Risk Analysis, Vol. 26, No. 6, 2006 DOI: 10.1111/j.1539-6924.2006.00833.x

Predicting Emergency Evacuation and Sheltering Behavior: A Structured Analytical Approach

Matt Dombroski,[1] Baruch Fischhoff,[1*] and Paul Fischbeck[1]

We offer a general approach to predicting public compliance with emergency recommendations. It begins with a formal risk assessment of an anticipated emergency, whose parameters include factors potentially affecting and affected by behavior, as identified by social science research. Standard procedures are used to elicit scientific experts' judgments regarding these behaviors and dependencies, in the context of an emergency scenario. Their judgments are used to refine the model and scenario, enabling local emergency coordinators to predict the behavior of citizens in their area. The approach is illustrated with a case study involving a radiological dispersion device (RDD) exploded in downtown Pittsburgh, PA. Both groups of experts (national and local) predicted approximately 80–90% compliance with an order to evacuate workplaces and 60–70% compliance with an order to shelter in place at home. They predicted 10% lower compliance for people asked to shelter at the office or to evacuate their homes. They predicted 10% lower compliance should the media be skeptical, rather than supportive. They also identified preparatory policies that could improve public compliance by 20–30%. We consider the implications of these results for improving emergency risk assessment models and for anticipating and improving preparedness for disasters, using Hurricane Katrina as a further case in point.

KEY WORDS: Disasters; evacuation; radiological dispersion devices; risk assessment; risk communication

1. INTRODUCTION

Hurricane Katrina showed the importance of understanding how the public will respond to official recommendations during a major emergency. That response reflects the interaction of many factors, including how citizens assess the risk to themselves and their loved ones, what resources they have available, how well the official response is organized, and what other consequences they anticipate (e.g., separation from family members, abandonment of pets, looting of homes, support—or violence—from fellow evacuees, support—or coercion—from authorities and media). Officials need to understand the interplay of these factors in order to create the most effective plans possible, and then to assess the limits of those plans.

We offer a general approach to predicting behavior in emergencies. It assumes that the risk analysis must be informed by the best available social science, while recognizing that general principles of behavior, identified by that research, are interpreted in the context of specific emergencies. For example, people understand some risks better than others, trust some officials more than others, and have greater freedom to act in some situations than others. Our approach addresses this challenge through an iterative process of formal modeling and expert elicitation. We begin with a formal risk analysis of the emergency situation, whose parameters include potentially relevant

[1] Carnegie Mellon University, Engineering and Public Policy, Pittsburgh, PA, USA.
[*] Address correspondence to Baruch Fischhoff, Carnegie Mellon University, Engineering and Public Policy, Pittsburgh, PA, USA; baruch@andrew.cmu.edu.

0272-4332/06/0100-1675$22.00/1

behavioral factors identified from the social science literature. We then elicit judgments regarding these behaviors from subject-matter researchers. These results are used to refine the terms of the scenario and focus it on those variables that appear most important in predicting its unfolding. Finally, we elicit predictions from a sample of emergency coordinators familiar with local conditions, using a structured survey. The approach is illustrated with a radiological dispersion device (RDD) exploded in downtown Pittsburgh, PA.

2. PRIOR STUDIES OF PUBLIC RESPONSE TO OFFICIAL WARNINGS

The social science research literature on public and institutional responses to disasters is extensive.[1–11] This section summarizes the patterns found in contexts similar to that of the RDD scenario described more fully below. These studies deal with Three Mile Island, accidental chemical releases, hurricanes, public shelters, and unanticipated air raid warnings. They show how public behavior depends on the match between warning systems [12,13] and emergencies.[13–15] They suggest that compliance can be improved by integrating hazard detection with hazard communication, having pretested messages available and decentralizing hazard responses.[14,16,17] We summarize some of these trends, and illustrate them with a few specific studies.

Mack and Baker studied public responses to three unanticipated air raid warnings during the height of the Cold War.[2][5] They found that, after high initial concern, most people established that the warnings were false alarms, mainly by conferring with others around them, rather than by relying on official sources. Few, if any, evacuated or engaged in other protective behavior. Other studies have also found the important role of informal communication.[15]

During the Three Mile Island accident in March 1979, the public received conflicting information from public officials, utilities, and the media, causing a high level of mistrust and confusion.[21] The Governor of Pennsylvania issued a limited evacuation order on the third day of the crisis. Prior to that order, about 8% of those within five miles of the plant had left; another 40% did so on that day, with a total of 60% eventually leaving.[8] Despite conflicting reports and poor traffic management,[25] the public evacuated in an orderly manner, consistent with the general finding that most people initially feel disbelief during a major disaster, but generally follow official advice after they confer with people who they trust.[22–24] However, the public retained high levels of mistrust for government, the utility, and the media months afterward.[8,26] Although only pregnant women and children were ordered to evacuate, people tended to leave in families, increasing the evacuation rate above that expected. Similar behavior has been observed in evacuations from chemical releases.[23]

