Inc.

Gore, A. (1997). Serving the American public best practices in performance measurement : Benchmarking Study Report. [Washington, D.C.]: National Performance Review. http://govinfo.library.unt.edu/npr/library/papers/benchmrk/nprbook.html

Gostin, L. O. (2005). *Jacobson v Massachusettes* at 100 years: Police power and civil liberties in tension. *American Journal of Public Health, 95*(4), 576-581. doi: 10.2105/AJPH.2004.055152

Gursky, E., Inglesby, T., & O'Toole, T. (2003). Anthrax 2001: Observations on the medical and public health response. *Biosecurity and Bioterrorism, 1*(2), 14. doi: 10.1089/153871303766275763

Hargan, E. D. (2008). Setting expectations for the federal role in public health emergencies. *Journal of Law, Medicine & Ethics, 36*(1 Suppl), 8-12. doi: 10.1111/j.1748-720X.2008.00253.x

Harvey, H., & Lister, S. (2004). *An overview of the U.S. public health system in the context of bioterrorism*. Washington, DC: Congressional Research Service Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA480138.

Hearne, S., Segal, L., & Earls, M. (2005). *Ready or not? Protecting the public's health from diseases, disasters, and bioterrorism*. Washington, DC: Trust for America's Health. http://healthyamericans.org/reports/bioterror05/

Helton, W. S. (2007). Deliberate practice in dogs: A canine model of expertise. *The Journal Of General Psychology, 134*(2), 247-257. doi: 10.3200/GENP.134.2.247-258

High, E. H., Lovelace, K. A., Gansneder, B. M., Strack, R. W., Callahan, B., & Benson, P. (2008). Promoting community preparedness: Lessons learned from the implementation of a chemical disaster tabletop exercise. *Health Promotion Practice*. doi: 10.1177/1524839908325063

Hodges, N. J., Kerr, T., Starkes, J. L., Weir, P. L., & Nananidou, A. (2004). Predicting performance times from deliberate practice hours for triathletes and swimmers: What, when, and where is practice important? *Journal of Experimental Psychology: Applied, 10*(4), 219-237. doi: 10.1037/1076-898x.10.4.219

Holder, H. D., & Blose, J. O. (1987). Impact of changes in distilled spirits availability on apparent consumption: A time series analysis of liquor-by-the-drink. *British Journal of Addiction, 82*(6), 623-631. doi: 10.1111/j.1360-0443.1987.tb01524.x

Homeland Security Exercise and Evaluation Program. (2007a). *Volume I: HSEEP overview and exercise program management*. Washington, DC: Government Printing Office.

Homeland Security Exercise and Evaluation Program. (2007b). *Volume I: HSEEP overview and exercise program management.* Washington, DC: Government Printing Office.

Hsu, E., Jenckes, M., Catlett, C., Robinson, K., Feuerstein, C., Cosgrove, S., . . . Bass, E. (2004). Effectiveness of hospital staff mass-casualty incident training methods: A systematic literature review. *Prehospital and Disaster Medicine, 19*(3), 191-199. doi: 10.1017/S1049023X00001771

Hutchinson, M. K., & Holtman, M. C. (2005). Analysis of count data using poisson regression. *Research in Nursing & Health, 28*(5), 408-418. doi: 10.1002/nur.20093

Institute of Medicine. (2002). *The future of the public's health in the 21st century*. Washington, D.C.: The National Academies Press.

Johnson, Y., Herrmann, J. A., Wallace, R. L., Troutt, H. F., & Myint, M. S. (2009). Development and implementation of a functional exercise to assess public health agency response to foodborn bioterrorism. *Journal of Homeland Security and Emergency Management, 6*(1), Article 49. doi: 10.2202/1547-7355.1570

Karazsia, B. T., & van Dulmen, M. H. M. (2008). Regression models for count data: Illustrations using longitudinal predictors of childhood injury. *Journal of Pediatric Psychology, 33*(10), 1076-1084. doi:10.1093/jpepsy/jsn055

