

# **Analysis Planning for a Domestic Weapon-of-Mass-Destruction Exercise**

**Brian McCue**

with  
**Christine A. Hughes**  
**Kathleen M. Ward**

**The CNA Corporation**

---

*Alexandria, Virginia*

The CNA Corporation (CNAC) is a private, nonprofit organization committed to working in the public interest. The Occasional Paper series is published by CNAC, but the opinions expressed are those of the authors and do not necessarily reflect the views of CNAC or our sponsors.

# **Analysis Planning for a Domestic Weapon-of-Mass-Destruction Exercise**

**Brian G. McCue**

**with**

**Christine A. Hughes**

**Kathleen M. Ward**

**Center for Naval Analyses**

---

*4825 Mark Center Drive • Alexandria, Virginia 22311-1850*



---

# Contents

<b>Introduction: Turning data into power.</b> . . . . .	1
<b>Overview.</b> . . . . .	5
Scenario design . . . . .	5
Analysis goals and influences . . . . .	5
Analysis areas . . . . .	6
Emergency public information . . . . .	7
Emergency public policy and decision-making . . . . .	8
Communications and connectivity. . . . .	9
Resource allocation . . . . .	11
Jurisdiction . . . . .	12
Anticipating the enemy. . . . .	13
<b>Observation.</b> . . . . .	15
<b>Reconstruction.</b> . . . . .	19
<b>Analysis</b> . . . . .	21
Measures of effectiveness . . . . .	21
"Issues," "indicators," and "instances". . . . .	22
A sample analysis product . . . . .	25
<b>Assessment</b> . . . . .	27
<b>Points arising in exercise analysis (and assessment)</b> . . . . .	29
Artificialities . . . . .	29
The flow of time . . . . .	30
"Fair experiment vs. fair fight". . . . .	30
<b>Appendix A: Exercise influences.</b> . . . . .	33
TOPOFF2000 . . . . .	33
Dark Winter. . . . .	36
September 11 and the anthrax incidents . . . . .	36
<b>Appendix B: Sample questionnaires.</b> . . . . .	39

---

---

<b>References</b> . . . . .	45
<b>Bibliography</b> . . . . .	47

---

## **Introduction: Turning data into power**

Francis Bacon said, some 400 years ago, “Knowledge is power”; a saying in the computer-age adds, “...but information is not knowledge, and data are not information” [1]. In Naval exercises, for example, *observers* collect data, *reconstruction* turns the data into information, and *analysis* turns the information into knowledge. That knowledge is then given back to the participants, who alone can turn it into power. Ideally, they do so through a process formalized as *assessment*.

Naval exercises help the Navy and Marine Corps use their resources most effectively and plan for the future. One way to achieve that objective is by analyzing those exercises. Some exercises involve only one ship or unit to allow crews to test their skills or try out a new system. Others require several units to engage in a specific mission. Still other, more complex, exercises are fleet-wide and involve differing missions. They may even include forces from allied countries. CNA analysts, active in all phases of these exercises, first work with the fleet commander and his operations officer to determine the objectives of analysis and specify the kind of data to be collected. After the exercise, they reconstruct the events, based on data acquired from a variety of sources, to provide a clear and comprehensive picture of what actually took place. They may concentrate on certain periods or aspects of the exercise that are of particular interest. The analysts can quantitatively evaluate the performance of various units and equipment, help assess tactics, estimate overall readiness and capabilities in various warfare areas, assess scenario outcomes, and find shortcomings in new tactics or equipment.

After the tragic events of September 11, 2001, many groups began planning ways to improve our nation's readiness for domestic terrorist attacks using weapons of mass destruction (WMDs). Many of these efforts were exercises whose objectives were to gain knowledge, train, or evaluate. This paper presents an analysis plan for a civilian-agency exercise that focuses on terrorist use of WMDs. The goals of the analysis are to help those

---

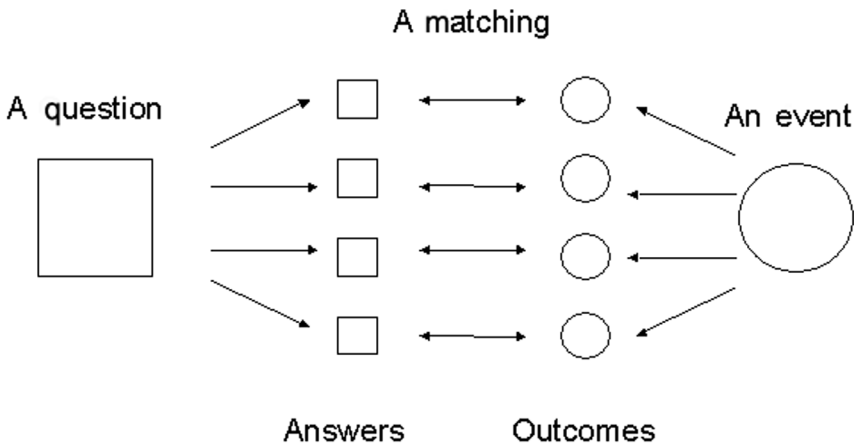
---

participating in the exercise understand problems, issues, resolutions, and solutions involved in such attacks. It is important for exercise participants to understand that analysis in this context is not evaluation; the analysis team is not going to evaluate performance—of personnel, systems, tactics, and so on. However, the analysis effort could make available information that participants may find of use in evaluating themselves or others.

We call it an exercise, but, in many respects, it may more usefully be thought of as an experiment. In an experiment (outlined in figure 1), one has:

- A set of questions and some possible answers
- An open-ended event and a set of possible ways that it could turn out
- A pre-determined matching of the possible outcomes to the answers.

Figure 1. Experiment illustration.



Then, when the event happens, one observes what happens and sees the answer(s) to which it points. Thus, an experiment is different from a *demonstration*, whose purpose is to display or transmit knowledge, or an *exercise*, whose purpose is to train or evaluate. The purpose of an experiment is to gain knowledge.

---

This document makes frequent reference to the TOPOFF 2000 exercise and its after-action report [2]. These statements are not to be taken as negative or critical regarding TOPOFF 2000; rather, they should be considered as signs of TOPOFF 2000's status as the cornerstone of civilian-agency, terrorist-WMD exercises, not just a building block.

Similarly, occasional references to other previous terrorist WMD-oriented civilian-agency exercises and to the responses to some real-world events are not to be construed as criticisms of those exercises and responses or of the people who undertook them.