In a review of hurricane evacuation rates, Baker concluded that the most important determinants are actual risk levels, citizens' beliefs that their homes are at high risk (e.g., if they live in mobile homes), and official recommendations and warnings.[27] Baker found that in four major hurricanes, evacuation orders issued for high-risk areas had evacuation participation rates between 80 and 97%, whereas evacuation rates were between 33 and 75% in lower risk areas without evacuation orders. Table I displays evacuation rates for three recent hurricanes, showing high rates for counties closer to major surge areas.[28]

Less research has been conducted on sheltering behavior, especially on sheltering in place (e.g., sheltering at one's work or home). One study found low usage rates for *public* shelters across a wide range of disasters (an average of 14.7% across 23 events ranging from 43.2% in Nanticoke to 0% in Three Mile Island). It found that the best predictors of shelter usage were evacuee characteristics (specifically, low socioeconomic status and elderly age), rather than characteristics of the event or official emergency preparations.[29]

These studies find no evidence of the panic that nonsocial scientists often predict for emergency situations.[30,31] This result echoes a pattern well established in the research literature: people rarely panic; if they do, it is highly localized.[3,8,10,11,15,32–36] For example, at the beginning of the London Blitz, British policymakers, experts, and psychologists were convinced that psychiatric casualties and panic would envelop Londoners when bombing began, yet real instances of panic turned out to be very rare.[35]

3. METHOD

This brief review shows that the response to warnings can range from disbelief[22–24] (the air raid sirens)

[2] On May 5, 1955, air raid warnings sounded in Oakland and San Francisco, CA when a squadron of U.S. military planes over the Pacific Ocean could not be identified.[18] On November 25, 1958, air raid warnings sounded in Washington, DC, after telephone workers accidentally triggered a civil defense warning system.[19] On September 22, 1959, air raid sirens sounded in Chicago celebrating the White Sox's clinching of the American League pennant.[20]

Table I. Hurricane Evacuation Rates for Several Major Hurricanes Hitting the United States

Hurricane	Description		Beach Surge Area	Mainland Surge Area	Nonsurge Areas
Opal	Oct. 1995, Category 3, $2.4B damage to Florida panhandle, United States	Max county	90	66	37
		Avg. county	85	57	30
		Min county	78	39	15
Georges	Sept. 1998, category 4, $3.2B damage to Puerto Rico, Florida Keys, Gulf Coast, United States		48	42	40
Floyd	Sept. 1999, category 4, $2.1B damage to Atlantic seaboard, United States	Max county	90	86	61
		Avg. county	69	52	41
		Min county	34	12	10

Notes: Numbers reported are evacuation rates for selected counties ordered to evacuate.(28) Only average evacuation rates were reported for Georges.

to nearly full evacuation (some hurricanes), depending on the specific context.(15) Unless a context has been observed before, expert judgment is needed to predict the expression of general behavioral principles in it (e.g., the effects of trust, familiarity, family units, and citizen characteristics). When eliciting such judgments, the context must be specified precisely enough that experts' responses can be used to derive explicit predictions and policies. We offer an approach to doing so.

Our approach begins by creating a formal risk model that predicts focal outcomes. In the case study presented below, the model predicts mortality from an RDD attack (see Reference 37 for details). This general model is then instantiated in an emergency scenario (ES1) by assigning a value to each of its elements. Local experts review the scenario for its plausibility in their context, commenting on variables that are missing, unclear, or superfluous. A revised emergency scenario (ES2) is then presented to academic experts in disaster behavior, who both perform a similar structural review and make predictions about the public's behaviors in it. The scenario is revised again, based on their structural comments. This revision (ES3) serves as the basis for a structured survey, used to elicit behavioral predictions from emergency coordinators familiar with the local area and population. Those estimates can then be used to develop emergency response plans and predict their consequences.(38)

4. RDD SCENARIO DEVELOPMENT

This section describes the development of the initial emergency scenario (ES1), based on the formal risk model.(37) An RDD is a conventional explosive wrapped in radiological material. It can cause trauma near the site, spread harmful radioactive material downwind, and force evacuation for an extended period. In our model, an RDD detonation occurs in downtown Pittsburgh. The model includes the physical and social processes believed relevant to predicting the associated health effects. These include the time of day, affecting where people are at the time of the explosion, and compliance rates with official instructions (if any), affecting subsequent population movements. They are represented in an influence diagram, a form of a directed graph whose nodes represent chance, decision, or consequence events and are linked by arrows, indicating interdependencies.(39,40)

Fig. 1 provides an overview of the critical variables in the model. After an RDD explosion, radioactive material is dispersed into the atmosphere and distributed by atmospheric processes. We consider two possible public behaviors: shelter in place[3] and evacuate. We predict these behaviors as a function of time of day, official information, and unofficial information (e.g., observed fires, news media). Exposure risks depend on individuals' actions, others' actions (affecting traffic), and dispersion processes.