Katz, A., Staiti, A. B., & McKenzie, K. L. (2006). Preparing for the unknown, responding to the known: Communities and public health preparedness. *Health Affairs, 25*(4), 946-957. doi: 10.1377/hlthaff.25.4.946

Kayes, A. B., Kayes, D. C., & Kolb, D. A. (2005). Experiential learning in teams. *Simulation & Gaming, 36*(3), 330-354. doi: 10.1177/1046878105279012

Keeler, T. E., Hu, T.-W., Barnett, P. G., & Manning, W. G. (1993). Taxation, regulation, and addiction: A demand function for cigarettes based on time-series evidence. *Journal of Health Economics, 12*(1), 1-18. doi: 10.1016/0167-6296(93)90037-f

Kerr, P. (2008). *Nuclear, biological, and chemical weapons and missiles: States and trends*. (RL30699). Washington, DC: Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA477531.

Kinner, K., & Pellegrini, C. (2009). Expenditures for public health: Assessing historical and prospective trends. *American Journal of Public Health, 99*(10), 1780-1791. doi: 10.2105/AJPH.2008.142422

Klaiman, T., & Ibrahim, J. (2010). State health department structure and pandemic planning. *Journal of Public Health Management & Practice, 16*(2), E1-E7. doi: 10.1097/PHH.0b013e3181b83475

Koh, H. K., Elqura, L. J., Judge, C. M., & Stoto, M. A. (2008). Regionalization of local public health systems in the era of preparedness. *Annual Review of Public Health, 29*(1), 205-218. doi: doi:10.1146/annurev.publhealth.29.020907.090907

Kolb, A. Y., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning & Education, 4*(2), 193-212. doi: 10.5465/AMLE.2005.17268566

Kolb, D. A. (1971). *Individual learning styles and the learning process*. Working paper. Sloan School of Management, Massachusetts Institute of Technology. Retrieved from http://library.mit.edu/F/ES43H1YDXK5ABNAMGL64GESBAQ3GI3BXGGGCYUBYDDP1R1P6QU-17720?func=find-b&find_code=OCA&request=14471565

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs: Prentice-Hall.

Kolb, D. A., Boyatzis, R., & Mainemelis, C. (2000). Experiential learning theory: Previous research and new directions. In R. J. Sternberg & L. F. Zhang (Eds.), *Perspectives on cognitive learning, and thinking styles* (pp. 227-249). Mahwah, NJ: Lawrence Erlbaum

Lamb, D. (2007). Could simulated emergency procedures practised in a static environment improve the clinical performance of a Critical Care Air Support Team (CCAST)?: A literature review. *Intensive and Critical Care Nursing, 23*(1), 33-42. doi: DOI: 10.1016/j.iccn.2006.07.002

Lawrence, J. M., Watkins, M. L., Ershoff, D., Petitti, D. B., Chiu, V., Postlethwaite, D., & Erickson, J. D. (2003). Design and evaluation of interventions promoting periconceptional multivitamin use. *American Journal of Preventive Medicine, 25*(1), 17-24. doi: 10.1016/s0749-3797(03)00097-7

Lee, Y.-I., Trim, P., Upton, J., & Upton, D. (2009). Large emergency-response exercises: Qualitative characteristics - A survey. *Simulation Gaming, 40*(6), 726-751. doi: 10.1177/1046878109334006

Levi, J., Segal, L., Gadola, E., Juliano, C., & Speulda, N. (2006). *Ready or not? Protecting the public's health from diseases, disasters, and bioterrorism*. Washington, DC: Trust for America's Health. http://healthyamericans.org/reports/bioterror06/

Levi, J., Vinter, S., & Segal, L. (2007). *Ready or not? Protecting the public's health from diseases, disasters, and bioterrorism.* Washington DC: Trust for America's Health. http://healthyamericans.org/reports/bioterror07/

Levi, J., Vinter, S., Segal, L., & St. Laurent, R. (2010). *Ready or not? Protecting the public's health from diseases, disasters, and bioterrorism* (pp. 52). Washington, DC: The Trust for America's Health. http://healthyamericans.org/reports/bioterror10/

