



HOMELAND SECURITY AND DEFENSE CENTER

THE ARTS
CHILD POLICY
CIVIL JUSTICE
EDUCATION
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INTERNATIONAL AFFAIRS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
SUBSTANCE ABUSE
TERRORISM AND
HOMELAND SECURITY
TRANSPORTATION AND
INFRASTRUCTURE
WORKFORCE AND WORKPLACE

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

[Jump down to document](#) ▼

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

Support RAND

[Purchase this document](#)

[Browse Books & Publications](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore the [RAND Homeland Security
and Defense Center](#)

View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND PDFs to a non-RAND Web site is prohibited. RAND PDFs are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the RAND Corporation monograph series. RAND monographs present major research findings that address the challenges facing the public and private sectors. All RAND monographs undergo rigorous peer review to ensure high standards for research quality and objectivity.

Evaluating the Reliability of Emergency Response Systems for Large-Scale Incident Operations

Brian A. Jackson, Kay Sullivan Faith, Henry H. Willis

Prepared for the Federal Emergency Management Agency



HOMELAND SECURITY AND DEFENSE CENTER

This research was sponsored by the Federal Emergency Management Agency and was conducted under the auspices of the RAND Homeland Security and Defense Center, a joint center of the RAND National Security Research Division and RAND Infrastructure, Safety, and Environment.

Library of Congress Cataloging-in-Publication Data

Jackson, Brian A., 1972-

Evaluating the reliability of emergency response systems for large-scale incident operations / Brian A. Jackson, Kay Sullivan Faith, Henry H. Willis.

p. cm.

Includes bibliographical references.

ISBN 978-0-8330-5005-2 (pbk. : alk. paper)

1. Emergency management—United States—Evaluation. 2. Preparedness—Evaluation. 3. Incident command systems—United States. 4. Assistance in emergencies—United States. I. Faith, Kay Sullivan. II. Willis, Henry H. III. Title.

HV551.3.J328 2010

363.34'80684—dc22

2010024680

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

© Copyright 2010 RAND Corporation

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (<http://www.rand.org/publications/permissions.html>).

Published 2010 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

Summary

Societies build emergency response systems to be there when damaging incidents—whether natural or caused by man—occur. Though the effectiveness of those systems in responding to everyday emergencies is easy to see, knowing how prepared they are to deal with large-scale incidents—which, fortunately, are far rarer—is much more difficult. Most of the time, responses to even large-scale emergencies go very well. But sometimes they do not, leading to questions about why response activities did not go as expected and what policy actions should be taken in response.

Practitioners and researchers in many fields have devoted significant effort to developing ways to measure emergency preparedness. Progress has been made—in the creation of systems to assemble data on preparedness inputs, national policy documents that begin to set standards for capability levels, and exercises designed to test preparedness systems—but the ability to measure preparedness has still been highlighted as an area requiring attention and innovation (FEMA, 2009b). This work helps address that shortfall by approaching preparedness assessment from a perspective that is very different from those used in most previous efforts.

We view the response operation for a large-scale emergency or disaster as a system, one that is built to address post-incident needs and potentially involves multiple separate organizations.¹ In this view, a response system is defined by a set of plans, resources, authorities, agencies, and their associated human resources. We draw on tools from the systems analysis and engineering fields for analyzing system performance as a way of looking at potential response performance at future incidents. This includes laying out what the system is intended to do and exploring what inputs and resources are required for it to deliver, a component common to most preparedness assessment efforts. But we also look at the system and assess what might go wrong—what breakdowns or “failure modes” might threaten the ability of the system to perform effectively. This second part

¹ This framing is consistent with the Emergency Management Accreditation Program’s definition of *emergency management program* as a “jurisdiction-wide system that provides for management and coordination of prevention, mitigation, preparedness, response and recovery activities for all hazards” (EMAP, 2007). Such a system “encompasses all organizations, agencies, departments, entities and individuals responsible for emergency management and homeland security functions,” though the focus in our work was on the system’s preparedness and response activities.