---

## Overview

### Scenario design

Ideally, analysts are involved in the exercise's scenario design and should encourage the designers to think of the exercise in terms of an experiment (as shown in figure 1), even if the exercise has substantial training aspects as well. The designers may tend to think in terms of evaluation, which may be appropriate for narrowly defined tasks that will occur as part of the larger experiment (e.g., getting an ambulance to the site of a call). The analysts' reconstruction will be useful in making such evaluations.

For tasks that involve top officials or impinge upon the six major analysis areas set forth in the following section, however, *analysis* is not *evaluation*. The analysis team should not evaluate anybody or anybody's performance. There are many reasons for this, but, from an analytic standpoint, the reason is that a decision that turns out badly in an exercise is not necessarily a bad decision. Maybe the decision-maker knew more than the exercise-creator did. In the case of terrorist attacks with weapons of mass destruction, so little is known that the analysis effort ought not to attempt to identify right and wrong choices. The closest it should come is that people may be asked to comment on their own decision-making. This situation stands in stark contrast to the condition of the military exercises that constitute so much of the experience base of those administering civilian exercises, especially those pertaining to terrorism. After thousands of years of experience with warfare and at least 2000 years of experience with military exercises, there is a huge knowledge base regarding these topics. The corresponding knowledge base regarding terrorist WMD attacks is virtually nil.

### Analysis goals and influences

The goal of exercise analysis is to foster improvement in understanding. One path to that goal is to ensure that the next

---

---

such exercise is better than the one at hand. This WMD-oriented exercise analysis plan draws on multiple influences, including especially TOPOFF 2000, Dark Winter, and the real-world events of September 11 and the later anthrax incidents. (See appendix A for details of these events.)

## **Analysis areas**

A number of separate efforts have created sets of analysis topics for civilian-agency terrorist WMD exercises. We examined six: the TOPOFF 2000 after-action review [2]; the TOPOFF 2000 article by Inglesby et al. [3]; the Defense Preparedness Program [4]; a training curriculum by Pelfrey [5]; the tabletop exercise Dark Winter, as reflected in O'Toole et al. [6]; Palarino's memorandum to Congress [7]; and our own study of the coverage of the hijacking and anthrax incidents in the autumn of 2001. Though these analyses differ somewhat in content and presentation, they display considerable underlying similarity, including a listing of topics that turned out to be important:

- Emergency public information
- Emergency public policy and decision-making
- Communications and connectivity
- Resource allocation
- Jurisdiction
- Anticipating the enemy.

We see these analysis areas as relatively distinct and believe that viewing an exercise in light of these areas will provide a useful organization of observations and ideas. That is not to say, however, that the areas are independent. Many links connect them, and some are joined intimately (e.g., emergency public information and emergency decision-making).

---

## **Emergency public information**

TOPOFF 2000 and the anthrax incidents underscore the importance of public information. We use the term “emergency public information” to indicate that such information might differ from business-as-usual public information. Further, the task of those responsible for public affairs might vary according to the type of emergency-disaster, crisis, or enemy attack. Thus, those responsible for public information may find that, despite the fact that they do their job every day, it becomes different—but no less important and possibly even more important—during a set of events like those to be simulated in the exercise.

O’Toole et al. state the ultimate importance of public information in their analysis of Dark Winter [6]:

The individual actions of U.S. citizens will be critical to ending the spread of contagious disease; leaders must gain the trust and sustained cooperation of the American people. Dark Winter participants worried that it would not be possible to forcibly impose vaccination or travel restrictions on large groups of the population without their general cooperation. To gain that cooperation, the President and other leaders in Dark Winter recognized the importance of persuading their constituents that there was fairness in the distribution of vaccine and other scarce resources, that the disease-containment measures were for the general good of society, that all possible measures were being taken to prevent the further spread of the disease, and that the government remained firmly in control despite the expanding epidemic.

They close by quoting Senator Sam Nunn, a Dark Winter participant:

“The federal government has to have the cooperation from the American people. There is no federal force out there that can require 300,000,000 people to take steps they don’t want to take.”

---

Numerous other events of the recent past also show the importance of public information, especially in cases involving, or potentially involving, terrorist WMDs [8].

### **Emergency public policy and decision-making**

Like public information, public policy and decision-making during an emergency (especially one that might be an enemy attack) might differ from day-to-day policy and decision-making. Two kinds of decision-making are so important they are accorded separate areas of analysis, described below. One regards decisions that address the potential of a human adversary—in contrast to nature—to think, plan, and react. Another, resource allocation, is treated separately not only because of its importance and clear boundaries, but also because it, uniquely, might pit American decision-makers against one another.

With these two areas of decision-making set aside, much of what remains could be called “political” decision-making. Though the boundaries of the “political” are hard to define sharply, it is easy to detect (and even predict) the presence of political considerations in some decisions. For example, public health considerations might make quarantine an obvious choice. But, as was observed regarding TOPOFF 2000,

Decisions regarding patient isolation, travel advisories, home curfews, the closure of airports and highways, and attempts to “quarantine” cities and states must be balanced against the practical feasibility of such measures, and their implications for civil liberties.

[3]

Some see “political” considerations as inherently venal or self-serving, but the essence of our system of government is that (through the ballot box, i.e., “politics”) it forces public officials to take the public will into account. Thus, “political” decisions, far from being something that ought not to happen, can in fact be seen as the entire basis of our system of government, especially when made in response to new conditions that require a quick resolution.

---

Emergency public policy and decision-making are closely related to public information, as observed by Inglesby et al. [3]:

The Governor's Emergency Epidemic Response Committee eventually decided to offer antibiotic prophylaxis to all EMS personnel, police officers, hospital workers and their families. The decision was intended to have two consequences: (1) to allow critical responders to remain healthy while directly managing the disaster; and (2) by ensuring the safety of their families, to maintain the willingness of critical responders to report to work. This decision caused consternation among some participants as it meant that the health department would need to justify giving prophylactic antibiotics to family members of critical personnel at the same time that the city was unable to provide life-saving antibiotics for some citizens who were already ill.

Emergency public policy and decision-making are also closely related to jurisdiction, as expressed by Governor Frank Keating of Oklahoma after his participation in Dark Winter:

My fellow governors are not going to permit you to make our states leper colonies. We'll determine the nature and extent of the isolation of our citizens....You're going to say that people can't gather. That's not your [the federal government's] function. That's the function, if it's the function of anybody, of state and local officials. [6]

Again, emergency public policy and decision-making are closely related to resource allocation and anticipating the enemy, but these important areas require analysis in their own right, as treated below.

