This product is part of the RAND Corporation occasional paper series. RAND occasional papers may include an informed perspective on a timely policy issue, a discussion of new research methodologies, essays, a paper presented at a conference, a conference summary, or a summary of work in progress. All RAND occasional papers undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.

OCCASIONAL
PAPER

The Problem of Measuring Emergency Preparedness

The Need for Assessing "Response Reliability" as Part of Homeland Security Planning

Brian A. Jackson

Homeland Security

A RAND INFRASTRUCTURE, SAFETY, AND ENVIRONMENT PROGRAM

This Occasional Paper results from the RAND Corporation's continuing program of self-initiated research. Support for such research is provided, in part, by the generosity of RAND's donors and by the fees earned on client-funded research.

Published 2008 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: http://www.rand.org
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Preface

Created in the wake of the September 11, 2001, terrorist attacks, the Department of Homeland Security came into being with the daunting core mission of taking action to protect the United States from terrorist attack and the simultaneous requirement to continue to perform the numerous other critical functions of all its component agencies. The complexity of the department's mission was further compounded by the fact that it depended not only on the success of the department's component agencies, but also on the efforts of a national homeland security enterprise comprised of organizations at the federal, state, and local levels, both inside and outside government. That there have been challenges in carrying out this endeavor in the years since should surprise no one. However, it has also been the fortunate reality that, whatever those challenges, at the time of this writing, there have been no major terrorist attacks within the United States since 9/11.

Transitions in presidential administrations are traditionally opportunities for the country to examine national policy goals, assess how we as a nation are trying to achieve them, ask whether what we are doing is working, and make adjustments where necessary. For homeland security, the upcoming presidential transition is even more important as it is the first change in administration since the creation of the Department of Homeland Security. To contribute to policy debate during this transition and to inform future homeland security policy development, the RAND Corporation initiated an effort to reexamine key homeland security policy issues and explore new approaches to solving them.

This paper is the first in a series of short papers resulting from this effort. The goal was not to comprehensively cover homeland security writ large, but rather to focus on a small set of policy areas, produce essays exploring different approaches to various policy problems, and frame key questions that need to be answered if homeland security policy is to be improved going forward. The results of this effort were diverse, ranging from thought experiments about ways to reframe individual policy problems to more wide-ranging examinations of broader policy regimes. These discussions should be of interest to homeland security policymakers at the federal, state, and local levels and to members of the public interested in homeland security and counterterrorism.

This effort is built on a broad foundation of RAND homeland security research and analysis carried out both before and since the founding of the Department of Homeland Security. Examples of those studies include:

- Brian A. Jackson, Peter Chalk, Kim Cragin, Bruce Newsome, John V. Parachini, William Rosenau, Erin M. Simpson, Melanie Sisson, and Donald Temple, *Breaching the Fortress*

Wall: Understanding Terrorist Efforts to Overcome Defensive Technologies, Santa Monica, Calif.: RAND Corporation, MG-481-DHS, 2007.
- Tom LaTourrette, David R. Howell, David E. Mosher, and John MacDonald, *Reducing Terrorism Risk at Shopping Centers: An Analysis of Potential Security Options*, Santa Monica, Calif.: RAND Corporation, TR-401, 2006.
- Henry H. Willis, Andrew R. Morral, Terrence K. Kelly, and Jamison Jo Medby, *Estimating Terrorism Risk*, Santa Monica, Calif.: RAND Corporation, MG-388-RC, 2005.

The RAND Homeland Security Program

This research was conducted under the auspices of the Homeland Security Program within RAND Infrastructure, Safety, and Environment (ISE). The mission of RAND Infrastructure, Safety, and Environment is to improve the development, operation, use, and protection of society's essential physical assets and natural resources and to enhance the related social assets of safety and security of individuals in transit and in their workplaces and communities. Homeland Security Program research supports the Department of Homeland Security and other agencies charged with preventing and mitigating the effects of terrorist activity within U.S. borders. Projects address critical infrastructure protection, emergency management, terrorism risk management, border control, first responders and preparedness, domestic threat assessments, domestic intelligence, and workforce and training. Information about the Homeland Security Program is available online (http://www.rand.org/ise/security/). Inquiries about homeland security research projects should be sent to the following address:

Andrew Morral, Director, Homeland Security Program, ISE
RAND Corporation
1200 South Hayes Street
Arlington, VA 22202-5050
703-413-1100, x5119
Andrew_Morral@rand.org

Contents

Figure

Summary

In the years since September 11, 2001, the question "is the United States sufficiently prepared for future natural disasters or terrorist attacks" has been prominent in national policy debate. Looking at past response operations, it is generally easy to find things that did not go as well as expected or areas where planning and preparedness efforts seemed to fall short. Though learning from past experience is useful and important, managing preparedness planning and policy "in the rearview mirror"—constantly reacting to the perceived shortcomings of the last response and recovery operation—is not the right path to developing a homeland security and emergency-preparedness policy that serves the nation's needs.

Policymakers and the public need ways to prospectively assess preparedness so they know what they can expect when disaster strikes. Doing so is also critical for resource management. Because of the nature of emergency situations, there will almost always be "more that could have been done" when responding to a particular incident or disaster. But in a world of finite resources, achieving ideal performance in such situations will almost never be possible. Trying to address every shortfall identified retrospectively in response actions therefore risks either creating unsustainable demands for increasing preparedness expenditures or focusing scarce resources on shortfalls that may be easy to see but that may not be the most important preparedness problems.

Over the years, there have been many efforts to assess preparedness. Some have focused on evaluating the resources and activities that are easiest to quantify to provide some insight into what response systems will be able to accomplish. We know that having the right equipment is important, so if equipment is not available to respond to types of incidents that concern us, then response operations are unlikely to go well. Other efforts have gone beyond inventorying resources to develop preparedness standards, assess less-tangible factors (such as training and leadership), or test plans and operations in exercises. Though they provide some insights into preparedness, these methods cannot answer the fundamental question of policymakers and the public: How certain should we as a nation be that the systems we have put in place to respond to damaging events will be able to deliver when called upon?

Whether there is a plan in place and supplies for assisting the victims of a disaster have been bought are key *inputs* to preparedness. However, those inputs will not produce the *outcomes* we want—actually providing relief to disaster victims when and where it is needed—if response organizations, infrastructures, and other components cannot deliver them. Confidence that response plans will be able to be executed as designed depends on the *reliability* of the system that is executing them. A plan to deliver supplies to an area that depends on one road and a single source of vehicles will be less likely to succeed than a plan that provides a

variety of options and routes so operations can be adapted easily in response to the changing operational situation.

To address this concern about response performance, *response reliability* should be evaluated as part of preparedness measurement efforts. Such an assessment would be based on the nature of the system of response organizations, capabilities, and resources and the factors that shape how well it responds. The starting point is to map the system and identify the different elements that shape its performance. Assessing the reliability of the system then requires identifying what could go wrong, estimating the likelihood of breakdowns, identifying their impact on performance (e.g., would a particular failure mode completely disrupt response operations or just reduce their efficiency or effectiveness?), and determining if planning has accounted for them and either built in hedging strategies or is flexible enough to compensate if they occur.

Although making quantitative estimates of response reliability will prove challenging in many circumstances, there are significant advantages to doing so:

- First, such information will help policymakers and the public better understand what past investments in preparedness have bought: A response system that is designed to be 95 percent reliable will look very different—and likely cost more—than one whose chance of success is much lower. The higher cost of the more-reliable system should not be viewed as inefficiency or waste because it is paying for a real performance advantage.
- Second, if assessment identifies major threats to the reliability of response activities, it could be the case that modest investments to address those threats could have high payoffs.

Of course we would prefer that our preparedness and response systems perform effectively at every incident; however, there is no simple answer to the question "how reliable should they be?" Very reliable systems that are able to perform in the most demanding circumstances cost more than those that can do the same things but less predictably. Investments needed to increase the reliability of response must therefore be measured against ways those funds could be used to achieve different homeland security or other national goals. The first step to making those trade-offs, however, is developing better ways to evaluate the reliability of the preparedness and response systems we have now, and the likely effect of further investments in those systems. Given the absence from policy debate of accurate and objective measures of preparedness and response reliability, the full implications of preparedness investments have not been considered. Framing preparedness policies using concepts such as response reliability has the potential to enrich national debate in this area—moving beyond mere argument about funding levels to a more productive discussion of the trade-offs between investments and the levels of performance that we can reasonably expect from our national preparedness system.