Levi, J., Vinter, S., St. Laurent, R., & Segal, L. (2008). *Ready or not? Protecting the public's health from diseases, disasters, and bioterrorism*. Washington, DC. http://healthyamericans.org/reports/bioterror08/

Lewis, J., Berg, J., & Gates, E. (2006). Evaluation of full scale exercise participation as a tool to improve hospital preparedness for a radiological terrorism event. *Annals of Emergency Medicine, 48*(4), S52. doi: 10.1016/j.annemergmed.2006.07.620

Lindley, M. C., Horlick, G. A., Shefer, A. M., Shaw, F. E., & Gorji, M. (2007). Assessing state immunization requirements for healthcare workers and patients. *American Journal of Preventive Medicine, 32*(6), 459-465. doi: DOI: 10.1016/j.amepre.2007.02.009

Lisko, S. A., & O'Dell, V. (2010). Integration of theory and practice: experiential learning theory and nursing education. *Nursing Education Perspectives, 31*(2), 106-108. doi: 10.1043/1536-5026-31.2.106

Lovelace, K., Bibeau, D., Gansneder, B., Hernandez, E., & Cline, J. S. (2007). All-Hazards preparedness in an era of bioterrorism funding. *Journal of Public Health*

Management and Practice, 13(5), 465-468. doi: 10.1097/01.PHH.0000285198.62049.1a

Lurie, N., Wasserman, J., & Nelson, C. D. (2006). Public health preparedness: Evolution or revolution? *Health Affairs, 25*(4), 935-945. doi: 10.1377/hlthaff.25.4.935

Mann, N. C., MacKenzie, E. J., & Anderson, C. (2004). Public health preparedness for mass-casualty events: A 2002 state-by-state assessment. *Prehospital Disaster Medicine, 19*(3), 245-255. doi: 10.1017/S1049023X00001840

McKinney, E. H., Jr., & Davis, K. J. (2003). Effects of deliberate practice on crisis decision performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 45*(3), 436-444. doi: 10.1518/hfes.45.3.436.27251

McLeod, K. S. (2000). Our sense of Snow: The myth of John Snow in medical geography. *Social Science & Medicine, 50*(7-8), 923-935. doi: 10.1016/s0277-9536(99)00345-7

Moore, S., Mawji, A., Shiell, A., & Noseworthy, T. (2007). Public health preparedness: a systems-level approach. *Journal of Epidemiology and Community Health, 61*(4), 282-286. doi: 10.1136/jech.2004.030783

Morey, J. C., Simon, R., Jay, G. D., Wears, R. L., Salisbury, M., Dukes, K. A., & Berns, S. D. (2002). Error reduction and performance improvement in the emergency department through formal teamwork training: Evaluation results of the MedTeams project. *Health Services Research, 37*(6), 1553-1581. doi: 10.1111/1475-6773.01104

Morrison, A. M., & Catanzaro, A. M. (2010). High-fidelity simulation and emergency preparedness. *Public Health Nursing, 27*(2), 164-173. doi: 10.1111/j.1525-1446.2010.00838.x

Morrow, C. B. (2007). Public health emergency preparedness and the evolution of public health in the 21st century. *Journal of Public Health Management & Practice, 13*(5), 439-440. doi: 10.1097/01.PHH.0000285193.93436.02

Morse, A. (2002). Bioterrorism preparedness for local health departments. *Journal of Community Health Nursing, 19*(4), 203-211. doi: 10.1207/S15327655JCHN1904_01

Moulaert, V., Verwijnen, M. G., Rikers, R., & Scherpbier, A. J. (2004). The effects of deliberate practice in undergraduate medical education. *Medical Education, 38*(10), 1044-1052. doi: 10.1111/j.1365-2929.2004.01954.x

National Security Council. (2007). *National Strategy for Homeland Security*. Washington, DC: Government Printing Office.