### **Communications and connectivity**

Nobody questions the importance of communications and connectivity in WMD emergency response, and few would question that there might be difficulties in these areas. Certainly there were difficulties encountered in TOPOFF 2000:

---

...decision-makers communicated mostly via conference calls. Almost all observers and participants reported that the process of decision-making by conference call was highly inefficient and led to indecision and significant delays in taking action. A series of very large conference calls took place that included officials from the city and state health departments, hospitals, CDC, Federal Emergency Management Agency (FEMA), and many others. At some points, as many as 50-100 persons were on these calls; many participants had never before met or worked together. The roles, authorities, and even identities of those participating in the calls, as well as the leadership of and agendas for the calls, were unclear. Such conference calls occurred throughout the exercise, one often literally running into the next. [3]

Palarino phrased the matter starkly after Dark Winter [7]:

To limit the number of deaths, communication, coordination, and command and control issues need to be addressed prior to an event.

.....

Coordination among large federal bureaucracies is a continual problem. Coordination becomes more difficult when local and state bureaucracies interact with federal agencies. During Dark Winter it was clear that the spread of disease could not be contained by use of a very limited supply of vaccine. Other measures should have been taken such as isolation and quarantine. Federal and state officials were at odds over whether or not to institute these types of issues. Closer coordination and cooperation between federal and state officials, and a better understanding by federal officials of what was occurring in the state may have slowed the spread of the disease.

Command and control was a significant issue during the exercise.

Very possibly, “communications and connectivity” provide the highest potential payoff of conducting an exercise, by virtue

---

of the severity and importance of its problems on the one hand, coupled with the comparatively large hope of finding ways to make improvements once specific trouble spots have been uncovered through the conduct of an exercise.

### **Resource allocation**

Resource allocation represents a unique category of decision-making because it can entail a struggle of the decision-makers against one another rather than against the enemy or nature, as noted by Inglesby et al. [3]:

By the termination of the exercise, at least 11 states were reporting cases of pneumonic plague and some were demanding that they be given antibiotics from the NPS [National Pharmaceutical Stockpile]. When other states began to report cases, Denver was told by the DHHS [Department of Health and Human Services] authorities that no further antibiotic Push-Packs [shipments from the NPS] would be available....

More generally, Palarino wrote (in the context of Dark Winter),

There is no surge capability in the U.S. health care and public health systems, or the pharmaceutical and vaccine industries. This institutionally limited surge capacity could result in hospitals being overwhelmed and becoming inoperable; could impede public health agencies' analysis of the scope, source and progress of the epidemic, the ability to educate and reassure the public, and the capacity to limit casualties and the spread of disease.[7]

Senator Sam Nunn also participated in Dark Winter and said afterward,

We're going to have absolute chaos if we start having war between the federal government and the state government. [6]

It is perhaps in this area that one must avoid the temptation to judge by results. Doing so is questionable even in the context

---

---

of real-world events, but in the context of make-believe events, judgment by results is logically untenable. Suppose the decision-makers, mindful of the threat of follow-on attacks, withhold additional Push-Packs, and then no further attacks occur. Were the decision-makers wrong? No—or rather yes, but wrong only in their attempt to read the minds of the scenario designers. In no sense can the fact that the decision-makers' choice turned out badly be construed as an indication that it was an incorrect choice; perhaps it was the correct choice, and the design of the scenario was incorrect.

### **Jurisdiction**

Metropolitan-area providers of emergency services typically have a set of interlocking Memoranda of Understanding that clarify jurisdictional issues, and TOPOFF 2000's agencies in Denver and Portsmouth were not exceptions. But terrorist attacks with WMDs bring into play entities and considerations not normally encountered, and not necessarily provided for in the Memoranda of Understanding. TOPOFF 2000 participants offered multiple commentaries as to the lack of clarity they felt regarding who was in charge of them, or who was in charge of a particular locale or activity.

Issues surrounding leadership, authorities and the processes of decision-making were widely cited as highly problematic and deserving of careful review. Overall, leadership roles and authorities in the crisis were uncertain. [3]

The former official agreed with the Senate staff member's assessment of the exercise. "Were there problems involving interagency coordination?" he said. "Sure. They were huge." [9]

Regarding Dark Winter, Palarino wrote:

Current organizational structures and capabilities are not well suited for the management of a biological weapon attack. Major "fault lines" exist between different levels of government (federal, state, and local),

---

between government and the private sector, among different institutions and agencies, and within the public and private sector. These disconnects could impede situational awareness, and contribute to loss of life. [7]

### **Anticipating the enemy**

TOPOFF 2000 didn't have an opposing forces (OpFor) group to represent the enemy, and the resulting absence of anybody to outsmart (or be outsmarted by) was seen by some as an area for improvement. The existence of an enemy makes the response to a terrorist attack qualitatively different from the response to a natural disaster. For example, the desire to keep the terrorists in the dark regarding plans counteracts the desire to keep the public informed. We sometimes take nature's fury as an indicator of malevolence, but of course it is not malevolent; we do not worry that posting an evacuation route will cause nature to aim thunderbolts at the bridges along the way. Terrorists, however, will react to our reaction to them, and this consideration must be taken into account. Media reports, some of them quite detailed, describing adjustments being made by the government in the wake of September 11, were criticized for giving too much information to the terrorists. In the context of immediate crisis management, this concern would be all the more valid.

O'Toole et al. quote James Woolsey after Dark Winter:

We are used to thinking about health problems as naturally occurring problems outside the framework of a malicious actor....If you're going against someone who is using a tool that you're not used to having him use—disease—and using it toward—quite rationally and craftily...an entirely unreasonable and god-awful end, we are in a world we haven't ever really been in before. [6]

For these reasons, terrorist WMD exercises conducted by civilian agencies may, in some cases, have an OpFor to represent the enemy. As the events of September 11 showed, terrorist enemies can be ferocious in the extreme, but in any one exercise it is important that the OpFor not overmatch the civilian agencies to

---

---

so great a degree as to preclude learning. Thus, representation of worst-case terrorist foes in exercises may have to wait until experience in earlier exercises has prepared the participants for such a threat.

---

## Observation

TOPOFF 2000 had a large number of controllers and observers, but the after-action review appears to be written largely on the basis of comments by the participants at the Hotwash, held some three weeks after the exercise. The participants disagreed afterwards on what had happened. These disagreements ranged from the minor (did the Improvised Explosive Device at Portsmouth (NH) go off at 0800 or 0815?) to the substantial:

While at the termination of the exercise the Expert Committee had not officially reached a final decision regarding whether the state borders should be closed or how these measures would be enforced, it is notable that some senior TOPOFF participants stated that the decision to close state borders had been made. [3]

To know what really happened in an exercise requires the presence of dedicated observers. Ideally, they would be persons familiar with the activities of those they are observing: firefighters would be assigned to observe firefighters, for example, but this is not necessary in all cases and if a mayor cannot be observed by another mayor, he or she must nevertheless be observed by somebody. Military reservists can make effective observers; so might emergency-services people from just outside the geographic limits of the exercise. Their departments probably have Memoranda of Understanding with departments inside the limits and taking part, and the outside departments would welcome the chance to attain some visibility into the exercise by providing some observers. Observers will require 4 to 8 hours of group training before the exercise. About 90 observers might be needed to observe an exercise of about the size of TOPOFF 2000 in the manner envisioned here.