CHAPTER ONE

Framing the Issue

In the years since September 11, 2001, the question "is the United States sufficiently prepared for future natural disasters or terrorist attacks?" has been a prominent one in national policy debate. Policymakers on both sides of the political spectrum have questioned our preparedness and whether more should be done to ensure that the country is prepared to respond appropriately to future events. Significant sums of money have been spent in the effort to achieve more and better preparedness. The country has executed a massive government reorganization to produce the Department of Homeland Security and has spent billions of dollars at all levels inside and outside government to buy new equipment, write new emergency-preparedness plans, and hold exercises and drills.

However, when Hurricanes Katrina and Rita struck in 2005, three years into this national effort to strengthen preparedness, it was clear that the investments that had been made to that time had not produced all the results the country expected. The poor response to the hurricanes—at all levels—demonstrated that there was more to do to ensure we were prepared for disasters. How prepared are we now? Three years have now passed since Katrina and Rita struck the Gulf Coast. More after-action assessments have identified ways to strengthen the national response system, more money has been spent, and more preparedness exercises have been planned and run. The nation is almost certainly more prepared in many ways now than it was then. But in spite of these efforts, the honest answer is that we do not fully know if we are prepared for future disasters.

Given the place of homeland security at all levels of recent policy and political debate, questions of preparedness have been both contentious and partisan. The number of political clashes in recent years concerning preparedness might lead readers to expect any number of conclusions or recommendations in a paper that begins with the statement "we do not know if we are prepared enough." Some may expect a critique of current preparedness efforts and arguments about whether the money that has been spent has been used wisely and effectively. Others may expect the argument that there are threats those preparedness programs have not considered, leaving the nation vulnerable to attack or disaster. Some may expect an argument that because of the nature of the threats the country faces, we can "never know" if we are sufficiently prepared until disaster strikes and we will therefore "always need more" investments in preparedness.

While these are certainly topics worth exploring and can lead to relevant questions for policy analysis and debate, this discussion addresses the more fundamental issue of what is needed to approach the question "are we prepared enough?" in a way that can be actually answered, and answered in a way that can contribute to guiding decisionmaking about building a preparedness system that can actually do what the nation expects. *The breakdowns that*

occurred after Katrina and Rita demonstrated not only capability and effectiveness shortfalls, but also that we did not know how prepared we actually were as a country. Policymakers and the public simply did not have the information needed to judge how the preparedness system was likely to perform.

Even today, decisionmakers still must assess preparedness and adjust policies largely "in the rearview mirror," looking at performance in actual events and responding to perceived failures.[1] While learning from real-world experience is important, decisionmakers need better ways to assess preparedness *prospectively* to make better choices as to how and where to strengthen it. The country also needs better ways to assess preparedness levels so citizens can set reasonable expectations about the performance of national, state, and local response systems and can make judgments about how confident they should be that the system will be able to deliver when they need it.

Developing reasonable ways to define such goals for preparedness efforts and calibrate the expectations of decisionmakers, emergency responders, and the public is also important because, looking at a response operation after the fact, there will almost always be "more that could have been done" or "things that could have been done better." While identifying possible areas of improvement is useful, judging preparedness efforts against that yardstick would not only be unfair to emergency planners, it also could lead to poor public policy. The right question is not "could more have been done?" but "how well did the system perform given what was expected from it and the investments that were made based on those expectations?" Given finite resources, the nation will never be able to do (and preparedness efforts should not be expected to provide for) everything that could possibly be done.

Given the scale of the challenge and the relative brevity of this discussion, the goal here is not to provide final answers or prescriptions for how preparedness can be accomplished. As will become quickly clear, there are many unanswered questions about how preparedness can be meaningfully measured. But to get to good answers, it is important to make sure the right questions are being asked. This paper is intended to help frame this broader set of questions and lay out some of the ingredients needed to answer them. To simplify the discussion, it focuses on response activities[2]—the near-term actions taken by responder organizations when a disaster or terrorist incident is occurring or in its immediate aftermath that are intended to limit its consequences—even though the term *preparedness* is frequently used more broadly to encompass more than just these very time-sensitive actions. It also provides some background on the national preparedness system and on current approaches for assessing emergency preparedness and then introduces the concept of *response reliability*, an alternative way of thinking about measuring preparedness that can answer the public's and policymakers' fundamental question: How certain should we as a nation be that the systems we have put in place to respond to damaging events will be able to deliver when called upon?

---

[1] See, for example, the discussion on preparedness outcomes measures in William O. Jenkins, "Homeland Security: DHS Improved Its Risk-Based Grant Programs' Allocation and Management Methods, but Measuring Programs' Impact on National Capabilities Remains a Challenge," Government Accountability Office, Testimony Before the Subcommittee on Homeland Security, Committee on Appropriations, House of Representatives, March 11, 2008.

[2] See discussion of these different mission areas and the tasks included in Department of Homeland Security, *Target Capabilities List: A Companion to the National Preparedness Guidelines*, September 2007.

CHAPTER TWO

Background

Emergency preparedness is fundamentally about being ready to take action when a damaging event happens—whether that event is a fire, a flood, or a terrorist attack. At the national level, the focus is predominantly on very large incidents—major disasters or terrorist attacks with far-reaching consequences. To answer whether or not the nation is prepared therefore requires identifying the types of damaging events that may happen, what we want to be able to do in response to them, and what capabilities and resources are needed to conduct the response. If a major earthquake occurs, we want to be able to evacuate the area affected and treat and care for the victims; to do so we will need transportation resources, medical supplies and responders, and facilities to provide shelter and food until the damage is addressed.

Whether we are prepared also depends on other characteristics of our desired response—whether the goals we are trying to achieve are limited or broad. The "higher quality" we want our response to be—i.e., resources that are able to get to the scene faster, to help more people, and address the needs of larger incidents or disasters—the more difficult it will be to prepare. In thinking about the quality of response activities, it is useful to begin with the much simpler case of "everyday" emergencies. Because fires in buildings generally get worse and do more damage the longer they are allowed to burn, we want fire trucks on the scene within a reasonable time and we want enough of them there to extinguish the size fires we are concerned about.[1] How much capability we want on the scene within a certain time will drive the answers to questions such as *How many fire stations and trucks are required? Where should they be positioned?* and *How should they be staffed?* The scope of the desired response is important, too. A fire-protection system that can handle ten fires burning simultaneously in an area would look very different than one that can handle only a single fire at a time.[2]

---

[1] Though response time is one factor related to success, faster is not always better. The effect of response time will differ from event to event and its significance will be driven in part by the time required for other elements of the response system to act. See, for example, discussions about this issue with respect to fire, emergency medical services, and law enforcement activities in RAND Fire Project, *Fire Department Deployment Analysis: A Public Policy Analysis Case Study,* New York: Elsevier, 1979, p. 81; Jonathan D. Mayer, "Response Time and Its Significance in Medical Emergencies," *Geographical Review*, Vol. 70, No. 1, January, 1980, pp. 79–87; and David Weisburd and John E. Eck, "What Can Police Do to Reduce Crime, Disorder, and Fear?" *Annals of the American Academy of Political and Social Science,* No. 593, 2004, pp. 42–65, respectively.

[2] Defining what outcomes response activities were designed to achieve was a key element of RAND's classic work on fire department operations in the 1970s (see, for example, RAND Fire Project, *Fire Department Deployment Analysis: A Public Policy Analysis Case Study,* New York: Elsevier, 1979). Though these issues are sometimes much more difficult for broader preparedness and homeland security problems, the same thought process is relevant for laying out what preparedness efforts are seeking to accomplish and identifying what is needed to do so.

Identifying the desired outcomes for responses to larger and more-complicated disaster situations is more difficult, but questions of speed, scope, and scale are critical:[3] "Prepared to feed all the citizens of a major city" would look very different from "prepared to feed everyone from a flooded residential subdivision." For major incidents, there will also be many things we want to accomplish simultaneously, including feeding and sheltering victims, treating injuries, stabilizing the hazardous situations, restoring the functioning of damaged infrastructures, and so on. While some straightforward and objective criteria can help to guide these decisions—e.g., food supplies need to be moved into a disaster area quickly enough that the victims they are intended to feed have not either gone hungry or left the area before the supplies have arrived—others will have to be guided by preferences and values as to what qualifies as the "right" amount of support and relief. But, even if what we want to be able to accomplish after a disaster strikes is a policy choice, that choice must be made before measures of our preparedness can even be conceptualized. The choices we make about what must be done drive the mixtures of capabilities and resources that need to be in place[4]—and therefore define the standards against which preparations must then be measured.[5]

The National Preparedness System

Major changes have been made in national thinking about emergency preparedness since the creation of the Department of Homeland Security. Planning efforts transitioned from the legacy *Federal Response Plan* through the *National Response Plan* to the *National Response Framework*.[6] The National Incident Management System was developed and its implementation was propelled by federal preparedness funding.[7] Implementing Homeland Security Presidential Directive–8, the Department of Homeland Security laid the foundations for a national preparedness system with a set of planning scenarios, guidelines, and the application of capabilities-based

---

[3] RAND has carried out such analyses for some types of disaster response, including Tom LaTourrette, Edward W. Chan, Jennifer Brower, Jamison Jo Medby, and K. Scott McMahon, *Emergency Response to Terrorist Attacks: An Analysis of Mission Performance Requirements,* Santa Monica, Calif.: RAND Corporation, MG-298-RC, 2006, not available to the general public.