Nelson, C., Lurie, N., Wasserman, J., Zakowski, S., & Leuschner, K. J. (2008). Conceptualizing and defining public health emergency preparedness. Santa Monica, CA: RAND Corporation. doi: 10.2105/AJPH.2007.114496

Nelson, C., Willis, H., Chan, E., Shelton, S., & Parker, A. (2011). Federal initiative increases community preparedness for public health emergencies. *Health Affairs, 29*(12), 8. doi: 10.1377/hlthaff.2010.0189

Nunn, S. (2002). The future of public health preparedness. *Journal of Law, Medicine & Ethics, 30*(3), 202. http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1748-720X

Office of Public Health Preparedness and Response. (2011). Funding, guidance, and technical assistance to states, localities, and territories. from http://www.cdc.gov/phpr/coopagreement.htm

Olson, D. K., Scheller, A., Larson, S., Lindeke, L., & Edwardson, S. (2010). Using gaming simulation to evaluate bioterrorism and emergency readiness education. *Public Health Reports, 125*(3), 468-477. http://www.publichealthreports.org/

Oulasvirta, A., Wahlström, M., & Ericsson, K. A. (2011). What does it mean to be good at using a mobile device? An investigation of three levels of experience and skill. *International Journal of Human-Computer Studies, 69*(3), 155-169. doi: 10.1016/j.ijhcs.2010.11.003

The Pandemic and All-Hazards Preparedness Act, Pub. L. No. 109-417 § 101 et seq, 120 Stat. 2831 (2006).

Perkins, B., Popovic, T., & Yeskey, K. (2002). Public health in the time of bioterrorism. *Emerging Infectious Diseases, 8*(10), 1015-1018. doi: 10.3201/eid0810.020444

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

Pistole, M. C., Kinyon, J., & Keith, C. B. (2008). Group experiential learning with undergraduate nursing students: An interdisciplinary approach. *Journal for Specialists in Group Work, 33*(2), 161-178. doi: 10.1080/01933920801977348

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate

practice for academic performance. *Contemporary Educational Psychology, 30*(1), 96-116. doi: 10.1016/j.cedpsych.2004.06.001

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12*(4), 531-544. doi: 10.1177/014920638601200408

Radhakrishna, R. B., Yoder, E. P., & Ewing, J. C. (2007). *Strategies for linking theoretical framework and research types.* Paper presented at the Proceedings of the 2007 AAAE Research Conference, Minneapolis, MN. http://aged.caf.wvu.edu/Research/NAERC-2007/intro.pdf

Reis, H., & Judd, C. (Eds.). (2000). *Handbook of research methods in social and personality psychology*. New York: Cambridge University Press.

Richmond, A., Hostler, L., Leeman, G., & King, W. (2010). A brief history and overview of CDC's Centers for Public Health Preparedness cooperative agreement program. *Public Health Reports (Washington, D.C.: 1974), 125 Suppl 5*, 8-14. http://www.publichealthreports.org/

Salas, E., Rosen, M. A., Held, J. D., & Weissmuller, J. J. (2009). Performance measurement in simulation-based training: A review and best practices. *Simulation Gaming, 40*(3), 328-376. doi: 10.1177/1046878108326734

Savoia, E., Biddinger, P., Fox, P., Levin, D. E., Stone, L., & Stoto, M. A. (2009). Impact of tabletop exercises on participants' knowledge of and confidence in legal authorities for infectious disease emergencies. *Disaster Medicine and Public Health Preparedness, 3*(2), 104-110. doi: 10.1097/DMP.0b013e3181a539bc

Savoia, E., Massin-Short, S. B., Rodday, A. M., Aaron, L. A., Higdon, M. A., & Stoto, M. A. (2009). Public health systems research in emergency preparedness: A review of the literature. *American Journal of Preventive Medicine, 37*(2), 150-156. doi: 10.1016/j.amepre.2009.03.023

Savoia, E., Testa, M. A., Biddinger, P. D., Cadigan, R. O., Koh, H., Campbell, P., & Stoto, M. A. (2009). Assessing public health capabilities during emergency preparedness tabletop exercises: Reliability and validity of a measurement tool. *Public Health Reports (Washington, D.C.: 1974), 124*(1), 138-148. http://www.publichealthreports.org/