These observers would function primarily by taking chronological notes. A pre-printed form can help them do so effectively, principally by ensuring that they record their name and the date and place of their observations. An observer would be needed

---

---

just about anywhere that there is a participant. Because some participants would move about, there needs to be some way for their observers to move with them, or at least to arrive at the same destination. Waivers or hold-harmless agreements might be needed for some observers.

Observers would probably also help with the administration of questionnaires, which to be effective, must be administered to participants daily. Perhaps questionnaires could be administered, or answered, electronically, but it is important they be done *daily*, because memories can fade rapidly. (Appendix B contains sample questionnaires.)

The OpFor might find it difficult to take observers with them, and, because they are not quite participants, they can be excused from doing so if necessary. However, observation of their actions is still important, and, if they choose not to be accompanied by observers, they have to take it upon themselves to submit accurate logs of their activities.

Finally, some participants will probably have valuable reactions that are best elicited by interviews. An after-exercise Hotwash is sometimes used as a source of data, but it is not a substitute for interviews of key participants during the exercise. If it does not occur immediately after the exercise, the vagaries of memory will weaken the results.

Some have suggested that the collection of data by logs, questionnaires, and interviews can be replaced by simply videotaping everything. Videotapes can be of great value as adjunct, or even confirmatory, data, but videotape is an extremely unfortunate choice as a primary means of data collection. It is time-consuming to review and impossible to sort electronically, and the audio portion is generally unintelligible.

All must recognize the importance of data collection to the exercise as a whole. The observers must be seen as having an importance equal to that of the participants and the controllers, and not as being VIPs, hangers-on, drones from DC, thrill-

---

seekers, members of the OpFor, or roleplayers.<sup>1</sup> Their acceptance would be furthered by wide promulgation of the policy expressed at the beginning of this paper: “Analysis is not evaluation: the analysis team is not going to evaluate anybody.”

---

1. This list is based on the author’s experience as an observer.

---



---

## **Reconstruction**

Reconstruction produces—from the data collected during the exercise—a fact-based, time-synchronized, deconflicted, and meaningful account of what actually happened. Such an account is notably lacking in the TOPOFF 2000 Exercise Observation Report, in which there appear at least two timelines (one from a controller and another from a participant), with no effort to resolve such inconsistencies as whether the Improvised Explosive Device went off at 0800 or 0815.

Reconstruction typically takes place in three steps:

- Independent and parallel reconstruction of events at each location, by analysts assigned to one or more locations.
- Group reconstruction, by all the analysts at once, of how the events at each location fit in with those at the other locations. This step typically engenders considerable revision of the individual analysts' initial reconstructions of the events at their locations.
- Creation of a single reconstruction report, usually undertaken by just one or two analysts.

The result is usually eye-opening to any single person-participant, observer, controller, or analyst, because any single individual's perspective is so limited.

Were an exercise of about the size of TOPOFF 2000 to be subjected to this process, there would probably be three analysts for each venue (Denver, Portsmouth, and the National Capital Region), and each step would take about two weeks.



---

## Analysis

Analysis takes the record of events provided by the reconstruction and seeks patterns in them. It does so in a manner that is *objective*. The requirement for objectivity is sometimes expressed as a desire for quantification, but the two are not really the same thing, and objectivity is the more desirable. Findings can be objective without being quantitative, as is explained below. Conversely, findings can be quantitative without being objective, as exemplified by the practice of asking participants or others for ratings on a numerical scale and then subjecting those numbers to arithmetic operations such as averaging.

### Measures of effectiveness

Prior to an exercise or other operational event, the intent to arrive at and apply quantitative measures of effectiveness (MOEs) is often expressed. Usually no such measures are ever advanced, because none of the proposed quantities is seen as truly capturing effectiveness. Quantities, typically chosen from those rejected during the quest for MOEs, are then advanced as measures of performance (MOPs), with the claim that there is something useful about them even though they cannot be defended as actually measuring effectiveness.

This frustrating sequence of events recurs so often because truly useful and defensible MOEs are in fact difficult or impossible to come by. At the minimum, formulation of a valid MOE requires the collection of a good deal of quantitative data (since it is rare to be able to create a quantitative output from qualitative inputs) and a thorough understanding of the process at hand. For example, baseball players' offensive performance is quantifiable in terms of runs batted in, batting average, and slugging percentage, but none of these is held to truly measure a player's offensive effectiveness. They are only measures of performance, and, while an assessment of effectiveness would take them into account, it would also take into account other statistics, such as stolen bases. In the end, the manager's decision

---

about the player would be based not only on these numbers, but also on qualitative facts about the player and on facts external to the player, such as how many other players who bat like he does (right, left, or switch-hitter) were already on the roster.

One often sees putative MOEs that are weighted sums. The weights are almost invariably chosen subjectively, and, in many cases, the inputs are themselves subjectively determined, e.g., by respondents applying a scale of 1 to 5. This method of formulating MOEs undermines their objectivity and also renders them useless as inputs to later calculations.

Armed combat, with which the human race has a regrettably long history, has so far proven resistant to the formulation of MOEs except in a few narrowly defined areas such as antisubmarine warfare [10]. With—again, thankfully—only scant experience with terrorist attacks that use weapons of mass destruction, there is little hope that meaningful or useful MOEs (or even MOPs) can be developed at present. However, this issue should be kept in mind as additional exercises take place, because at some point in the future the formulation of some MOEs might be possible.

### **“Issues,” “indicators,” and “instances”**

Absent any hope, for the reasons described above, of formulating MOEs or MOPs, an alternative framework for objective analysis of a terrorist WMD exercise is needed. We suggest a method based on issues, indicators, and instances.

The *issues* can—and have—been advanced earlier in this document in the discussion of analysis areas. They include the quality, quantity, and credibility of emergency public information, the organization of emergency public policy and decision-making, the need for workable communications and connectivity (including not just the hardware, but the user element), the need for orderly, comprehensive, and agreed-upon lines (both geographical and organizational) of jurisdiction, the need for correct allocation of resources across time and space, and the need to

---

outsmart the thinking, malevolent enemy. For a specific event, input regarding issues can also be solicited from participating agencies, though we suspect that most lists will strongly resemble ours.