[4] Analogies to this element of preparedness—how complementary inputs have to come together at the right time and in the right proportions for the overall effort to be successful—can be found in past efforts to assess military readiness and sustainability. See, for example, S. Craig Moore, J. A. Stockfisch, Matthew S. Goldberg, Suzanne M. Holroyd, and Gregory G. Hildebrandt, *Measuring Military Readiness and Sustainability,* Santa Monica, Calif.: RAND Corporation, R-3842-DAG, 1991.

[5] Homeland security and emergency preparedness are not the first areas in which these types of questions have been raised. In military planning—which can be viewed as "preparedness for different types of conflicts"—the questions of what to plan for and what it means to be prepared have been a key element of debate and policy analysis surrounding defense planning. A classic document in this area is Alain C. Enthoven and K. Wayne Smith, *How Much Is Enough? Shaping the Defense Program, 1961–1969*, New York: Harper & Row, 1971, which was recently reissued by RAND. Since then, a variety of ideas has been used to shape what military policies were designed to "prepare for," ranging from broad ideas, such as the ability to fight one or more major wars simultaneously, to more-specific ideas regarding the different types of specialized contingencies in which defense organizations should be ready to intervene.

[6] *Federal Response Plan*, 9230.1-PL, January 2003; Department of Homeland Security, *National Response Plan*, December 2004; Department of Homeland Security, *National Response Framework*, January 2008.

[7] Department of Homeland Security, *National Incident Managment System*, December 2004. A revision of this document was in process at the time of this writing.

planning approaches drawn from defense planning.[8] One of the major goals of these efforts was to frame a national approach for thinking about preparedness for large-scale incidents, and to do so in a way that was compatible with planning for smaller-scale incidents and the variation in hazard exposures and requirements across the country.

Within the National Preparedness System, documents such as the *Universal Task List* and *Target Capabilities List* catalogue activities that the national response system should be capable of accomplishing in response to very large incidents.[9] They range from broadly applicable capabilities, such as providing mass care and feeding to the affected population, to specialized roles, such as isolation and quarantine for events involving communicable diseases.[10] Based on planning assumptions drawn from the National Planning Scenarios,[11] these documents also include some specifications of how much of different capabilities should be in place nationally and how quickly they must be available when a disaster strikes. Using these frameworks to set priorities for making decisions about resource investments in homeland security—e.g., setting the goals of federal grant funding provided to states and localities—and activities such as exercises as part of the Homeland Security Exercise and Evaluation Program links them to future efforts to improve and assess preparedness.[12]

How Is Preparedness Assessed Now?

Given concerns about preparedness that have persisted over the last seven years, there have been a variety of efforts to assess preparedness overall and to measure preparedness for particular types of incidents. These have included both the development of standards to provide

---

[8] For a description of the various pieces of the National Preparedness System, see Department of Homeland Security, "National Preparedness Guidelines," Web page, September 11, 2008 (as of October 23, 2008: http://www.dhs.gov/xprepresp/publications/gc_1189788256647.shtm). The approach being applied in these national-level preparedness efforts draws on the concepts of capabilities-based planning that were developed in the military community to assist in planning for a range of scenarios under uncertainty. See, for example, Paul K. Davis, *Analytic Architecture for Capabilities-Based Planning, Mission Systems Analysis, and Transformation,* Santa Monica, Calif.: RAND Corporation, MR-1513-OSD, 2002.

[9] Office of Domestic Preparedness, Department of Homeland Security, *Universal Task List 2.0*, December 21, 2004; Department of Homeland Security, *Target Capabilities List: A Companion to the National Preparedness Guidelines*, September 2007. Beyond the policy documents themselves, recent versions of which are available from the Department of Homeland Security, a review of these elements of the department's efforts in the broader context of the department's activities is available in William O. Jenkins, "Emergency Preparedness and Response: Some Issues and Challenges Associated with Major Emergency Incidents," Testimony Before the Little Hoover Commission, State of California, GAO-06-467T, February 23, 2006.

[10] Efforts have been undertaken by others to define preparedness in specific subset fields or disciplines as well. For example, see Christopher Nelson et al., "Conceptualizing and Defining Public Health Emergency Preparedness," *American Journal of Public Health* 97, No. S1 (2007): S1–S5, for an example breakdown for public health activities.

[11] The National Planning Scenarios are a set of large-scale incidents of national concern that were chosen as targets for preparedness.

[12] For example, a recent Federal Emergency Management Agency briefing identifies the goal of the next National Preparedness System development as moving from a system enabling preparedness management to one that can evaluate effectiveness (Federal Emergency Management Agency, "National Preparedness System: Building a Nation Prepared," briefing at USACE/FEMA Senior Leadership Seminar 2008, April 8, 2008).

guides for assessing preparedness activities and the gathering of data about both current preparedness levels and ongoing initiatives for improving them.[13]

As discussed above, the National Preparedness System has included activities focused on developing standards, measures, and metrics for preparedness. Other efforts, some of which predate the formation of the Department of Homeland Security and the 2001 terrorist attacks, produced standards for components of preparedness programs.[14] A number of regulatory standards define requirements for some elements of emergency preparedness and response,[15] and a number of nongovernmental organizations also play major roles in standards development. Examples of these organizations include a variety of groups associated with health care and practice and nonprofit organizations such as the National Emergency Management Association and the National Fire Protection Association (whose *Standard on Disaster/Emergency Management and Business Continuity Programs*[16] has been broadly adopted). Although many of these standards lay out accepted criteria and frameworks for planning processes, in general, they do not include detailed guidance to fully evaluate the results of those planning efforts and to assess whether or not an area is prepared. They have also been characterized as "qualitative" and focusing on predominantly on written plans.[17]

Other assessment efforts focus on counting the ingredients of preparedness that are most readily countable. To be successful, a geographic area needs to have a plan in place to respond, needs to have the equipment that the plan calls for, and must have enough people to perform all the roles needed to implement it. Evaluation efforts therefore frequently look for a shortage or absence of key ingredients, since we know that if major pieces are missing, response efforts are likely to break down. Common indicators include programmatic measures, such as whether there is a plan covering all the incidents or disasters of concern, counting how many responders have completed particular training programs, assessing stocks of key equipment and supplies, counting available response workers, and so on. These types of assessments are frequently implemented as checklists[18] that response organizations (or external bodies) can use

---

[13] A variety of these programs and the issues associated with them were reviewed by the National Academies Institute of Medicine in an assessment of the Metropolitan Medical Response System that was published in 2002 (Committee on Evaluation of the Metropolitan Medical Response System Program, Board on Health Sciences Policy, Institute of Medicine, *Preparing for Terrorism: Tools for Evaluating the Metropolitan Medical Response System Program,* Fredrick J. Manning and Lewis Goldfrank, eds., Washington, D.C.: National Academies Press, 2002). The following discussion draws on this framework in considering preparedness evaluations presented in that work.

[14] For a review, see Ben Canada, "Homeland Security: Standards for State and Local Preparedness," RL31680, Congressional Research Service, January 2, 2003.

[15] For example, the Occupational Safety and Health Administration's Hazardous Waste Operations and Emergency Response regulation (29 CFR 1910.120).

[16] National Fire Protection Association, *Standard on Disaster/Emergency Management and Business Continuity Programs*, NFPA-1600, 2004.

[17] "With only a few exceptions, the committee deemed these standards to be of limited utility in assessing the preparedness of local communities for coping with a CBR [chemical, biological, radiological] terrorist event . . . . Most of the standards . . . are qualitative in nature [and] most of them also focus on the adequacy of written plans." (Committee on Evaluation of the Metropolitan Medical Response System Program, Board on Health Sciences Policy, Institute of Medicine, *Preparing for Terrorism: Tools for Evaluating the Metropolitan Medical Response System Program*, Fredrick J. Manning and Lewis Goldfrank, eds., Washington, D.C.: National Academies Press, 2002, pp. 93–94).