We create scenarios by sampling a plausible value from each node in the influence diagram.(42) The physical variables in our model have been extensively specified in previous work, making the selection of plausible values straightforward, once an attack is specified.(37,43) However, the relevant behavioral research has rarely been rendered into the form needed

[3] Sheltering in place will require a second stage recommendation for people to ventilate their homes and offices as soon as the plume passes.(41) We do not evaluate compliance with that recommendation in this article. It is an important topic, which could be pursued with the same methodology.

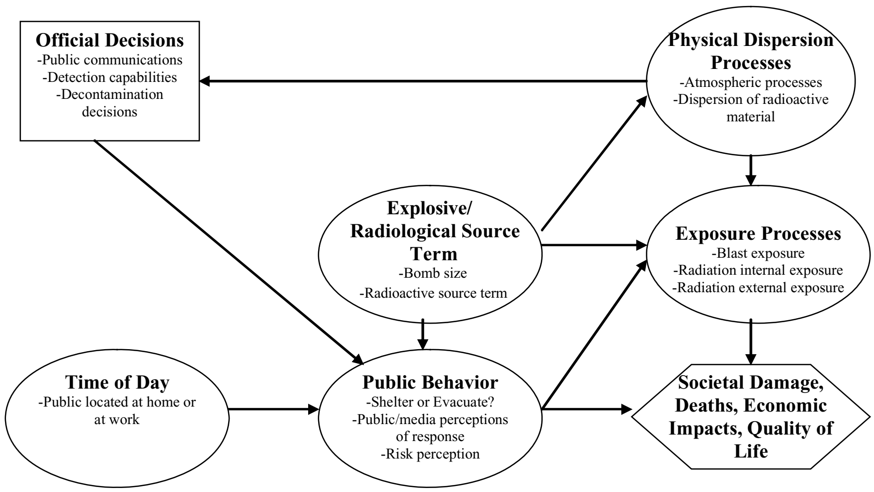

Fig. 1. A schematic diagram of factors determining the outcomes of an RDD event.

for modeling, leading to the expert elicitation reported here.

Our focal scenario assumes that the RDD detonates in downtown Pittsburgh, during a workday, causing casualties at ground zero and dispersing radioactive material downwind, where first responders detect it. A public statement is released half an hour into the incident, but with no mention of radiation. Once radiation is found, the media relay that news.[4] Public officials then issue a statement and recommendation, adapted from the National Council on Radiation Protection and Measurements (NCRP) set of prepared communications for such situations.(43) This initial scenario (ES1) was reviewed with several Pittsburgh-area first responders, leading to the revised scenario (ES2). ES2 appears in Appendix A.

5. OPEN-ENDED INTERVIEWS WITH ACADEMIC EXPERTS

The academic experts reviewed the RDD scenario (ES2), in open-ended interviews, following the mental models methodology, in which respondents are asked increasingly specific questions, directing them to the topics in the formal model, but always expressing themselves in their own terms.(44) We invited comments regarding factors outside the model. These experts also answered quantitative questions, predicting compliance with official orders.

[4] The media often monitor emergency communication channels, so it is likely that as soon as a HAZMAT team reports to the scene, the media will begin reporting the possibility of a hazardous materials release before official statements confirm it.

5.1. Interview Protocol

The academic experts had the opportunity to read the RDD scenario (ES2) in advance. The interview began by reading the scenario, asking them to visualize the situation, thinking aloud as they reviewed the scenario details, commenting on its realism and potentially significant omissions. They then considered variants created by changing the values of three potentially relevant variables:

- ability to see and hear the explosion;
- media influence (supportive or skeptical);
- official recommendations (shelter in place or evacuate).

The experts considered, in turn, variants on the base scenario with the following six combinations of values on these three variables: (1) seeing/hearing the explosion, *supportive* media reports, official advice to *shelter in place*; (2) *not* seeing/hearing the explosion, *supportive* media reports, official advice to *shelter in place*; (3) *not* seeing/hearing the explosion, *skeptical* media reports, official advice to *shelter in place*; (4) seeing/hearing the explosion, *supportive* media reports, official advice to *evacuate*; (5) *not* seeing/hearing the explosion, *supportive* media reports, official advice to *evacuate*; (6) *not* seeing/hearing the explosion, *skeptical* media reports, official advice to *evacuate*.

Predicted compliance with official orders was elicited in terms of best guess, lower bound, and upper bound estimates—under the assumption that the public had heard both the media reports and official advice. These judgments were used to create triangular distributions.

5.2. The Sample

The goal of refining the scenario led us to look for diverse expertise, hoping to ensure that the model's variables were scrutinized from multiple perspectives. About 30 potential experts were identified from our literature review. We narrowed the list to 15; two primary and one alternative from each of five key research areas: (1) disaster sociology, (2) disaster psychology, (3) risk communication, (4) emergency preparedness and community response, and (5) disaster journalism. Each expert had at least three relevant publications. All primary experts agreed to participate. They were told that $100 would be donated to a charity of their choice in exchange for their participation in interviews lasting between 60 and 90 minutes in length.