These issues are all, necessarily, somewhat abstract. To ensure an objective treatment of their manifestations in the exercise, one must specify *beforehand* one or more *indicators* regarding each issue. These can be positive or negative. For example (see also table 1 for a more complete list of indicators, arranged according to the areas of analysis), one issue regarding emergency public information is the spokespersons' preservation of their credibility. In TOPOFF 2000, as reviewed in appendix A, the epidemiologist lost credibility at the press conference by not giving out an important piece of information (that simple dust masks can arrest the spread of pneumonic plague) until pressed by questions from a reporter. Thus, it is not unreasonable to set up an indicator of credibility: "gives out important information without prompting." The opposite, of course, would be an indicator of reduced credibility.

After the reconstruction is complete, it is examined for *instances* of each indicator. For example, suppose that the reconstruction's account of a press conference following the explosion of a radiological device reveals that the spokesperson did not mention the wind direction until asked to do so by a reporter. His failure to do so would then be an *instance* of the "gives out important information without prompting" *indicator* of the credibility *issue* of the emergency public information area of analysis.

The reason to choose indicators beforehand is to avoid the charge (not to mention the actuality) that one is being a "Monday morning quarterback," capitalizing on hindsight's wonderful clarity. Thus, the indicators listed in table 1, though preliminary and subject to further refinement, should be finalized before the exercise. It is also important to note that, although one or two indicators allude to incorrect decisions, any instances of these are to be identified only by the decision-maker in question, not by anybody else and certainly not by any analyst.

---

Table 1. Issues and indicators<sup>a</sup>

Area	Issue	Indicator
Emergency public information	+ Timeliness, clarity, consistency, honesty (v. correctness), relevance +Or, a good reason why not	- No information given to public - Long delays in giving out info + Recognition of situation in which info must be restricted to help in “anticipating the enemy”
	+ Sources that start and remain credible	- Wrong information (for no good strategic reason) - Important facts revealed only in answer to questions -Multiple and inconsistent spokespersons
Emergency public policy	- Slow in making obvious choices - Choosing incorrectly in obvious cases	- Long meetings on easy questions -Delays in implementing decisions once made - Decisions of which they think better later
Communications and connectivity	- Trouble in communicating	- Failure to establish communications promptly - Decisions made without all the known facts - Uncertain of identity of party one is communicating with
	+ Ease in communicating	+ Participants have all the needed facts + Participants agree as to what has been decided
Jurisdiction	- Overlaps in asserted authority	- Disputes over who’s in charge -Disputes over which (whose) set of rules/laws applies
	- Gaps in accepted responsibility	- No one taking responsibility
	+ Hierarchical relationships	+ Agreed-upon vertical chains of command
Resource allocation	-Trouble deciding	- Slow decisions, especially late decisions
	+ Recognition of the problem as such	- Decisions of which they think better later
Anticipating the enemy	+ Recognition of 2-sided situations when they arise	- Failures to see action of OPFOR for what it is
		- Decisions of which they think better later

a. Note: + and - indicate positive and negative issues and indicators.

---

## A sample analysis product

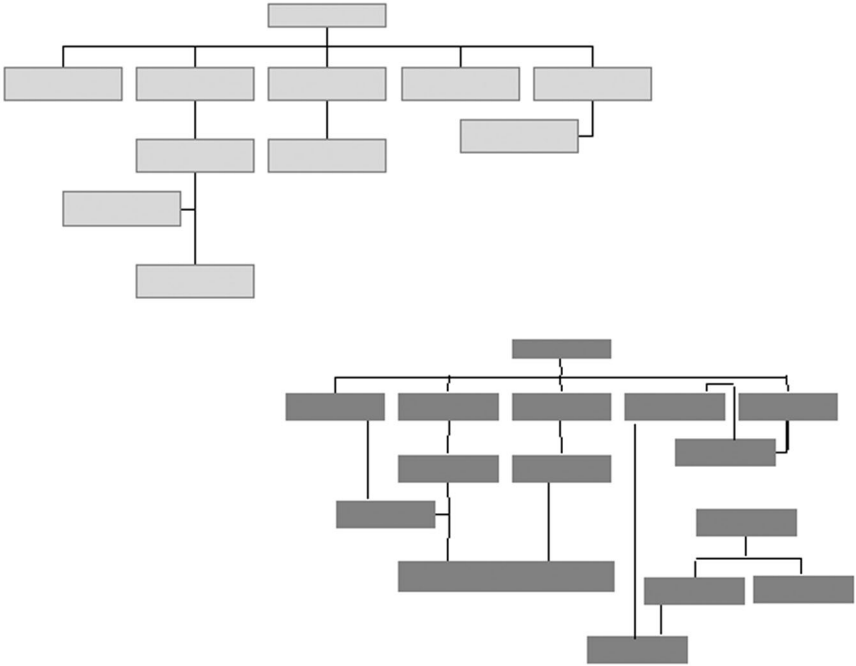
The word “analysis” originally referred to the process of taking things apart, and the preceding discussion demonstrates this fact: we started with some large areas of analysis (e.g., emergency public information) and some rather abstract notions regarding what we are to look for (e.g., credibility), and we have broken these ideas down to the point that we can comb the reconstruction for some well-defined behaviors (e.g., failure to give out important information until asked).

To a major degree, a table like the one above, supplemented by an additional column showing when and where (and where in the reconstruction) each instance occurred, would be an important analysis product in itself. It would help to structure participants’ (and others’, e.g., Congresspersons’) understanding of what happened, what it means, and what changes it suggests. However, analysis is usually best presented in products that embody some amount of synthesis, putting the disassembled pieces back together, arranged by topic. A good analysis product does so in a telling way. We present one example here.

Figure 2 shows two organization charts. A chart of this kind would be made from the respondents’ answers to the questionnaire items regarding reporting, i.e., “who reports to you?” and “to whom do you report?” In an orderly arrangement of a jurisdiction, reporting relationships would be much like that in the “healthy chart” to the upper left. The chart to the lower right, in contrast, shows several pathologies. (Note that, depending on who is in it, the box on the lower right that doesn’t report to anybody may or may not be one of these—it could be that of a Governor.)

---

Figure 2. Organization charts



---

## Assessment

After the reconstruction is complete and the analysis has at least been drafted, it is useful to conduct an assessment conference. Recall that assessment is the step that turns knowledge into power. This conference is not a Hotwash, because its purpose is not to collect the recollections of the participants. On the contrary, the assessment conference occurs after the analysis report is complete, at least in draft form. The participants are sent this report or draft before they come to the conference; then, at the conference, the analysts brief the participants on the results of the reconstruction and analysis efforts.