Schuman, H., Walsh, E., Olson, C., & Etheridge, B. (1985). Effort and reward: The assumption that college grades are affected by quantity of study. *Social Forces, 63*(4), 945-966. doi: 10.1093/sf/63.4.945

Sonnentag, S., & Kleine, B. (2000). Deliberate practice at work: A study with insurance agents. *Journal of Occupational & Organizational Psychology, 73*(1), 87-102. doi: 10.1348/096317900166895

Soumerai, S. B., McLaughlin, T. J., Ross-Degnan, D., Casteris, C. S., & Bollini, P. (1994). Effects of limiting Medicaid drug-reimbursement benefits on the use of psychotropic agents and acute mental health services by patients with schizophrenia. *New England Journal of Medicine, 331*(10), 650-655. doi: doi:10.1056/NEJM199409083311006

Szreter, S. (2002). Rethinking McKeown: The relationship between public health and social change. *American Journal of Public Health, 92*(5), 722-725. doi: 10.2105/AJPH.92.5.722

Tilson, H., & Gebbie, K. M. (2004). The public health workforce. *Annual Review of Public Health, 25*(1), 341-356. doi: 10.1146/annurev.publhealth.25.102802.124357

Tucker, J. B. (1999). Historical trends related to bioterrorism: An empirical analysis. *Emerging Infectious Disease, 5*(4), 498-504. doi: 10.3201/eid0504.990406

Tuffiash, M., Roring, R. W., & Ericsson, K. A. (2007). Expert performance in SCRABBLE: Implications for the study of the structure and acquisition of complex skills. *Journal of Experimental Psychology: Applied, 13*(3), 124-134. doi: 10.1037/1076-898x.13.3.124

U.S. Department of Health and Human Services, & Centers for Disease Control and Prevention. (2005). *Program Announcement AA154 - 2005*. Retrieved from http://emergency.cdc.gov/planning/guidance05/pdf/annoucement.pdf.

U.S. Department of Health and Human Services, & Centers for Disease Control and Prevention. (2007). *Public Health Emergency Preparedness Cooperative Agreement; Budget Period 8 Performance Measures: Definitions and Guidance*. Retrieved from http://emergency.cdc.gov/cdcpreparedness/coopagreement/pdf/fy07guidance_definitions_v1_122607.pdf.

U.S. General Accounting Office. (2002). *Diffues security threats: USPS air filtration systems need more testing and cost benefit analysis before implementation*. (GAO-02-838). Retrieved from http://www.gpo.gov/fdsys/pkg/GAOREPORTS-GAO-02-838/pdf/GAOREPORTS-GAO-02-838.pdf.

U.S. Government Accountability Office. (2007). *Public health and hospital emergency preparedness programs: Evolution of performance measurement systems to measure progress*. (GAO-07-485R). Washington D.C.: Retrieved from http://www.worldcat.org/title/public-health-and-hospital-emergency-preparedness-programs-evolution-of-performance-measurement-systems-to-measure-progress/oclc/191819984&referer=brief_results.

Unger, J. M., Keith, N., Hilling, C., Gielnik, M. M., & Frese, M. (2009). Deliberate practice among South African small business owners: Relationships with education, cognitive ability, knowledge, and success. *Journal of Occupational & Organizational Psychology, 82*(1), 21-44. doi: 10.1348/096317908X304361

Ward, P., Hodges, N. J., Starkes, J. L., & Williams, M. A. (2007). The road to excellence: deliberate practice and the development of expertise. *High Ability Studies, 18*(2), 119-153. doi: 10.1080/13598130701709715

William, D. (1996). Standards in examinations: A matter of trust? *Curriculum Journal, 7*(3), 293 - 306. doi: 10.1080/0958517960070303

Williams, A. M., Ericsson, K. A., Ward, P., & Eccles, D. W. (2008). Research on expertise in sport: Implications for the military. *Military Psychology, 20*(1 supp 1), 123 - 145. doi: 10.1080/08995600701804863