5.3. Interview Quantitative Results

Fig. 2 shows each expert's high, low, and best estimates of compliance rates for the six scenarios.[5] For example, with Scenario 1, Expert 1's best-guess prediction was that 60% of citizens who had seen or heard the explosion would comply with an official recommendation to shelter in place, given that it was supported by the news media. However, this was an uncertain prediction, as seen by the low and high estimates of 0 and 80%. Expert 1 provided the same estimates for Scenario 2, meaning that it made no difference whether citizens had seen or heard the explosion, and for Scenario 3, meaning that it made no difference whether the media were skeptical. Expert 1 anticipated much greater compliance if the recommendation were to evacuate (Scenario 4), even if citizens had not seen or heard the explosion (Scenario 5). However, skeptical media reduced the best estimate from 95 to 90% and the lower estimate from 90 to 80%. Following the judgments of the other experts across the scenarios reveals their personal theories regarding the importance of the different factors. Looking across the experts shows the range of opinions for each scenario.

Various patterns emerge from these data. Given the small quota sample, we will describe them heuristically, without any attempt at inferential statistics. One clear pattern is that, for all three direct scenario comparisons, the experts expected greater compliance with an evacuation order than with a shelter-in-place order. They were nearly unanimous in predicting high levels of compliance with an evacuation order for people who had seen or heard the explosion and supportive media. In the evacuation scenario, half the experts' low estimates envisioned no less than half of people evacuating. Substantial ranges of estimates were fairly common, indicating a degree of individual uncertainty, in addition to disagreement across experts. These ranges may reflect uncertainty about both how people will respond in the specified scenario and the effects of unspecified variables. The scenario development process was intended to identify the key variables. However, any simplification risks omissions, just as adding variables risks reducing the attention afforded to each.

Most experts had similar judgments, regardless of whether or not people had seen or heard the explosion. A skeptical media prompted most experts to anticipate lower compliance rates. Median compliance for a shelter-in-place order was 60 and 53%, with supportive and skeptical media, respectively. The corresponding medians for an evacuation order were 90 and 83%.

Probability density functions (pdfs) are generated by equally weighting the individual triangular distributions and summing the resulting pdfs. Fig. 3 show them for the shelter and evacuation scenarios, respectively. The experts expected much greater (although still not perfect) compliance with an evacuation order. In their verbal commentary, a common theme was that people typically want to get away from a problem, especially when they are caught away from home. The experts were divided regarding people's willingness to shelter in place, producing a bimodal distribution, with values clustering around 20 and 60%.

5.4. Prescriptive Suggestions

Most experts doubted that seeing or hearing the explosion would seriously affect compliance rates, a belief reflected in Fig. 2. Their individual confidence and collective agreement suggest that plans and scenarios should pay little attention to this variable. Some experts spontaneously added that people would be more reluctant to evacuate from home than from work. Experts generally agreed that seeing or hearing the explosion would have less impact on compliance than whether people were at home or at work. As a result, in ES3, we replaced that variable with the location one. In the Cold War context, Ikle and Kincaid speculated that people knew enough about long-term radioactive contamination to

[5] Expert 9 did not offer quantitative assessments for the skeptical media scenarios (Scenarios 3 and 6), but did provide qualitative comments. Time constraints precluded Expert 3 from providing an assessment for Scenario 6 (where people did not see or hear the explosion, an evacuation order was given, and the media were skeptical).