Assessment addresses the implications of an exercise's findings. Although the analysts may well be aware of them, they cannot make them part of the analysis *per se*, because they follow from the experiment's findings and a knowledge of the real world, including operational, political, and programmatic realities, and the analysts' assignment is to analyze only the experiment. At the assessment conference, the participants and others (e.g., stakeholders of various kinds, and/or officials who did not participate) discuss the analysts' findings and articulate their views of how these findings relate to larger reality.

The assessment conference ought to produce a report separate from that of the analysis effort. The choice of who signs this report can be of crucial importance, for it is this report that bridges the gap from the make-believe world of the exercise back into the real world.



---

## **Points arising in exercise analysis (and assessment)**

One could write an entire book on exercise analysis, but a few topics merit mention here because they arise twice:

1. After the exercise, when doing their analysis, the analysts must take into account that certain actions were forced on the participants by the controllers, deemed by the controllers to turn out in a way other than what might really have happened, done with unrealistic ease by the participants, or simply deemed to have happened with no effort at all on the part of the participants.
2. The issues really arise earlier, during the exercise, because they strike the participants as being unreasonable. “Who cares what the participants think?” is not a viable attitude, because the success of the exercise depends in large part on the participants’ “willing suspension of disbelief,” and if disbelief becomes too hard for them to suspend, they will throw up their hands and behave unrealistically, if at all, spoiling the value of the exercise.

Therefore, the topics that merit some contemplation by the exercise designers include artificialities, the flow of time, and “fair experiment vs. fair fight.”

### **Artificialities**

In numerous respects, the exercise will of necessity be unrealistic—to begin with, organizers are not allowed to disrupt city life, much less set off a real WMD, or cause actual casualties. Instances of unrealism, especially those that might influence the outcome, are called artificialities. In a broad sense, practically any issue arising in exercise analysis is concerned with artificiality, because if it weren’t for artificialities, one could take the outcome of the exercise at face value.

---

A good example—not only of an artificiality, but of the difficulties entailed in dealing with it—was provided by the 1994 terrorist WMD exercise Mirage Gold, as outlined by Kaplan [9]:

“For instance,” [an] official said, “six weeks before this supposedly ‘no-notice’ exercise, the FBI leased 11 T-1 phone lines and installed them in an empty warehouse that it planned on using as a command post.

“Now there are two ways to interpret this,” he added. “The cynical interpretation is that they prearranged things so there’d be no snafus. The charitable interpretation is that when you’re doing an exercise, you’re not going to order AT&T’s global communications network to install all these lines, but you might be able to do that in a real emergency, so why not simulate it?”

“Both interpretations are probably true,” he said. “This is a basic problem with all these tests. You can’t avoid artificiality. How much does that distort your results? It’s hard to say.”

## **The flow of time**

Time inside the exercise is not the same as time outside the exercise: the former starts, stops, and sometimes might flow unnaturally quickly, or even jump backward, all at the behest of the exercise controllers. By accident or design, it can even happen that “exercise time” is different at different places. There is nothing intrinsically wrong with these anomalous flows of time, as long as the analysts are aware of them. There isn’t even anything wrong with backing up and doing some event again, as long as everybody is aware of what is happening, and as long as the “done-over” version is not taken as the only valid outcome: both outcomes occurred, and each deserves attention.

## **“Fair experiment vs. fair fight”**

Sometimes the participants’ actions are constrained by considerations other than the practicalities of conducting an

---

exercise. In particular, sometimes certain courses of action are forbidden by Exercise Control, even though they would make a great deal of sense in a real setting.

Regardless of the rationale given, the frustrated participants (e.g., the OpFor) are likely to say, “they wouldn’t let us do it because they said it wasn’t fair,” and to follow up on this observation with a detailed discussion of why considerations of “fairness” have no place in armed conflict. This attitude almost always comes from an understandable confusion between a misplaced desire to have a “fair fight” and a fully justifiable desire to have an exercise that is a “fair experiment.” The phone-line artificiality example cited above is a good one in this connection, too. The purpose of installing the telephone lines is not to make the contest against the OpFor more equitable; it is to make the exercise a more valid representation of what would actually happen in a real emergency, which is that the Federal Government would make the telephone company drop everything and install telephone lines overnight.

Similarly, the purpose of repeating some event is usually to learn something. The OpFor must be given a clear understanding of this, because their natural reaction will be to say, “They wanted to do it over and over until they found a way to win.”

Finally, especially in an exercise about terrorist WMD attacks, the OpFor are likely to be restrained from exerting what they view as their full threat potential. The reason is that an overwhelming attack in which everybody gets killed is not illuminating, and it doesn't even provide good training. For example, terrorists might conceivably set off a nuclear explosion [11], but a better case to exercise is the also-possible scenario in which they merely detonate a radiological dispersal device. Similarly, they might infect a few thousand people with smallpox, a highly contagious disease that kills a considerable fraction of its victims and cannot be cured once the symptoms develop. But, in the early going, a more fruitful exercise would address the case in which they attempt an anthrax attack, because this disease is non-contagious and treatable.

---



---

## **Appendix A: Exercise influences**

### **TOPOFF 2000**

TOPOFF 2000 was a May 2000 national-level exercise featuring simulated coordinated terrorist attacks on Denver, Colorado, and Portsmouth, New Hampshire, as well as federal play in the National Capital Region. A major purpose, as indicated by the name, was to involve top officials—in this case the president, governors, and mayors, all of whom were responsible to no higher authority, only their constituents.

TOPOFF 2000, though in many respects a great success, was subject to some telling critiques. One frequent comment was that the after-action review provided insufficient evaluation, guidance, and prescription for improvement. Though TOPOFF was Congressionally mandated and funded, a senior Senate staffer who followed it closely complained that the Senate never got the after-action review. One reason was that TOPOFF 2000 did not aspire to answer general questions. Its only “experimental” aspects were limited to questions about the use of radios at Portsmouth and whether the venues—one having taken part in a Federal terrorism-related program and one not—would differ in their levels of success in dealing with their attacks. Neither question seems to have been resolved to anybody’s satisfaction.