[18] Examples of such checklist-type approaches include the Federal Emergency Management Agency's Capability Assessment for Readiness tools, some of the instruments developed by the Centers for Disease Control and Prevention for assessing public health preparedness, tools produced for the National Disaster Medical System, various Department of Homeland

for evaluation purposes or as surveys[19] to gather broader cross-sectional data on the status of preparedness in a state, in a region, or in the nation overall.

More-recent efforts include the development of measures and metrics that focus not just on inputs that are in place but the capabilities those inputs are intended to deliver. The *Target Capabilities List* similarly includes a variety of preparedness and performance measures and metrics, which are a combination of checklist-like assessments (e.g., yes/no judgments about whether a plan is in place) and quantitative measures of performance.[20] Recent efforts focused on public health preparedness have similarly moved toward building approaches for assessing capability levels. Although comprehensive information on current data collection and assessment efforts within the National Preparedness System is not available publicly, efforts based on frameworks from the *Target Capabilities List* and other documents are clearly in progress.[21] A recent example of a more in-depth assessment effort is the Department of Homeland Security's "Nationwide Plan Review Phase 2 Report," which focuses on preparedness planning for dealing with hurricanes and evacuations in response to the breakdowns encountered in the Hurricane Katrina and Rita response.[22]

Although checklists and survey methods are useful for cataloguing elements of preparedness, they have key shortcomings. A challenge with mechanisms that focus on counting resources and verifying that processes have been completed is that there is a major risk of focusing on verifying *quantities* of resources or activities and ignoring the potential differences in *quality* that may actually be more important to eventual success in response operations. While two areas may both have a preparedness plan, there is a major difference between a high-quality plan that was carefully assembled and one that was produced merely to satisfy a

---

Security processes for the collection of information from organizations at the state and local level, and approaches for auditing used by internal government inspectors general for program assessment (see Federal Emergency Management Agency, "State Capability Assessment for Readiness," Report to the United States Senate Committee on Appropriations, December 10, 1997; "Jurisdiction Handbook," State Homeland Security Assessment and Strategy Process, fiscal year 2003, and other resources available online at http://www.shsasresources.com/; Office of the Inspector General, Department of Homeland Security, "Audit of the National Urban Search and Rescue Response System," OIG-06-54, August 2006; reviews in Committee on Evaluation of the Metropolitan Medical Response System Program, Board on Health Sciences Policy, Institute of Medicine, *Preparing for Terrorism: Tools for Evaluating the Metropolitan Medical Response System Program*, Frederick J. Manning and Lewis Goldfrank, eds., Washington, D.C.: National Academies Press, 2002, pp. 94–95; and Harry A. Mayer, "First Responder Readiness: A Systems Approach to Readiness Assessment Using Model Based Vulnerability Analysis Techniques," masters thesis, Naval Postgraduate School, Monterey, Calif., September 2005, pp. 25–39.)

[19] For example, surveys carried out by RAND in support of the Advisory Panel to Assess Domestic Response Capabilities for Terrorism Involving Weapons of Mass Destruction ("the Gilmore Commission") and reported in the commission's reports (available online at http://www.rand.org/nsrd/terrpanel/); Lois M. Davis, Louis T. Mariano, Jennifer E. Pace, Sarah K. Cotton, and Paul Steinberg, *Combating Terrorism: How Prepared Are State and Local Response Organizations?* Santa Monica, Calif.: RAND Corporation, MG-309-OSD, 2006); Bellwether Group, Inc., "Disaster and Terrorism Preparedness: An Examination of the State of Corporate Preparedness in the US Among Eight Leading Corporations," December 2005; James Graham, Steve Shirm, Rebecca Liggin, Mary E. Aitken, and Rhonda Dick, "Mass-Casualty Events at Schools: A National Preparedness Survey," *Pediatrics,* Vol. 117, No. 1, January 2006, pp. e8–e15; Kevin Jack Riley and Bruce Hoffman, *Domestic Terrorism: A National Assessment of State and Local Preparedness*, Santa Monica, Calif.: RAND Corporation, MR-506-NIJ, 1995.

[20] Department of Homeland Security, *Target Capabilities List: A Companion to the National Preparedness Guidelines*, September 2007.

[21] See, for example, Charles Stone, "The National Preparedness System (NPS): Moving Preparedness into a Net Centric Environment," briefing at the Net Centric Operations Conference, March 5–8, 2007 (as of October 23, 2008: www.dtic.mil/ndia/2007netcentric/Stone_NPS_NetCentric.pdf).

[22] Department of Homeland Security, "Nationwide Plan Review Phase 2 Report," June 16, 2006.

requirement and that has not been looked at since it was written. Similarly, even if staff members were trained in previous years on what would be required to respond to rare and unusual incidents (e.g., bioterrorism incidents involving exotic diseases), they may not retain enough of the knowledge for that training to have any real effect on how they might respond to such an event in the future. Turnover in key staff can similarly result in knowledge "draining away" over time and preparedness falling off.[23]

Some preparedness measurement efforts have tried to assess some of the intangible inputs and factors that affect performance. For example, in examining public health preparedness, some efforts have asked about how the experience of leaders and staff, gathered data on the triggers for activating certain kinds of plans,[24] or asked specifically whether individual organizations can perform stated tasks.[25, 26] Some seek to link the measurement of plan components to evidence as to whether the activities could be executed and to existing data that support the position that the capabilities listed in the inventory are "real" rather than just existing "on paper." Nonetheless, checklist and inventorying mechanisms are most valuable for tabulating the *inputs* needed to respond to particular events, but they will not necessarily capture whether the preparedness system has the capability to use those inputs to achieve the desired response *outcomes* when an event actually occurs.[27]

In one sense, stocks of supplies and equipment, available staff, and a plan for utilizing them can be thought of as defining a "theoretical maximum" performance for the system—what could be done if all the inputs were used most effectively and efficiently when disaster strikes. If there are 100 breathing apparatuses for operations in an environment contaminated with hazardous materials, only 100 responders could be involved in operations at a specific time. However, staff being unfamiliar with the plan for such response, a lack of necessary training for using the equipment, key supplies like filters for the equipment being out of date, or other shortfalls in the *quality* of preparedness efforts can reduce the level of performance that can realistically be expected to well below that ideal level.[28]

---

[23] A notable example of this potential disparity was documented in RAND work assessing public-health preparedness and the ability of jurisdictions to respond to calls regarding cases of particular concern on a 24/7 basis. "All but one local health department in a convenience sample of 19 indicated that they had a process in place to receive and respond to calls on a 24/7 basis. However, operational assessments demonstrated that only 2 of 19 local health departments consistently met the prescribed standard." (Christopher Nelson, Nicole Lurie, and Jeffery Wasserman, "Assessing Public Health Emergency Preparedness: Concepts, Tools, and Challenges," *Annual Review of Public Health,* Vol. 28, 2007, p. 7).

[24] See, for example, Centers for Disease Control and Prevention, Public Health Practice, "State and Public Health Preparedness and Response Capacity Inventory: A Voluntary Rapid Self-Assessment," Version 1.1, December 2002.

[25] Trust for America's Health, "Public Health Laboratories: Unprepared and Overwhelmed," June 2003.

[26] For a review of a variety of written public health assessments, including discussion of the complications inherent in assessing their results, see Christopher Nelson, Nicole Lurie, and Jeffery Wasserman, "Assessing Public Health Emergency Preparedness: Concepts, Tools, and Challenges," *Annual Review of Public Health,* Vol. 28, 2007, pp. 1–18.

[27] See, for example, discussion in Committee on Evaluation of the Metropolitan Medical Response System Program, Board on Health Sciences Policy, Institute of Medicine, *Preparing for Terrorism: Tools for Evaluating the Metropolitan Medical Response System Program*, Frederick J. Manning and Lewis Goldfrank, eds., Washington, D.C.: National Academies Press, 2002.

[28] For a more extensive discussion of this difference between theoretical and actual performance in the context of assessing the readiness of response units, see Harry A. Mayer, "First Responder Readiness: A Systems Approach to Readiness Assessment Using Model Based Vulnerability Analysis Techniques," masters thesis, Naval Postgraduate School, Monterey, Calif., September 2005.

In most such assessments, there is also tension between the assessment of individual elements of preparedness in isolation (e.g., an explicitly tactical-level focus on specific response activities) and a broader, more strategic-level view of how different elements function together as part of response activities. Where efforts seek to measure likely performance in future events, the focus of assessment is frequently on one individual response function at a time rather than on providing an overall framework for "rolling up" detailed measures in a way that can provide an evaluation of likely overall response effectiveness in a way that is useful for high-level response planners and decisionmakers.