Yeager, V. A., Menachemi, N., McCormick, L. C., & Ginter, P. M. (2010). The nature of the public health emergency preparedness literature 2000-2008: A quantitative analysis. *Journal Of Public Health Management And Practice: JPHMP, 16*(5), 441-449. doi: 10.1097/PHH.0b013e3181c33de4

Curriculum Vitae

Donna Beth Knutson

Educational Background

2009 - 2013 (expected)	Ph.D. Public Policy and Administration Focus in Law and Public Policy **Walden University, Minneapolis, MN** *Dissertation:* "Drills and Exercises as Interventions To Improve Public Health Emergency Response" (Advisor: James Mosko, Ph.D)
1983-1985	Master's of Science Degree in Education Emphasis in Exercise Physiology **University of Akron, Akron, Ohio**
1979-1983	Bachelor's of Science Degree in Physical Education **Ohio University, Athens, Ohio**

Professional Background

2012 - Senior Advisor to the Director, Financial Management Office, CDC
Centers for Disease Control and Prevention, Atlanta, GA
Office of the Chief Operating Officer
Office of the Director

2010-2012 Senior Advisor to the Associate Director for Program and the Associate Director for Policy, CDC
Centers for Disease Control and Prevention, Atlanta, GA
Office of the Director

2009-2010 Senior Advisor to the Deputy Director, CDC
Centers for Disease Control and Prevention, Atlanta, GA
Office of the Director

2008-2009: Deputy Director, National Center for Injury Prevention and Control
Centers for Disease Control and Prevention, Atlanta, GA
Coordinating Center for Environmental Health and Injury Prevention

2008-2008: Director, Division of State and Local Readiness (Acting)
Centers for Disease Control and Prevention, Atlanta, GA
Coordinating Office of Terrorism Preparedness and Emergency Response

2003-2008: Senior Advisor and Associate Director of Program Development
Centers for Disease Control and Prevention, Atlanta, GA
Coordinating Office of Terrorism Preparedness and Emergency Response

2002-2003 Vice President, Public Health Consultant
Scientific Technologies Corporation, Atlanta, GA

1998-2002 Executive Director
Council of State and Territorial Epidemiologists, Atlanta, GA

1995-1998 Section Chief
Centers for Disease Control and Prevention, Atlanta, GA
National Center for Chronic Disease Prevention and Health Promotion, Division of Cancer Prevention and Control, Program Services Branch

1994-1995 Program Consultant
Centers for Disease Control and Prevention, Atlanta, GA
National Center for Chronic Disease Prevention and Health Promotion, Division of Cancer Prevention and Control, Program Services Branch

1989-1994 Health Education Specialist
Centers for Disease Control and Prevention
National Center for Chronic Disease Prevention and Health Promotion, Division of Chronic Disease Control and Community Intervention, Health Interventions Translations Branch.

1992-1994 Wisconsin Department of Health and Social Services
1989-1992 Maryland Department Of Health and Mental Hygiene

1985-1989 Project Director
Health Promotion/Risk Reduction Project of Summit County, Akron, Ohio; a health education consortium of the Summit County, City of Akron, and City of Barberton Health Departments

1984-1985 Health Educator
Zanesville-Muskingum County Health Department, Zanesville, Ohio

1983-1984 Graduate Assistant
University of Akron, Akron, Ohio

Publications:
US Civilian Smallpox Predaredness and Response Program, 2003.
Strikas RA, Neff LJ, Rotz L, Cono J, Knutson D, Henderson J, Orenstein WA. Clin Infect Dis. 2008 Mar 15;46 Suppl 3:S157-67.

Strategies for follow-up and treatment services in state breast and cervical cancer screening programs. Lantz PM, Richardson LC, Macklem DJ, Shugarman LR, Knutson DB, Sever LE.
Womens Health Issues. 1999 Jan-Feb;9(1):42-9.

Wisconsin's experience with the national nicotine patch giveaway.
Lickteig S, Knutson D, Yoast R, Remington PL. Wis Med J. 1993 Nov;92(11):631-2.