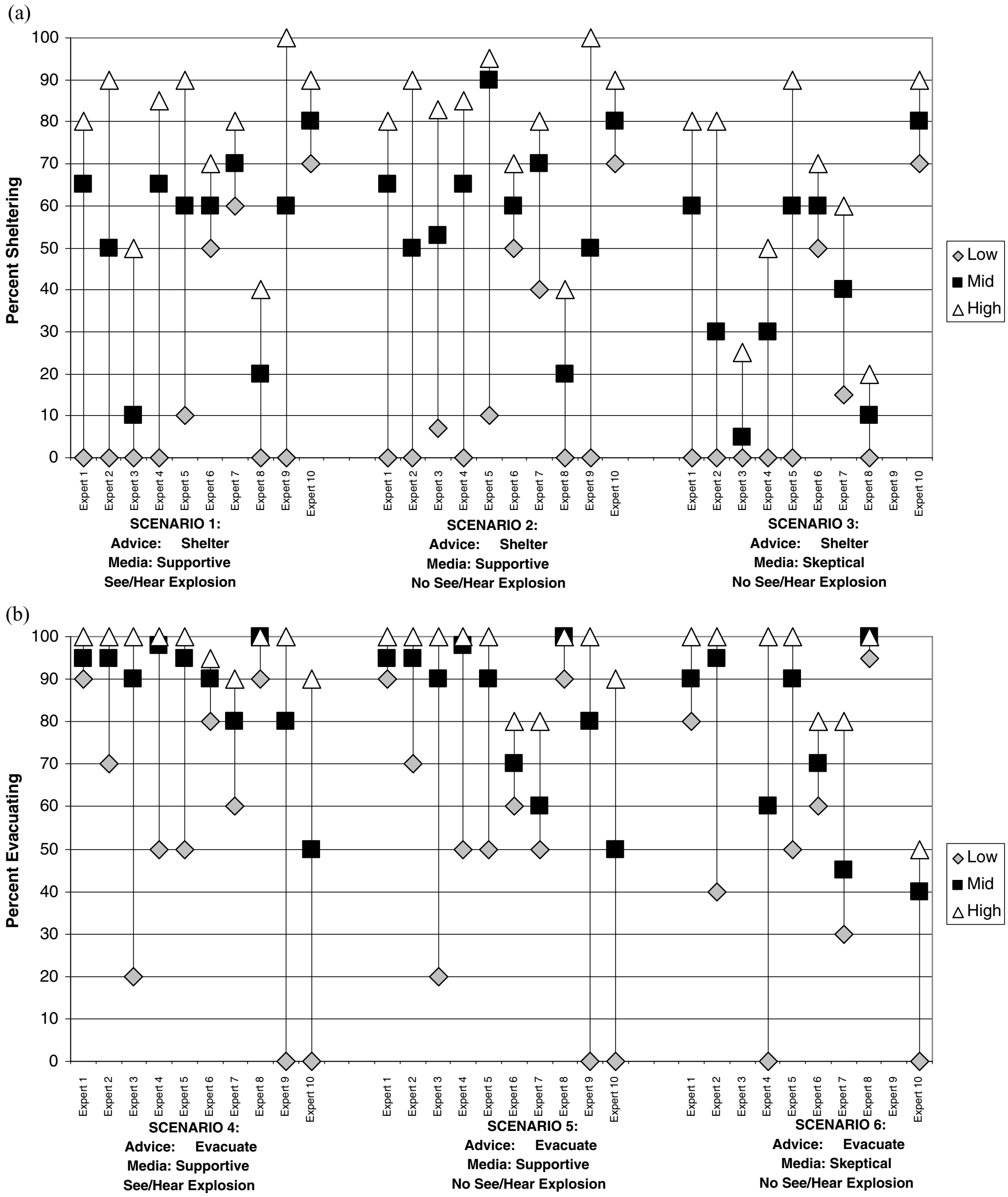

Fig. 2. Academic expert estimates of compliance rates with an order to (a) shelter in place and (b) evacuate.

be reluctant to leave, not knowing when they could return.[1] However, Ziegler, Brunn, and Johnson found that this prospect was not raised when people explained their decisions to stay during the Three Mile Island crisis, with the most frequent reason being "no apparent reason to evacuate" (38%).[25] The fact that the Three Mile Island accident developed slowly, unlike the RDD scenario presented here, is another possible source of difference. Clearly, more research on sheltering behavior would be helpful.

The experts explained the influence of skeptical news media in ways consistent with their judgments. They believed that compliance with either recommendation could be improved by using communications better than the one we adopted from the NCRP. Many experts believed that a media-training program could reduce media skepticism. Another suggestion was having the official message address protective measures taken at local schools, so that worried parents could evaluate whether they had to act immediately in

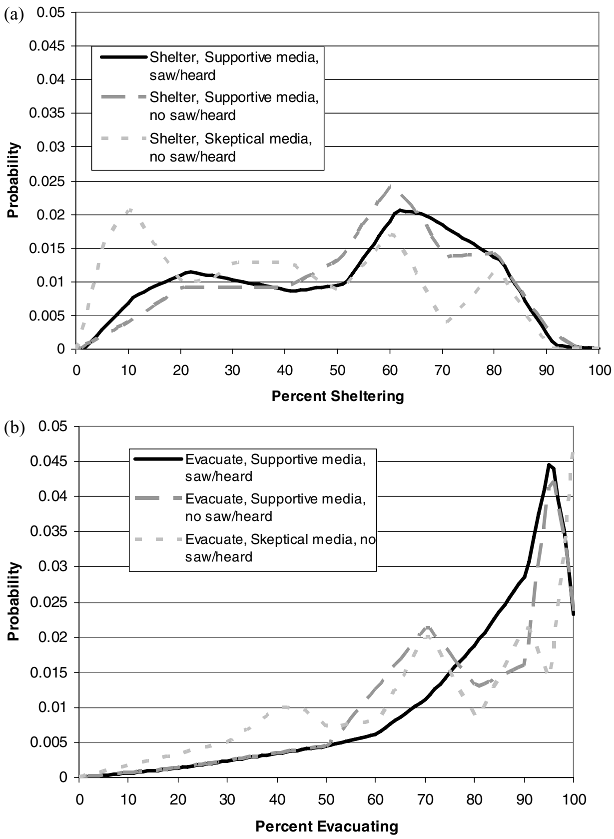

Fig. 3. Aggregated academic expert judgments of public behavior. A pdf for sheltering is shown in (a) and for evacuating in (b).

order to protect their children. In order to give such official assurance credibility, some experts suggested involving parents in school emergency planning. Many experts suggested that workplace drills could make sheltering and evacuation more effective and orders more credible. We asked the emergency coordinators to evaluate these recommendations in the subsequent structured survey.