The limits of many of the means of assessing preparedness have led to interest in the use of exercises—activities that range from abstract table-top events at which leaders from response organizations talk through what would happen if a particular incident occurred to very concrete operations where responders actually deploy to simulated events and physically practice the tasks they would need to perform to actually respond. In principle, a well-designed, sufficiently realistic exercise could fully test response activities and evaluate how organizations would perform in a real event—and adequately assess individual response capabilities or actions in isolation. As a result, whether or not a plan has been exercised is frequently used as a proxy measure for assessing its preparedness value.

In practice, whether or not exercises are useful as a measure of preparedness depends on how they are designed.[29] Exercises are versatile tools that can be used for a number of purposes, including assessment, policy development, individual and organizational training, and multiagency coordination planning.[30] Frequently, exercises are expected to play multiple roles—e.g., training staff, helping craft policy, and assessing preparedness—which are not always compatible and may require significant differences in exercise design. Furthermore, there are limits to how realistic any exercise can be—conducting a fully realistic exercise for a major disaster would be prohibitively costly, both financially and because of the disruption it would cause in the area where it was held. Established and standardized practices for extracting data from exercises so their output can be useful for making broader statements about preparedness outside the specific scenario that was the basis for the effort or about preparedness writ larger is also a concern.[31]

Because of the costs of large-scale events, often, smaller-scale exercises are held that only cover "pieces" of response activities (e.g., evacuating one large building or facility rather than

---

[29] For an overview of different types of exercises and their strengths and limitations for assessing preparedness levels, see Christopher Nelson, Nicole Lurie, and Jeffery Wasserman, "Assessing Public Health Emergency Preparedness: Concepts, Tools, and Challenges," *Annual Review of Public Health,* Vol. 28, 2007, pp. 1–18.

[30] See, for example, discussion in Homeland Security Exercise and Evaluation Program, "Terminology, Methodology, and Compliance Guidelines," undated. RAND has used exercises in efforts aimed at the full range of these purposes. See, for example, Marc Dean Milot, Roger C. Molander, and Peter A. Wilson, *"The Day After . . ." Study: Nuclear Proliferation in the Post–Cold War World,* Volume II*, Main Report*, Santa Monica, Calif.: RAND Corporation, MR-266-AF, 1993; Brian A. Jackson, James W. Buehler, Dana Cole, Susan Cookson, David J. Dausey, Lauren Honess-Morreale, Susan Lance, Roger C. Molander, Patrick O'Neal, and Nicole Lurie, "Bioterrorism with Zoonotic Disease: Public Health Preparedness Lessons from a Multi-Agency Exercise," *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science,* Vol. 4, No. 3, 2006, pp. 287–292; and David J. Dausey, Nicole Lurie, Alexis Diamond, Barbara Meade, Roger C. Molander, Karen A. Ricci, Michael A. Stoto, and Jeffrey Wasserman, "Bioterrorism Preparedness Training and Assessment Exercises for Local Public Health Agencies," TR-261-DHHS, Santa Monica, Calif.: RAND Corporation, TR-261-DHHS, 2005.

[31] For example, considerable guidance is provided by the Homeland Security Exercise and Evaluation Program on exercise design and evaluation, but the focus is on evaluation and data collection within the context of individual exercises and scenarios (see https://hseep.dhs.gov/).

the entire city). While scaling down in this way is useful for containing the cost and disruption that exercises can cause, it is not necessarily the case that if an organization can run an evacuation of 50 people using one bus, it is prepared to scale up that operation a thousand fold to empty residents from a medium-sized city. The artificialities that exist in many exercises can mean that an exercise that is very useful for training purposes or to identify some shortcomings of preparedness efforts may actually provide little insight into how an organization or group of response agencies will perform during an actual disaster or contingency.

CHAPTER THREE

Potential Options and Solutions

Although the unpredictable nature of emergency situations makes it difficult to know for certain if everything will go as planned until after a specific incident has occurred,[1] relying only on "seeing what happens when a crisis hits" to assess preparedness efforts cannot provide all the ingredients needed to craft good homeland security policy. Policymakers and the public need to be able to measure how well the system is prepared to perform, not just watch how well it performs after the fact. Though current measurement approaches provide information on many of the key inputs to preparedness and have made some progress toward outcome assessment, they do not make it possible to reasonably anticipate response-system performance before events actually occur and preparedness is tested against reality. This traps policymakers in a reactive mode, always responding to the shortcomings of past responses rather than making proactive decisions to shape homeland security efforts. Being able to do these measurements is also necessary to account for resources that have already been spent on preparedness efforts, to assess what they have accomplished, and to identify where additional investment may be needed to meet public expectations for performance after disasters.

Rather than just asking if a plan has been written or an exercise held, policymakers and citizens should be asking a different and much broader question: *How confident should we be that the systems we have put in place to respond to damaging events will be able to deliver when we call on them?* The answer to this question requires going a step beyond just asking what we want to be able to do after an incident and what resources and capabilities are needed to do so. It does not matter if supplies have been bought or responders trained if, when a disaster hits, they cannot get to the scene to do their jobs. Asking how confident we should be that we can achieve desired response outcomes is a systems question: It requires knowing how the system of organizations, capabilities, and resources we expect to respond works and the factors that shape how well it does so, and then assessing how reliably it can operate under the demands of different types of damaging events.[2]

[1] See, for example, discussion in Richard A. Falkenrath, "The Problems of Preparedness: Challenges Facing the U.S. Domestic Preparedness Program," Cambridge, Mass.: John F. Kennedy School of Government, Harvard University, December 2000, pp. 15–16.

[2] For an analogous application of these types of ideas, see Shari Welch and Kirk Jensen, "The Concept of Reliability in Emergency Medicine," *The Journal of Medical Quality*, Vol. 22, No. 1, January/February 2007, pp. 50–58.

An Example Reliability Assessment

The importance of assessing the reliability of response systems can be illustrated with a simple example from evacuation planning. In a hypothetical emergency plan, a city builds its capability to evacuate its citizens on buses in case of a major disaster. Counting the number of seats on a bus, one might conclude that it would be possible to evacuate the residents of a small apartment building in one bus.[3] But, if that bus breaks down as a hurricane approaches the city, then it will be impossible to transport *any* of those people to safety. If no backup buses are available, the possibility that the bus will break down is a threat to response success at that site. If the chance of the bus not functioning is 20 percent, then, at best, the response plan will be 80 percent reliable.[4]

Assessing response reliability requires determining what might go wrong and anticipating what the impact of particular events would be on the success of the operation.[5] In some cases, breakdowns will derail the entire response effort—they will cause catastrophic failures where there are limited options for adaptation or improvisation to reconstitute capabilities and effectiveness. In others, they may just delay when response can be initiated, limit the number of people who can be served, or reduce the effectiveness of the operation—i.e., affect the quality of the response. If an area's evacuation plan included four buses, the loss of one would not halt response entirely, but it would reduce response capacity by 25 percent, requiring more time to evacuate the same number of people, all other factors being equal. Depending on the circumstances, a delay could equate to complete failure (e.g., if an evacuation was delayed so much that it could not begin before an approaching hurricane struck the city), but in other circumstances the consequences will be far less serious. For noncatastrophic failure modes, there may also be options for responders to dynamically adjust "on the fly" and find ways around the failure. Depending on if and how rapidly such adjustments can be made, the overall impact of a breakdown could be reduced to a minimal level. Understanding the scope of potential consequences of different response breakdowns is important. In most cases, faults that would result in failure of the entire response pose much more risk—and therefore merit greater remedial attention—than those that would just reduce total capacity or effectiveness.

Response breakdowns can be caused by the nature of the incident itself, particularly for local response organizations that are based in the affected area. High winds might strip down the antenna systems on which communications systems rely, making it impossible to dispatch response resources effectively after the storm has passed. If an incident affects members of response organizations or their families, staff critical to executing response plans may be unavailable. Alternatively, flooding or an earthquake might damage road infrastructures, hindering movement around a disaster area or preventing external aid from reaching the areas that

---

[3] See, for example, analogous discussion in the "Citizen Evacuation and Shelter-in-Place" response mission described in Department of Homeland Security, *Target Capabilities List: A Companion to the National Preparedness Guidelines*, September 2007, pp. 377–393.

[4] The chance of failure is *at least* one in five because other factors could increase it beyond just the risk posed by equipment failure. This is discussed in more detail below.