However, because TOPOFF 2000 does clearly portray how events might unfold, it points to areas that should be analyzed in future exercises. For example, TOPOFF 2000 participants stated in the after-action review that they had difficulty communicating with one another. Conference calls that took place were cited as unwieldy and unproductive. Even at the time, “The roles, authorities, and even identities of those participating in the calls, as well as the leadership of and agendas for the calls, were unclear.” The use of handheld radios was a specific focus of analytic interest in one of the venues, but the analysis largely fell through for want of data. Observers could not tell how many attempts were needed to make a connection, because they could

---

not tell a repeat attempt from an attempt to reach a different party.

Other comments in the TOPOFF 2000 after-action review address confusion regarding who was in charge. For example, in a scene with the person who had died of septicemic plague, some agencies asserted control because the scene contained evidence regarding a crime, and their mandate included investigating crimes; others asserted control because of the scene's public health implications. Other confusion arose because of questions regarding federal, state, and local jurisdiction.

Another paper written after TOPOFF 2000 (separately from the official after-action review) pointed out that public-health decision-making is normally done by seeking consensus among a large number of experts. This process represents one extreme of the trade-off between accuracy and timeliness in decision-making. Some participants attributed these difficulties to the decision-making processes of public health agencies. According to Inglesby et al.[3]:

One observer commented that "in public health, most decision-making is through democratic processes and consensus building, but for some decisions, this cannot work." Another advised that "conference calls need to be led, you cannot go around the table and let everyone speak their opinion."

"Decisions made on Saturday were reversed on Sunday morning, then reversed again Sunday afternoon," commented one individual. "Reversing decisions back and forth is the antithesis of crisis management and efficient decision-making." Another observer remarked, "The time frame that public health is accustomed to dealing with is not what is needed for bioterrorism. In [this type of crisis], one needs to make decisions quickly. You don't have the luxury of time to do more research."

One observer remarked: "With thousands standing outside hospitals awaiting prophylaxis, public health

---

officials were citing papers...Some from the CDC, state and local health agencies tried to look at this as a standard epidemiological investigation. In absolutely no way would this [scenario allow] a normal epidemiological investigation.”

The observations made by outside observers of TOPOFF 2000’s Denver venue include some telling points not made in the after-action review. The most salient of these is the participants’ non-recognition of the differences between a natural disaster and a terrorist attack. Despite the (simulated) evidence, for example, a doctor answering reporters’ questions in a press conference fell back on a public-health view of *Yersinia pestis*, asserting that “pneumonic plague is rare”: in nature, it *is* rare, but an intentional attack would create a large proportion of pneumonic cases, and, at the time of the press conference, more than 500 such cases (all simulated, some with “paper patients”) had been seen by the Denver hospitals.<sup>2</sup> Similarly, it took days for the one victim of septicemic plague, who was found dead with an aerosol sprayer and who had a known association with a suspicious individual who had given him a sprayer and paid him to install it in the Denver Center (the epicenter of the outbreak), to be connected with the other cases of illness.

CNA observers also noticed that officials’ responses at the (simulated) press conference did not inspire confidence: for example, the fact that simple masks could halt or slow the spread of pneumonic plague did not come out until the officials were prompted by a reporter’s question.

---

2. In nature, the pneumonic form constitutes about 2 percent of the cases. Given that Denver had 500 cases of the pneumonic form, an additional 21,000 cases of the bubonic form would be expected in a natural outbreak. The absence of these accompanying bubonic cases (and a few thousand septicemic cases) should have been a strong indication that an unnatural disaster was under way.

---

## **Dark Winter**

Dark Winter was a two-day seminar-style exercise in June 2001, featuring play by a simulated National Security Council (NSC). An important aspect of the exercise was that the participants portraying the NSC, though not current NSC members, were themselves holders or past holders of high-level positions in the government or military. Likewise, the exercise featured a supporting cast composed of knowledgeable players who were veterans of high-level policy, operational, or media positions.

The scenario was that of a smallpox release in the United States. Considerable care went into the creation of a scenario that was sensible from the standpoint of biology. The purpose of the exercise was to have the mock NSC discuss such issues as the allocation of the nation's scant supply of smallpox vaccine; the imposition of isolation, quarantine, or travel restrictions (including whether this issue should be decided at the federal or state level); the economic impact of the attack; the measures taken in response to it; and the foreign policy implications of a hostile act, including whether the nation is at war.

Dark Winter raised numerous questions regarding the need to establish and keep the goodwill of the public; the role of scientific understanding in decision-making; the shortage of vaccine and the need to distribute it wisely; the relationship of federal and state policies, laws, and authority; and the treatment of the outbreak as an unnatural disaster created by a thinking, planning enemy [6].

## **September 11 and the anthrax incidents**

Certain aspects of these real-world events, subsequent to TOPOFF 2000, served to underscore the lessons of that exercise. This seemed especially true in the case of the anthrax incidents.

As in TOPOFF 2000, multiple agencies asserted authority over the Brentwood mail-sorting facility, some seeing it as being within their purview as a threat to public health, and others

---

seeing it as being within theirs because it was a crime scene. Also, and strikingly in parallel to the press conference in TOPOFF 2000, public health officials initially attempted to portray the outbreak of anthrax as routine, “an isolated case,... not linked to any threats of bioterrorism against the United States,” despite the fact that it was an outbreak of the rare inhalational form of the disease, not seen in the U.S. in the past 25 years and seen in the U.S. only 18 times in the entire 20th century, and despite the fact that this odd occurrence took place in the wake of the terrorist attacks of September 11. Finally, officials’ inconsistent and indefinite statements were seen as jeopardizing their credibility, which could not be regained nearly as quickly as it was lost.



---

## **Appendix B: Sample questionnaires**

This appendix contains preliminary draft questionnaires of the type that might be used in the exercise. An important feature of these questionnaires is that they are quite general: “one size fits all,” and there need not be separate questionnaires for fire-fighters, police, mayors, federal officials, and so on. However, agencies may wish to create and administer such questionnaires for their own purposes.

Questionnaires could be administered in some electronic, web-based manner rather than on paper, but experience has shown that it will not do to ask participants to respond to the questionnaires on their own time. At the end of the exercise day, they are tired, dirty, and anxious to return home, and, by morning, memories will have faded. More value is gained from the exercise by having a slightly shorter exercise day and then administering the questionnaires at the end (and receiving them back before the participants leave) or, in the case of around-the-clock events that do not end at the end of the workday, having a “yellow flag” pause during which participants (including the OpFor) cease all exercise activity and stop to fill out forms, and the activity of any simulated radiation, fires, or germs is suspended until play resumes.

Delaying the completion of forms beyond the end of any 24-hour period greatly threatens the completeness and accuracy of the responses. Memory is fallible, and days tend to blend.