6. EMERGENCY COORDINATOR QUESTIONNAIRE

Emergency coordinators have indirect familiarity with the research literature, some of which finds its way into their training. They have direct familiarity with the citizens who will face an attack in their area. A good scenario should help them to envision the circumstances in which those citizens might find themselves. The previous research was intended to create such a scenario, consistent with natural and social science understanding of an RDD attack. It formed the basis for the structured survey with local emergency coordinators, designed to get their opinions, with sufficient precision and in sufficient number, to provide inputs to risk analyses.

6.1. Study Design

As described above, we identified three variables as potentially having significant impact on public behavior and health effects:

- where the public was located (i.e., at home or at work);
- media influence (i.e., supportive or skeptical);
- official recommendations (i.e., shelter in place or evacuate).

The questionnaire had respondents evaluate a baseline scenario (ES3 in Appendix B) with one value for each variable (at work, supportive media, and shelter in place), then consider the effects of changing each variable in seven plausible combinations. For each scenario, respondents provided a best guess, high estimate, and low estimate of the percentage of people who would comply with the official advice, assuming that they had seen or heard local media reports and official press releases. This strategy focuses respondents on individual variables and their combinations (which were not emphasized with the academic experts' scenarios because the emphasis of the academic interviews was to identify individual variables affecting behavior).

The second part of the survey asked how four strategies, identified by the academic experts, would affect compliance: media training, emergency drills at work, parent participation in school emergency planning, and first-responder risk-communication training.

The survey concluded with demographic questions that might be related to respondents' judgments, including experience with exercises or emergencies like the scenario, number of years in emergency management, organizational experience (HAZMAT, fire, police, EMS, etc.), and work zip code.

The survey was distributed to 116 staff of the Allegheny County Emergency Management Coordinators (EMC) and the Southwestern Pennsylvania Chapter of the American Red Cross, who were asked to complete it at home and return it by mail. They were told that $15 would be donated to the American Red Cross or the Salvation Army for each returned survey and that completion should take 30 to 60 minutes. A reminder email was sent two weeks later. Thirty-three questionnaires were returned for a response rate of

Table II. Best-Guess Compliance Rates for Eight Scenarios, Elicited from Emergency Coordinators.

Scenario Number	Official Order	Media Influence	Location of Public	Mean Compliance Rate (%)	Standard Deviation
1	Shelter	Supportive	Work	58.2	19.2
2	Shelter	Supportive	Home	71.4	18.4
3	Shelter	Skeptical	Work	46.9	23.0
4	Shelter	Skeptical	Home	57.4	20.2
5	Evacuate	Supportive	Work	80.0	15.4
6	Evacuate	Supportive	Home	67.3	18.9
7	Evacuate	Skeptical	Work	71.0	16.6
8	Evacuate	Skeptical	Home	60.8	20.3

28%. One was incomplete and discarded. Respondents reported an average of nine years of work experience in emergency management (three had less than a year; 12 had more than 15 years). They had participated in an average of four incidents or exercises that they judged as similar to the research scenario ES3 (five had been in zero; six in at least ten). Due to the concentration of chemical industry in the Pittsburgh area, local emergency management is especially qualified to deal with hazardous materials, expertise relevant to this scenario.

6.2. Confirmatory Questionnaire Results

6.2.1. Emergency Coordinator Compliance Estimates

Table II summarizes best-guess compliance rates for the eight scenarios. Like the academic experts, the emergency coordinators expressed a considerable range of beliefs regarding compliance rates. As with the experts, the emergency coordinators expected greater compliance when the order called for evacuation than for sheltering in place. Pair-wise comparisons of best-guess compliance rates found three significant differences. These respondents expected *greater* compliance (1) with an evacuation order for people at work than at home, both with a supportive media (13.7% difference) and a skeptical one (10.4%), both $p < 0.0005$; (2) with a sheltering order for people at home than at work, both with a supportive media (14.3% difference), $p < 0.0005$, and a skeptical one (10.6%), $p < 0.05$; and (3) when the media are supportive rather than skeptical, when the order is to evacuate work (9.2%), $p < 0.005$, evacuate home (6.0%), $p < 0.05$, shelter at work (12.3%), $p < 0.005$, or shelter at home (16.0%), $p < 0.0005$. Thus, they believe that both location and the media strongly influence decision making with either recommendation.