[5] For a broader discussion of this type of analysis, see Harry A. Mayer, "First Responder Readiness: A Systems Approach to Readiness Assessment Using Model Based Vulnerability Analysis Techniques," masters thesis, Naval Postgraduate School, Monterey, Calif., September 2005.

need it. In such cases, the location of key response resources and their robustness to the types of incidents that might strike the area will be a key contributor to the reliability of response.[6]

Though a hurricane wiping out a key communication system is a very tangible event that could clearly hurt effectiveness, response reliability can be hurt by much less tangible breakdowns, such as shortcomings in management and coordination. For example, in large events, capabilities must frequently be drawn from multiple response organizations both locally and across the country.[7] While national planning and standardization efforts (such as the development of the National Incident Management System) are designed to make it possible for different organizations to work well together, often cooperation and coordination do not go smoothly and key functions "fall through the cracks" between organizations. After-action reports assessing the response to Hurricanes Katrina and Rita highlighted such problems as important drivers of the poor performance of the response to those storms.[8] If an effective response to a large event relies on components that must be brought together and integrated from many different response organizations—from many federal agencies, organizations from multiple levels of government, private-sector entities, and nongovernmental organizations—how well those interorganizational links function will be a major determiner of how confident the public can be that the preparedness system will be able to deliver what is called for in disaster-response plans.[9]

An illustration of the type of fault analysis that is needed to assess the reliability of response operations is shown in Figure 3.1, which illustrates the basic example of a centralized bus evacuation of a population, intended to safely transport a population out of areas where they are in

---

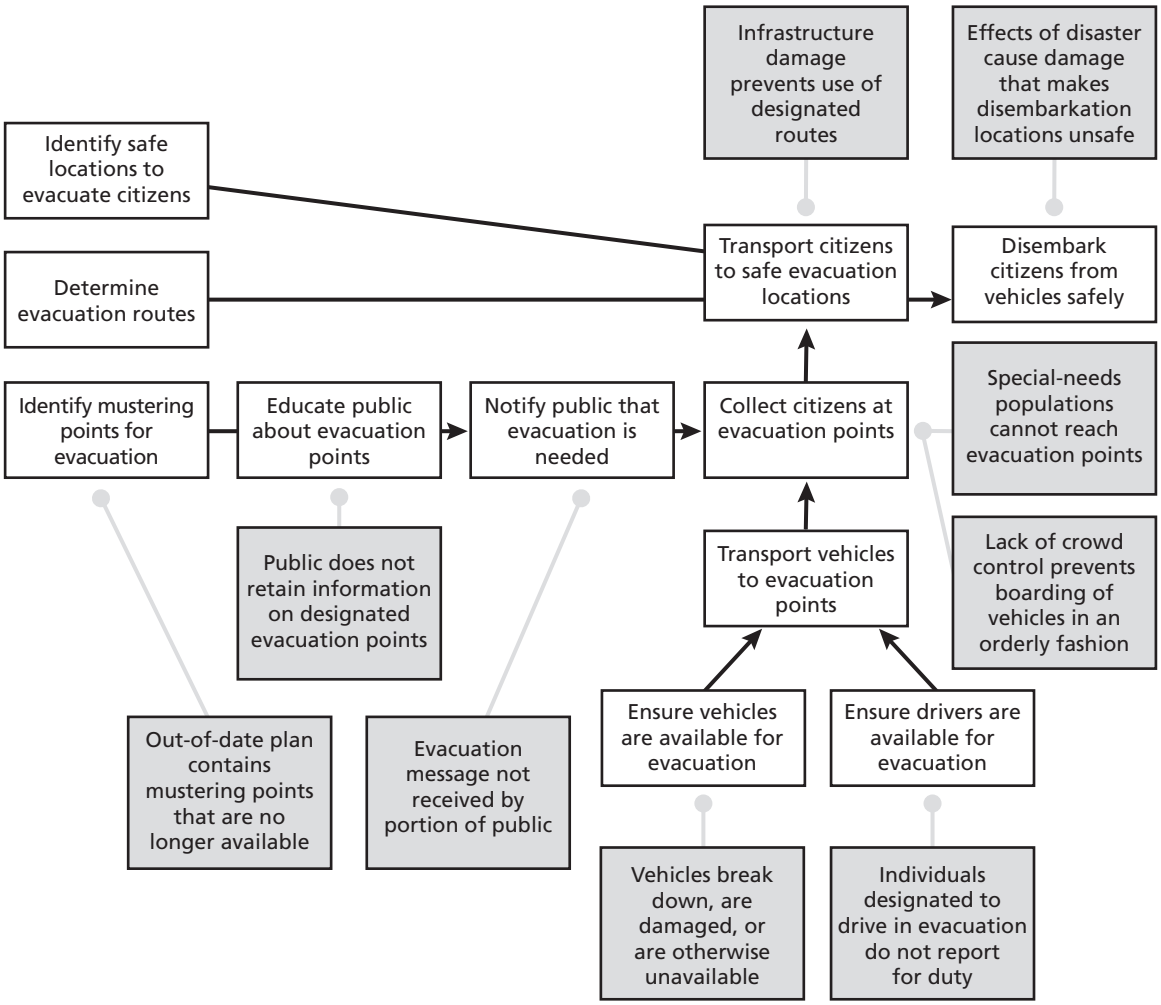
[6] All of these "failure modes" for response systems were observed in case studies of past disaster-response operations done by RAND as part of a larger project examining the specific issue of how responder safety is managed in large-scale incident operations. The results of that study are reported in Brian A. Jackson, John C. Baker, M. Susan Ridgely, James T. Bartis, and Herbert I. Linn, *Protecting Emergency Responders,* Volume 3: *Safety Management in Disaster and Terrorism Response,* Santa Monica, Calif.: RAND Corporation, MG-170-NIOSH, 2004.

[7] Disaster researchers have identified this breakpoint—where the conditions are such that requirements go beyond those of any single response organization and multi-agency cooperation becomes imperative—as the divide between everyday emergencies and disasters or catastrophes (E.L. Quarantelli, "Emergencies, Disasters and Catastrophes Are Different Phenomena," Disaster Research Center, University of Delaware, undated, as of October 13, 2008: http://www.udel.edu/DRC/preliminary/pp304.pdf). See also Randall A. Yim, "National Preparedness: Integration of Federal, State, Local, and Private Sector Efforts Is Critical to an Effective National Strategy for Homeland Security," U.S. General Accounting Office, Testimony Before the Subcommittee on Economic Development, Public Buildings, and Emergency Management, Committee on Transportation and Infrastructure, House of Representatives, GAO-02-621T, April 11, 2002.

[8] See, for example, Select Bipartisan Committee to Investigate the Preparation for and Response to Hurricane Katrina, "A Failure of Initiative," U.S. Government Printing Office, February 15, 2006; The White House, "The Federal Response to Hurricane Katrina: Lessons Learned," February 2006.

[9] Assessing the strength and potential effectiveness of the links between separate organizations before an event occurs is clearly an analytically challenging prospect, but such an assessment could be approached from a number of different directions. More-countable measures include formal agreements among those organizations, the frequency of joint exercises, past history of (effective) participation in joint response operations, and so on. Because the individuals who are part of organizations and whose relationships provide the links between them are important, social-network–type analyses of key personnel interfaces between them could also contribute. Finally, more-qualitative survey or other techniques that evaluate both past experiences and future expectations of the effectiveness of interactions could also play a part. It is immediately clear that these kinds of analyses would be most tractable for a small number of organizations within a defined geographic area; complexity increases combinatorially up to national-level responses. Studies would be required to explore how many and what types of assessments would be most valuable for these higher-level response operations while controlling the costs imposed by large-scale assessments.

Figure 3.1
Example Fault Analysis for Identifying Risks to Evacuation Reliability



RAND *OP234-3.1*

danger to locations out of harm's way. The general steps of the response operation are shown in the white boxes, from required planning tasks (e.g., identifying mustering points for evacuation) through the steps needed to actually trigger the evacuation and carry it out.[10] Possible events that could occur that might disrupt the operation are included in the gray breakout boxes linked to the steps of the response. Rather than being exhaustive, the events shown are intended to illustrate some of possible elements that would have to be considered in assessing evacuation plans.[11] For example, the first identified failure mode is that, due to planning being

[10] These steps are analogous to the descriptions, tasks, measures, and metrics in the "Citizen Evacuation and Shelter-in-Place" response mission described in Department of Homeland Security, *Target Capabilities List: A Companion to the National Preparedness Guidelines*, September 2007, pp. 377–393.

[11] The process of identifying such failure modes would draw on past experience in similar or analogous operations. Lessons learned from previous operations can provide a taxonomy of what can go wrong so the risks to future operations can be identified. An example of such a taxonomy for large-scale operations in particular is presented in John R. Harrald, "Agility

out of date, mustering points that are included in the plan are no longer available (e.g., a parking lot that has since been developed). While this failure would not threaten the entire operation, it would disrupt it and could result in a significant reduction in effectiveness. In contrast, evacuation vehicles being unavailable (e.g., if the site where they were staged was flooded or if a firm contracted to provide them could not deliver when called upon) could cause the entire operation to fail. Other failures risk the quality of the evacuation operation (e.g., breakdowns in crowd control or damage at the designated points for evacuees disembarking result in citizens injured during the evacuation process).