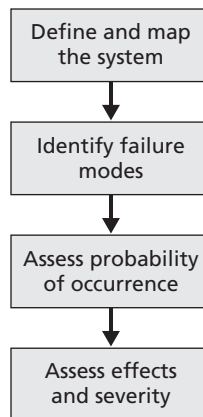
is the heart of our analysis and distinguishes our approach from most preparedness assessment methods. The combination of these two approaches can help to answer the fundamental question that the public and policymakers have about response systems: *How much confidence should we have that they will function as planned when the next large-scale incident or disaster occurs?*

Assessing the Reliability of Response Systems

To answer that question, what is needed is a measure of the likelihood that a response system will perform well—or, put another way, that events that prevent it from performing well will not occur—at a future incident. We have labeled that measure *response reliability*, the probability that a response system will be able to deliver at or above a given level of capability at a future emergency incident. Our framing of response reliability takes into account that the scale, scope, and complexity of an incident matters. A given response system may perform predictably well for events up to a certain level of required performance, but above that level problems may considerably reduce the chances that the system will be able to meet the requirements of the incident.

To evaluate response reliability, we adapted analytical techniques from the reliability analysis and risk assessment fields—specifically, fault tree analysis and failure mode, effects, and criticality analysis (FMECA). Building on the ways these techniques are applied to technical systems, we framed four steps for analysis of response systems for large-scale incidents (and diagram them in Figure S.1).

Figure S.1
The Four Steps of Response Reliability Analysis



1. **Define and Map the System.** Understanding what might go wrong in a system requires knowing how it is put together. Laying out the different functions (in the case of response operations) that must be performed and how they link together defines the structure and bounds of the analysis. For example, evacuating an area requires transporting people who have no transportation of their own, which involves not just vehicles and drivers but also (1) responders to manage gathering the people and their orderly boarding, (2) public communications capability to get the message out to people that an evacuation is under way, (3) information collection capabilities to identify where the people who need assistance are, and (4) a functioning incident management system to fit all the activities together and make sure that they are completed in time to be valuable. Breakdowns in the system could be caused either within individual components or at the seams between components that depend on one another. Defining the system requires understanding what it means for each part of the system to work well and determining how reductions in performance would affect outcomes.
2. **Identify Failure Modes.** Failure modes are defined as “the observable manners in which a component fails” (Ebeling, 1997, p. 168), which in this case would be the ways that performance of different parts of the response system would break down. Examples of failure modes for response activities include staffing problems, human errors, equipment breakdowns (e.g., damage to response vehicles), and so on. Some failures might occur as a response operation was being initiated, while others might occur at later points in the response. Failures may be due to random events (e.g., equipment failures even though all appropriate maintenance had been done), have a clear human cause (e.g., maintenance had not been done), or be caused by external events (e.g., the incident damaged the vehicles prior to deployment). In our work, we drew on the practitioner literature, response after-action reports (AARs), past RAND research, and other published sources to help identify the ways that response operations can break down. We account for the effect of failure modes on each part of the system by determining whether each mode is specific to one response function or capability or has more general effects on multiple parts of the system.
3. **Assess the Probability of Occurrence of Different Failure Modes.** Given many things that could hurt the functioning of a system, one differentiator among them is how likely they are to happen. The probability that a specific failure will occur during a response operation could be estimated a number of different ways; for example, the estimate might be based on real-world data on the occurrence of failures in past responses, or it might be based on estimates elicited from practitioners or subject-matter experts. Different ways of estimating the probability of failure have their own strengths and weaknesses with respect to accuracy and ease of implementation. Depending on how failure

modes have been defined, some calculation may be involved in determining the probability of a specific mode. For example, if the failure mode of concern for response is a communications system breakdown and there are both primary and backup systems, then the probability of the failure would be the probability *both* systems failed.