6.2.2. Emergency Coordinator Behavioral Distributions

Fig. 4 aggregates response distributions for the eight scenarios, using the same procedures as before (triangular distributions, equal weighting). Figs. 4a and 4e (sheltering and evacuating from work given supportive media) resemble corresponding ones for the academic experts: a bimodal distribution for sheltering (Fig. 4a) and high (but still incomplete) compliance for evacuation (Fig. 4e). When the media are skeptical, both distributions move downward, as seen in Figs. 4c and 4g. The emergency coordinators also agreed with the academic experts in predicting greater willingness to shelter at home and less willingness to evacuate home, with 10–15% differences (Figs. 4b, 4d, 4f, and 4h).

We split the sample according to six demographic variables: (1) four or more similar incidents ($n = 13$) versus fewer incidents ($n = 19$); (2) working for more than 15 years in emergency management ($n = 12$) versus working less ($n = 19$); (3) American Red Cross ($n = 9$) versus EMC ($n = 23$); (4) fire fighters ($n = 22$) versus others ($n = 10$), (5) working in more than one area of emergency management (HAZMAT, EMS, fire, police) ($n = 14$) versus working in just one ($n = 18$), and (6) Pittsburgh workers ($n = 18$) versus others ($n = 13$). We compared their best-guess estimates, using two-sample t-tests.[44] Six group comparisons for eight scenarios meant 48 statistical tests. Only four of those produced statistically significant differences (at $\alpha = 0.05$). Because of the chance of observing that many by chance variation, we will not report them.

6.2.3. Prescriptive Polices Suggested by Experts

Respondents estimated compliance rates with a shelter-at-work order for each of the four initiatives proposed by the academic experts, assuming supportive media and implementation one year prior to the

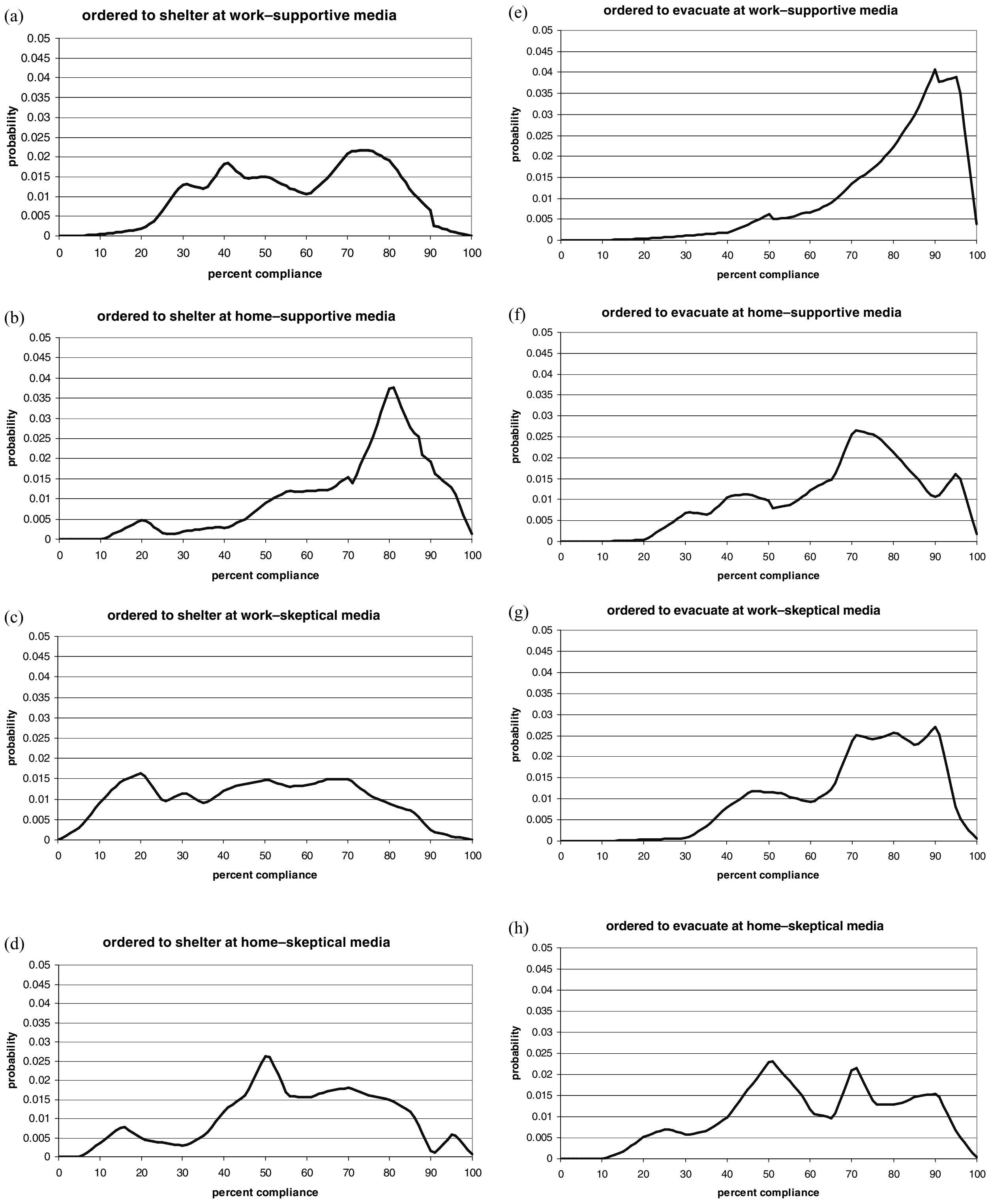

Fig. 4. Probability density functions for compliance with official orders to shelter in place or evacuate.