Depending on the characteristics of the city or state preparing an evacuation plan, individual potential failure modes may be more relevant or more likely than others. Evacuation planning for a municipality that maintains its own bus fleet, keeps its equipment in perfect working order, and stores the vehicles in a location protected from most possible disasters would be very different from a town that is counting on vehicles provided by a previously unknown overseas firm whose winning bid for the contract was unrealistically low. The risk that vehicles will not be available when needed is likely much lower in the first case than in the second. The estimated reliability of the second town's plan should therefore be much lower, reflecting that major risk to the operation.

Thinking Through Assessing Response Reliability

Though the concept of response reliability gets directly at what citizens expect from emergency response systems—that they will be there when they are needed and perform predictably when they are called on—that does not mean that conducting such reliability assessments will be easy. For emergency response systems that operate every day, these concepts are straightforward enough to apply: For everyday emergencies, the frequency of breakdown can simply be measured directly. For example, a variable of concern for medical response is that response units arrive on the scene quickly enough to treat patients with time-sensitive conditions such as heart attacks. Since such events happen every day, the percentage of responses where emergency medical services units fail to arrive within the desired time window can be tracked, and changes can be made to fix problems that are detected.[12]

Thinking about assessing the risk of failure for responses to large events is much more difficult because to do so, one must make judgments about the potential performance of systems that will only be called on rarely, if ever. Thinking would have to begin with the building of a basic model for a given response system that is intended to deliver particular capabilities (similar to the basic chart in Figure 1). It also requires defining the needs the response system is intended to address and identifying how the different pieces of that system fit together and depend on one another. Ways that system could break down would then need to be identified, and estimates made of the likelihood of those failure modes arising.

The illustration in this paper was developed around citizen evacuation, but it is easy to see how an analogous process could be undertaken to identify failure modes for activities such

---

and Discipline: Critical Success Factors for Disaster Response," *The ANNALS of the American Academy of Political and Social Science,* Vol. 604, 2006, pp. 256–272.

[12] See, for example, a recent discussion of such an initiative in Atlanta reported in Robert Davis, "Atlanta Becomes a Template for Improving EMS," *USA Today*, August 20, 2007.

as dispensing pharmaceuticals during a disease outbreak, search-and-rescue operations in a flooded area or collapsed building, and so on. For a given large-scale incident, the relationships among and between different response activities would then need to be identified to see if the effort of carrying out a task in multiple locations or simultaneously with other tasks would produce additional potential breakdown points (e.g., from resource shortfalls or limits in available trained staff). Since large-scale events happen rarely, failure modes and estimates of their likelihood would in many cases most likely have to be made based on relevant everyday experience, on performance in realistic drills and exercises, or simply on good-faith estimation to make it possible for policymakers and the public to decide if they are satisfied with the protection the system provides.

To make a formal estimate of the reliability of a response system, estimates of the chance that that individual failures would occur would then be combined to develop an overall estimate of the reliability of the entire system. Doing so would involve applying approaches such as building more-complete fault trees for the response and applying project risk analysis, techniques from the project-management and engineering fields developed to systematically assess the performance of complex systems. Such techniques were developed to assess failure risks quantitatively and to make it possible to evaluate how different courses of action might push the chance of success up or down even slightly. This can be done for technical systems because failure rates for individual components can frequently be measured directly and complex modeling can be used to calculate precise estimates of system performance under uncertainty.

Since emergency response efforts are human systems and must contend with the unpredictability of post-incident environments, it would be unrealistic to expect that their reliability could be estimated with comparable levels of precision to technical systems. To expect to be able to measure small differences in the reliability of response systems—e.g., distinguishing whether an additional investment in equipment or training would raise the chances of success for an operation a few percentage points—would almost certainly be unproductive. Realistically, attempts to assess response reliability might only be able to reasonably distinguish larger overall changes, or even that likely reliability fell into one bin rather than another in a coarser qualitative ranking.[13] Conversely, assessments would have to take into account the strengths arising from the fact that these are human systems: For some failure modes, options will be available to "adapt around" the failure to keep an operation going that otherwise might have fallen apart.[14] While these complexities do limit the ways such reliability estimates can be used in decisionmaking, it only limits the value of making those assessments slightly; as is the case in many areas of policy analysis, the process of systematically assessing reliability as objectively as possible may provide as much value as the numerical value or ranking the process produces at the end.

Current national preparedness efforts do already address some of what is needed to assess overall response reliability. For example, the exercise evaluation guidance produced by the Homeland Security Exercise and Evaluation Program includes "root-cause analysis" as part of

---

[13] For example, it was a 4 rather than a 5 on an ordinal scale or a "high" rather than a "medium" in an even coarser rating system.

[14] However, to the extent that preparedness systems are regimented and institutionalized and this flexibility reduced, this advantage may be reduced. See, for example, discussion in John R. Harrald, "Agility and Discipline: Critical Success Factors for Disaster Response," *The ANNALS of the American Academy of Political and Social Science,* Vol. 604, 2006, pp. 256–272.

after-action reviews of exercises to determine the fundamental causes of breakdowns identified during simulated response operations.[15] The results of such root-cause analyses, in addition to helping identify areas needing improvement, could also be used to inform more-general assessments of the likely reliability of an area's response efforts. It is likely that other sources of data also already exist, including after-action reviews of actual response operations, that could provide relevant insights to inform assessment.

Advantages of Using Reliability Measures for Preparedness Evaluation

Including reliability assessment in preparedness evaluation would have two key advantages in thinking about national preparedness policies.

First, whether or not an "exact" response reliability can be calculated, the process of identifying and assessing threats to response performance may identify high-payoff investments that would otherwise be missed. If there is a key factor that is limiting overall response reliability, the risk it poses could hurt the value of all other investments (e.g., if your buses are very likely to break down, spending lots of new money training drivers may have no effect on performance in an actual event). Focused investment to fix such a problem would provide a particularly high payoff since doing so would magnify the value of both past and future investments in response capability. Conversely, if such major threats to reliability exist and are not recognized, resources may continue to be poured into areas that look attractive, but may not result in better performance when incidents actually occur. Identifying major threats to reliability could also identify preparedness needs that should be a high priority in the near term, whatever their cost. If there is a weak link in a preparedness system that could cause its functioning to break down entirely, fixing that vulnerability should be a near-term target.

These concepts make such an assessment particularly relevant to capabilities-based planning efforts within the National Preparedness System (and preparedness efforts at the state, local, and nongovernment levels). Though capabilities-based planning provides a structure for identifying and building capability portfolios that are relevant across a wide range of incidents or contingencies, there is always the question of "how much of a given capability is enough."[16] By linking individual capability levels to their impact on overall response reliability, this approach provides one way of identifying the point where added investments in one capability no longer are producing commensurate improvements in overall preparedness and resources should be directed elsewhere.

Second, without a way of discussing the reliability of different response activities, policymakers and the public are missing a key ingredient for calibrating their expectations about both what the response system can do and how much it costs to achieve different levels of performance. Reliability measures are one way to explain why the expenditures of response organizations in one area may be higher than in another. A response system that can deliver 95 percent reliability will almost certainly be more costly than one whose likelihood of per-

[15] Homeland Security Exercise and Evaluation Program, "Volume III: Exercise Evaluation and Improvement Planning," February 2007.

[16] This is a different question from the more general "how much preparedness is enough" overall question, discussed below.

forming is closer to 50/50. If an organization is spending more so it can provide more-reliable response, that difference in performance should be recognized so the resulting higher spending is not misinterpreted as inefficiency or excess. Similarly, reliability measures can provide a language to better explain the potential effects of policy changes. If the effect of a funding cut is primarily on the likely reliability of response, its effects may be invisible until disaster strikes and the system fails to function. While it may still be a reasonable policy choice to cut funding, good policymaking requires that it be done with an understanding of its effects.