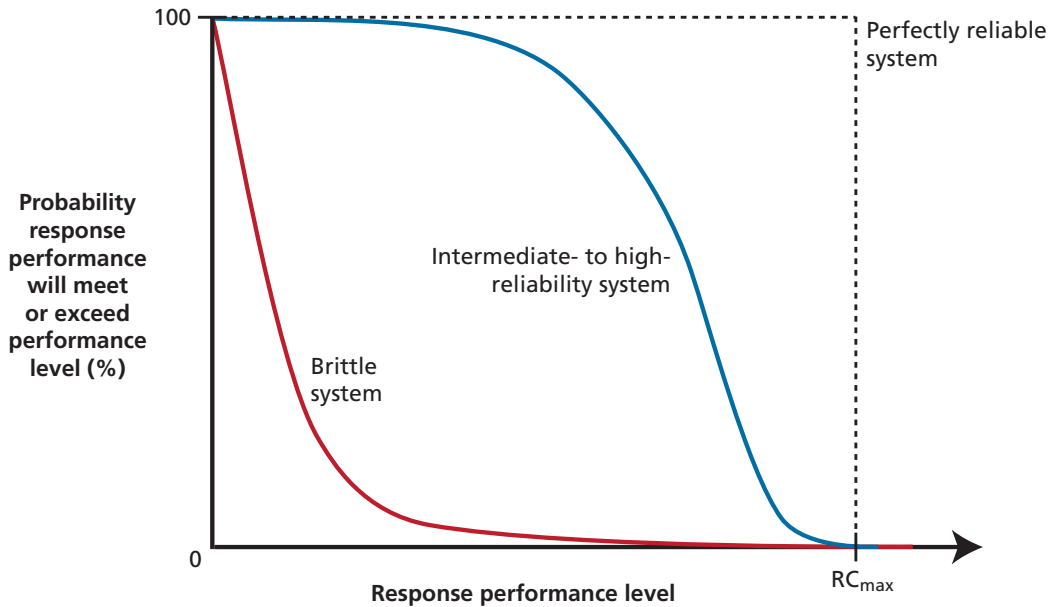
4. **Assess the Failure Mode Effects and Their Severity.** Other differentiators among failure modes are their effect and severity. In FMECA, this assessment is done at the system level, by asking, “What is the effect of the failure mode’s occurrence on overall system performance?” Different events can have a variety of effects. In considering the effect of failures on emergency response operations, we viewed failures as falling into two main classes: (1) response–termination failures—that is, failures that would stop a response operation entirely—and (2) capability–reduction failures—that is, failures that make a response operation less effective but do not halt response (e.g., an event that reduces the number of victims a hospital can accept after an incident). Effectiveness-reduction failures may cause a reduction in system performance either directly or via their effects on other response functions—for example, difficulties implementing incident command could have an effect on many other response activities. The severity of an effectiveness-reduction failure is determined by the size of its impact relative to the requirements of the response operation.

FMECA is one of a number of methods in reliability analysis for combining information on the incidence and consequence of failures into an overall assessment of a system. For our work, we found this process attractive because the approximations that are made allow each failure mode to be treated independently, somewhat simplifying use of the technique for assessing a complex response system.

In our study, we explored FMECA from two complementary perspectives.

First, we did a complete theoretical analysis of a very simple response system to fully demonstrate the methods and show how both qualitative and quantitative assessment of response reliability could help inform preparedness planning and resource allocation. Statistical modeling of a response operation affected by a variety of different failure modes made it possible to show how the probability that the system would perform effectively at incidents of increasing levels of required performance could be calculated. Figure S.2 illustrates how graphs of a response system’s reliability (on the vertical axis) delivering increasing levels of capability or performance (moving right on the horizontal axis) can provide an overall snapshot of its likely future performance. The curves in the figure show three exemplary cases: a system that is perfectly reliable (the gray dotted line) and so functions perfectly up to the highest level of performance it has been designed for (the maximum response capacity, RC_{\max}); a system with serious problems whose probability of good performance drops very rapidly as the required performance level increases (the red line, labeled a “brittle system”); and a

Figure S.2
Illustrative Reliability Curves for Response Systems of Varied Performance



RAND MG994-S.2

more realistic system (the blue line) that performs well over a large range of incidents but has difficulty for those close to its maximum designed capacity. For a real response system, such an analysis could be done with sufficiently detailed results from the process described above, identifying failure modes and estimating their probabilities of occurrence and the severity of their consequences.

Second, we analyzed a more realistic response for a chlorine release incident, drawing on AARs from past response operations as a data source. This element of the work was designed as a proof-of-concept demonstration of the analysis using real-world data that response organizations already produce. We constructed a model of the response to a chlorine incident, covering the elements of all response tasks from incident management through assisting exposed victims. We identified failure modes for each part of the model, including critical interdependencies among different parts of the response operation. We then analyzed a convenience sample of 70 AARs, describing 65 separate response operations. All but two of the AARs described actual incidents, with the remainder describing exercises. We examined each AAR for the presence of different failure modes during the response operation and any descriptive information on the consequences of each failure's occurrence. This second phase of the work simultaneously demonstrated the promise and the challenge of the analytic approach when applied to real-world response systems.

Using the Results of Reliability Assessment in Preparedness Evaluation and Decisionmaking

The goal of this work was to show that adapting techniques from reliability engineering and risk analysis for evaluating the performance of technical systems can contribute to better ways of evaluating preparedness and anticipating the likely future performance of emergency response systems in large-scale events. We believe that we have achieved that goal, and have demonstrated that both the *process* of such analyses and their *results* can potentially contribute to preparedness planning and evaluation in different but complementary ways.