Note that two questionnaires are to be filled out after the event is over. One is analogous to those filled out by the participants during the exercise—it asks what they did, and so on. The other consists of questions regarding the conduct of the exercise. It will of use in the creation and administration of future exercises and might well be of use in interpreting what happens in this exercise if, for example, people's actions turn out to have been heavily influenced by some exercise artificiality. For actual use, these questionnaires would contain a good deal more “white space” than is shown here.

---

---

**Questionnaire for use after the first day**

Name: \_\_\_\_\_

Normal workday telephone number (\_\_\_\_) \_\_\_\_\_; e-mail  
\_\_\_\_\_

EXERCISE EVENT duty station: \_\_\_\_\_

EXERCISE EVENT telephone number (\_\_\_\_) \_\_\_\_\_

EXERCISE EVENT radio frequency or channel \_\_\_\_\_, call sign  
\_\_\_\_\_

What was the emergency?

When did it occur?

Where did it occur?

How did you find out about it?

What did you do?

Whom did you tell?

What did you ask them to do?

By what means did you communicate with them?

What did you do yourself?

What were you asked to do, and by whom?

By what means did they communicate with you?

---

**Questionnaire for use after subsequent days**

Name: \_\_\_\_\_

Normal workday telephone number (\_\_\_) \_\_\_\_\_; e-mail  
\_\_\_\_\_

EXERCISE            EVENT            duty            station:  
\_\_\_\_\_

EXERCISE    EVENT    telephone    number    (\_\_\_)  
\_\_\_\_\_

EXERCISE EVENT radio frequency or channel \_\_\_\_, call sign  
\_\_\_\_\_

With whom are you in touch about this emergency?

Who's working for you?

How do you communicate with them?

For whom are you working?

How do they communicate with you?

What do you need?

Are you getting it? If so, where does it come from?

Have you told the public anything? If so, what?

With which decision-making processes have you been involved?

Who was involved in them with you?

How did you communicate with them?

What was decided?

Why did it come out that way?

---

**Questionnaire for use after the last day**

Name: \_\_\_\_\_

Normal workday telephone number (\_\_\_\_) \_\_\_\_\_; e-mail  
\_\_\_\_\_

EXERCISE                  EVENT                  duty                  station:  
\_\_\_\_\_

EXERCISE EVENT telephone number (\_\_\_\_) \_\_\_\_\_

EXERCISE EVENT radio frequency or channel \_\_\_\_\_, call sign  
\_\_\_\_\_

What was the emergency?

What did you, yourself, do about it?

What did the people working for you do?

How did your actions help those for whom you were working?

How did your actions, and those of the people working for you and the people who were working with you, help the population as a whole?

What, if anything, do you wish you had done differently?

---

**Questionnaire about the conduct of the exercise**

Name: \_\_\_\_\_

Normal workday telephone number (\_\_\_\_) \_\_\_\_\_; e-mail  
\_\_\_\_\_

Did the exercise improve your ability to do your job? Explain.

What, in the exercise, was harder for you than it would have been in a real emergency of the kind portrayed?

What, in the exercise, was easier for you than it would have been in a real emergency of the kind portrayed?

Whom would you have wanted to reach, but you couldn't because they weren't participating in the exercise?



---

## References

- [1] John M. Ford, "Fugue State" (story), in *Under the Wheel: Alien Stars*, Volume III, Baen 1987
- [2] Office for State and Local Domestic Preparedness Support (OSLPDS), Office of Justice Programs (OJP), Department of Justice (DOJ), and the Readiness, Response, and Recovery Directorate, Federal Emergency Management Agency (FEMA), *Top Officials (TOPOFF) 2000 Exercise Observation Report*, August 31, 2001
- [3] Thomas Inglesby, Rita Grossman, and Tara O'Toole, "A Plague on Your City: Observations from TOPOFF"; *Biodefense Quarterly*, Volume 2, Number 2 (September 2000) ([http://www.hopkins-biodefense.org/pages/news/quarter2\\_2.html](http://www.hopkins-biodefense.org/pages/news/quarter2_2.html))
- [4] Domestic Preparedness Program (DPP), Nunn-Lugar-Domenici (NLD) Domestic Preparedness Program exercises
- [5] William V. Pelfrey, William K. Kelley, John W. May, Jr., "Training Strategy for Office for State and Local Domestic Preparedness Support with Implications for WMD Training," September 2001
- [6] Tara O'Toole, Michael Mair, and Thomas V. Inglesby, "Shining Light on 'Dark Winter'," *Clinical Infectious Diseases* Volume 34 (2002), pp. 972-983 (<http://www.journals.uchicago.edu/CID/journal/issues/v34n7/020165/020165.html>)
- [7] R. Nicholas Palarino, "Briefing memorandum for the hearing, Combating Terrorism: Federal Response to a Biological Weapons Attack, scheduled for Monday, July 23, 2001, To Members of the Subcommittee on National Security, Veterans Affairs, and International Relations" (<http://>

---

[www.house.gov/reform/ns/web\\_resources/  
briefing\\_memo\\_july\\_23.htm](http://www.house.gov/reform/ns/web_resources/briefing_memo_july_23.htm)

- [8] Lawrence K. Altman, "At the Health Department, the Messengers Still Stumble," *The New York Times*, October 8, 2002
- [9] Fred Kaplan, "Nuclear Shadow-Mobile Teams Hunt for Atomic Threats," *Boston Globe*, 11 June 2002
- [10] Philip M. Morse and George E. Kimball, *Methods of Operations Research*, Washington, Operations Evaluation Group, 1946, available from the Center for Naval Analyses. 1951 edition recently reprinted by the Military Operations Research Society, Alexandria
- [11] John McPhee, *The Curve of Binding Energy*. New York, Farrar Straus & Giroux, 1974

---

## Bibliography

David S. Alberts and Richard E. Hayes, *Code of Best Practice for Experimentation*, Washington DC, DoD Command and Control Research Program, 2001.

Jeanne Guillemin: *Anthrax: The Investigation of a Deadly Outbreak*. University of California Press, 1999.

Helen Karppi and Brian McCue, "Experiments at the Marine Corps Warfighting Laboratory," *Marine Corps Gazette*, Volume 85 number 9 (September 2001).

Richard Kass, "Joint Warfighting Experiments: Exploring Cause and Effect in Concept Development," *Phalanx*, Volume 35 number 1 (March 2002).

Zhores A. Medvedev, *Nuclear Disaster in the Urals*. New York, W.W. Norton & Company, 1979.

Michael Schrage, *Serious Play*, Boston, Harvard Business School Press, 2000.

United States Marine Corps Warfighting Laboratory, *Experimentation Procedures (X-File 5-12x)*, January 2001.