Because of these advantages, an explicit effort to assess the reliability of response systems is an important element of homeland security planning at all levels. For organizations charged with responding to major incidents, an assessment of the reliability of response systems could be useful both as an input to planning and in making the case for new investments to improve preparedness. Though examinations of this kind of reliability assessment are likely routine in many organizations for the "everyday emergencies" that make up the bulk of their activities, an analysis of how systems might work—or fail—in larger events and why they might do so provides a specific rationale for the allocation of resources to strengthen those systems. Whether the case is being made to a governor or to the Department of Homeland Security through a grant program (for state and local response organizations in the governmental homeland security system) or whether it is being made to corporate executives (for business continuity or "internal preparedness activities" in the private sector), an argument that a specific investment affects the reliability of response efforts (and why it does so) provides a specific and substantial rationale for that investment.[17]

However this type of analysis could contribute to supporting requests for more preparedness resources, explicit efforts to measure response reliability and determine how new resources affect it are also potentially useful to the funders of preparedness. Because large-scale incidents and disasters fortunately occur rarely, it is often difficult to impartially assess or objectively verify exactly what expenditures on preparedness are "buying." As a result, determining if one preparedness investment (in training, for example) is the best use of the marginal homeland security dollar (which could otherwise be spent on equipment or in other ways) is difficult, if not impossible. Accountability regarding the past use of funds is also difficult, and decisionmakers are challenged to determine if adding more resources to preparedness is truly the best use of the public's or the shareholders' money. Laying out how good different parts of the system are believed to be now and how good they are expected to be after a particular investment in equipment, people, or planning forces explicit, "on-the-record" judgments that might otherwise not be made. Therefore, reliability assessment could help to both build accountability into the system and make it easier for homeland security funding efforts to improve over time. Systematically shaping preparedness efforts requires more than simply observing after each major disaster that the system did not function as hoped for and, similar to the criticism that military planners frequently focus on "fighting the last war," can result in dispropor-

[17] Although developing ways to discuss response reliability can improve policy choice, it is important to recognize that maximizing the reliability of response systems should not necessarily be the goal. In spite of the damage that emergencies large and small can produce, it is reasonable and expected that different systems and response activities will have different levels of reliability. Increasing the reliability of response efforts costs money, and the investment needed to make the outcomes of response efforts more certain must be traded against other possible uses for those funds. As a result, deciding how much reliability we need must be driven by the preferences of decisionmakers and the populations they serve. The public and policymakers must decide "how much reliability is enough."

tionate emphasis being placed on the individual failure modes that caused the breakdown in performance.

CHAPTER FOUR

How Might the Impact of This Approach Be Evaluated?

Thinking about actually assessing the reliability of response systems for large-scale incidents—where operations can involve organizations from many response disciplines, geographic areas, and levels of government—is far from a trivial exercise. As a result, the first step in assessing the value of such an approach must be a focused systems-analysis and modeling effort of representative response systems to provide a basis for estimating the reliability of those systems for different types of incidents and to explore the value of making such assessments for improving planning.[1]

A variety of data sources already exist that could provide a starting point for such efforts: Comprehensive emergency plans frame the outline of what the response systems they describe look like, after-action reviews of past exercises and responses operations can help identify failure modes, and other collections of emergency response best practices and strategies[2] can help identify the ways that response organizations have hedged against or adapted around those failure modes in the past. The extensive framework of necessary response capabilities included in the *Target Capabilities List* and initial metrics provides a palette of "starting materials" for defining the variables and structure for needed models. Previous work on simulation-based studies of different elements of emergency-response activities similarly provides a template for model design.[3]

An effort to build and analyze such systems models for real-world response systems will have to take into account the practicalities of how response operations are run "in real life"—as one reviewer of this paper pointed out, responses to major incidents often involve improvisation based on preexisting plans, rather than the deterministic following of those plans. However, though adding to the complexity of the effort, such problems are not insurmountable—and are similar to those faced when systems-analysis techniques were first applied to national defense problems decades ago.

Research and analytical efforts are needed to help develop appropriate measures to assess the reliability of different response elements. The reason most current efforts focus on things that can be counted is because that counting is comparatively easy—and assessment that goes beyond such input measures will be more difficult almost by definition. If current data sources

[1] Because they are primarily analytical and simulation activities, the resource demands for exploring these ideas should be comparatively modest, particularly when compared to the national-level investment in preparedness activities themselves.

[2] For example, information contained in the Department of Homeland Security Lessons Learned Information System.

[3] For example, modeling efforts exploring disease spread and the effectiveness of particular interventions, throughput in "service delivery type" response activities such as the dispensing of pharmaceuticals, and others.

do not have all the ingredients for these sorts of assessments, how can the blanks be filled in? Figuring out what new information needs to be collected, how evaluation of exercises might need to be changed, and other ways of gathering more information are important parts of assessing this potential change to preparedness evaluation. In some cases—as is the case in many fields—where data is simply unavailable (or unavailable at reasonable cost), expert judgment and opinion will have to fill in.[4] Systematically using self-reported data or expert judgment in assessments is difficult, but processes can be designed to minimize the potential for subjectivity or other effects skewing assessments.[5]

Exploration into different ways these concepts could be used in preparedness assessment at different levels is also needed. Though the discussion in this paper has been framed largely in terms of national-level preparedness, assessments based on response reliability could also be made for individual organizations, cities, states, or regions. For some preparedness efforts—e.g., in private-sector organizations that are familiar with thinking about the reliability of service delivery in their everyday activities or individual-responder organizations that routinely measure things like response time and scope of service delivery—determining how to craft and apply such ideas could be straightforward. Their application to the activities of multiple organizations and to the national-level preparedness planning that this discussion has focused on will require more testing and experimentation.

The final element in evaluating the utility of reliability assessment in preparedness evaluation would be looking at how such assessments can be used to frame policy evaluation and debate. For example, a pilot effort applying these concepts in evaluation could focus on an assessment of one city or metropolitan area and a retrospective look at its investments in preparedness over the last several years. By building a model of the area's preparedness system and examining how new expenditures on planning, exercises, staff, or equipment fit into that system, the amount of additional reliability those investments "bought" could be estimated. The utility of assessments built on that foundation could then be compared with existing ways such investments would otherwise be evaluated.

Beyond just helping to assess the value of preparedness investments after the fact, a key part of the argument for examining response reliability is that it could improve prospective decisionmaking for both policymakers and the public. A priori, it appears that simply identifying critical failure modes and framing outcome objectives in ways that link explicitly to the investments made to achieve them have significant advantages. However, an assesment of how much effort a broader endeavor to measure response reliability merits should take into consideration the scope of its other benefits. For government audiences, evaluation efforts would therefore need to examine whether analyses focused on how preparedness investments improve

---

[4] See, for example, discussion regarding similar problems in the military context reported in S. Craig Moore, J. A. Stockfisch, Matthew S. Goldberg, Suzanne M. Holroyd, and Gregory G. Hildebrandt, *Measuring Military Readiness and Sustainability,* Santa Monica, Calif.: RAND Corporation, R-3842-DAG, 1991.

[5] Such problems arise in evaluations in many areas. For example, when RAND was developing and implementing measures for assessing health care quality across the nation, designing reporting and evaluation systems that addressed these concerns were central to the effort. Other structured techniques, such as the Delphi method (a systematic process for elicitation of expert opinions), could also be applied to this problem. The Delphi approach was developed for just these types of problems—where uncertainties make predictions difficult and techniques are needed to systematically aggregate expert opinion in a way in which group dynamics and other confounding factors would skew the outcome (for example, see Bernice A. Brown, *Delphi Process: A Methodology Used for the Elicitation of Opinions of Experts*, Santa Monica, Calif.: RAND Corporation, P-3925, September 1968).

response reliability are more useful than the ways these choices are considered in policy debate. Does focusing on response reliability suggest different priorities for expenditures? Will it actually improve potential response outcomes? Answering these questions would require both determining the perceived value of these approaches to decisionmakers and, through pilot efforts or other policy evaluations, studying whether the outcomes of using them actually are superior.

National policy debate on preparedness does not just involve policymakers, however: It involves the public as well. Though it might seem unrealistic that concepts such as response reliability could be used to effectively frame policy choices for the broader electorate, a recent RAND experiment suggests otherwise. In focus groups drawn from the general public, RAND researchers used an analytical technique called assumption-based planning to frame discussions of policy choices in the Global War on Terrorism and, by doing so, focused discussion on the outcomes the country was trying to achieve and how policies did or did not contribute to doing so.[6] Framing preparedness policies using concepts such as response reliability could similarly elevate national debate in this area—moving beyond arguments about funding levels to a more productive discussion of the trade-offs between investments and what levels of performance we should reasonably expect from the national preparedness system.

---

[6] Robert J. Lempert, Horacio R. Trujillo, David Aaron, James A. Dewar, Sandra H. Berry, and Steven W. Popper, *Comparing Alternative U.S. Counterterrorism Strategies: Can Assumption-Based Planning Help Elevate the Debate?* Santa Monica, Calif.: RAND Corporation, DB-548-RC, 2008.