The first step of the process, defining and mapping the response, takes an explicitly systems-oriented approach to how an entire response operation functions. In our model, we do not distinguish which responders will perform the tasks in each part of the overall system, in terms of which organizations they are a part of or which response disciplines they are trained in. By ignoring the insignia on the uniforms of individual participants in the response, this approach lays out in black and white the potential interdependencies among organizations and how seams between them could result in response failure. In discussing our work with one response practitioner, the comment was made that “though we are supposed to be breaking stovepipes, we still do a lot of our planning within single agencies—and this captures the problems that can still create.”

The second step, systematically identifying failure modes for each part of the response model, provides a structured method for doing the type of “what-if” questioning done by experienced planners, and also for capturing the results of that process so they can be explicitly included in an organization’s plan and the knowledge spread among its staff. Working down to the level of individual root failure modes also makes it easier to identify solutions to problems that are discovered, since different failure modes—even ones within the same response function—can have very different “fixes.” Even just counting up failure modes and determining the breadth of their effects can help inform prioritization, with failure modes that have broad effects on performance standing out as causes for particular concern.

The third and fourth steps—assessing the probability, effects, and severity of the consequences of individual failure modes—get at the information needed to identify specific priorities and to assess the value of different preparedness investments. In our work, we drew on existing AARs from response operations to test this part of the analysis with real-world data. The AARs we examined proved to be a challenging data source. But we were nevertheless able to apply the basic analytical process we describe, and this process made it possible to extract useful data from a very heterogeneous dataset. Though we were seeking that data to inform qualitative and quantitative measures for response performance, practitioners who we interacted with also suggested other uses for such datasets. For example, for a specific jurisdiction, data showing that fail-

ures were adding up in a specific area could be used as a way to suggest which parts of the response system might need “preventive maintenance”—refreshers in training, particular focus in near-term exercises, and so on—to reduce their chances of recurrence in the future. Such applications could help to address requirements for exercises and corrective action programs in relevant emergency management standards (e.g., EMAP, 2007; NFPA, 2007).

In considering potential future implementation of these methods for broader preparedness assessment, a variety of other data sources may be superior to AARs for providing the information needed. Some such systems—for example, current preparedness assessment systems and remedial action management programs at the national level (FEMA, 2009b, p. ii) or local equivalents—might provide even better data on specific failure modes and their consequences, which could inform higher-resolution analysis of real response systems. These methods have the potential to contribute to current efforts to improve preparedness assessments (such as those required by the Post-Katrina Emergency Management Reform Act [P.L. 109-295]). Similarly, though our proof-of-concept work here used historical data from AARs, these approaches could be applied to more real-time datasets on response performance. Doing so would be consistent with the Federal Emergency Management Agency’s goal to “support a living reporting mechanism that will provide an up-to-date resource on the current state of preparedness” (FEMA, 2009b, p. 1) in the nation.

Comparing the results of our reliability analysis of a real-world response operation (using AARs) with our illustrative analysis of a simple response system using simulated data, we could not take our analysis as far “in practice” as we could “in theory.” In part, this was due to shortcomings in the AARs as a data source; small changes in the type of information included in such documents—i.e., capturing some estimate of the seriousness of the consequences of response issues in AARs—could make them much more useful for this type of analysis. Nonetheless, the results of our analysis and simulation using a simpler response scenario demonstrate the broader potential of reliability analysis to contribute to preparedness planning and evaluation. Though the data available to us did not support highly quantitative analysis of the chlorine response scenario, to the extent that response reliability curves can actually be estimated for real-world response operations, they could help provide a direct answer to the question—“What is the chance that things will work next time?”—that most current preparedness assessment methods cannot.

Having such a measure would help to inform policy debate of preparedness issues in a number of ways. Quantifying response reliability would help to make clear how much reliability the public should expect given current investments in preparedness, clearly articulate the cost of increasing it, and provide a means to compare different possible investments to do so—from surgically fixing known failure modes to just buying more capability to put additional slack into the system to respond to an unknown future. Reliable data on or solid estimates of response systems’ reliability would help to

focus preparedness policy debate and inform consideration of the truly key questions in this area: not just “How much money should we spend?” but “How exactly should we spend it?” and not just “Do we need to spend more?” but “How do we know when we have invested enough?”