RDD emergency. For each initiative, they predicted significant ($p < 0.001$) increases above baseline compliance (58%), with mean estimated compliance rates of 72% for media training, 74% for practice at work, 71% for risk-communication education, and 73% for parent involvement in school emergency plans. They estimated that implementing their two top polices would increase compliance to an average of 81%.

7. DISCUSSION

7.1. General Approach

Responsible emergency plans must make realistic assumptions regarding public behavior. We have demonstrated an approach for extending risk analysis to address this need. It begins by creating an analytical model incorporating scientific knowledge about physical processes (e.g., atmospheric dispersion, dose-response relationships), noting the roles of human behavior in determining health outcomes. Social science research is reviewed to identify the factors most relevant to predicting those behaviors. The model is then instantiated with scenarios created by sampling plausible values for each of its variables.

Academic experts predict the focal behaviors (here, compliance rates) in the contexts described by the scenarios. Each expert provides a range of estimates, so that it is possible to capture uncertainty at the individual level, as well as disagreement at the group level. The experts also explain their estimates, evaluate the plausibility of the scenario, and suggest ways to improve warning systems. These qualitative and quantitative judgments are used to refine the model and the representative scenarios, focusing them on central issues.

Quantitative estimates of the focal behavior (here, compliance rates) in the revised model and scenario are then elicited with a structured survey, suitable for use with a larger sample. Here, we selected local emergency coordinators as experts, given their understanding of local conditions and expertise in hazardous material incidents. The survey includes a revised baseline scenario, alternative scenarios created by varying individual parameters (identified by the academic experts), and questions about the efficacy of selected interventions (also identified by the academic experts). This division of labor assumes that the academic experts are best suited to draw attention to variables that have generally proven important, while the local experts are best suited to anticipate how things actually play out. Conducting analogous surveys with members of the public would reveal how credible such initiatives seem to them. If they disagree with the experts, then they may know something that the experts do not or they may need extra convincing to accept the initiatives.

7.2. Specific Results

Both the academic experts and the emergency coordinators predicted fairly high compliance rates, about 70–80% for an evacuation order and 60–70% for sheltering in place (averaged across scenarios). They predicted about 10% higher compliance for sheltering at home and evacuating work, compared to sheltering at work and evacuating homes. They predicted about a 10% reduction in compliance should the media be skeptical, rather than supportive. The academic experts did not think that it mattered much whether people had actually seen or heard the explosion (leading us to drop to this variable from the model). The emergency coordinators thought that sheltering compliance could be increased by 10–15% by implementing any one of four preparatory measures identified by the academic experts: media training programs, evacuation and sheltering drills at work, parent participation in emergency planning at schools, and first-responder risk-communication programs. They estimated that their two top programs could produce a 20–25% improvement (taking mean compliance from 58 to 81%).

Although our two groups of experts considered somewhat different scenarios, there was enough similarity that it might be instructive to compare their judgments formally. A Kolmogorov-Smirnov goodness-of-fit test (K-S test)[46] was conducted comparing the two cumulative probability distributions for the six most similar scenarios. There were no statistically significant differences. Recognizing that the small sample sizes (10, 32) provided statistical power for detecting only large differences, we looked for the greatest apparent differences between the two groups' judgments. There were apparent differences in sheltering predictions, suggesting the value of research to improve understanding of such behavior.

7.3. Implications for Behaviorally Realistic Risk Analysis of Hypothetical Emergencies

Risk predictions for many emergencies hinge on citizens' responses to orders. The resulting estimates can be used as parameters in models guiding risk management, taking proper care to reflect the uncertainty expressed by individuals and the disagreement expressed by the group. Doing so addresses the customary goal of risk analysis in a behaviorally realistic way: identifying the best possible policies, with an understanding of how much faith to place in them. Although the present results reflect a specific attack, our model is readily adapted to other Pittsburgh-area scenarios. With incrementally greater effort, the model could be applied to RDDs in other locales,[37,38] while the approach could be adapted to other hazards. Using this model, Dombroski and Fischbeck[37] showed that in some worst-case RDD event scenarios, risk can vary by more than a factor of four, depending on public compliance with official orders. Other emergency scenarios, such as hurricanes and nuclear disasters, might be much more sensitive. Potential impacts can be examined by eliciting expert judgments in a form suited to incorporation in risk models.

Of course, models can also be used to improve plans as well as evaluate them. In our scenario, like many others, casualties will depend on traffic flows, as they affect emergency personnel's ability to get to injured people and others' ability to evacuate. Urbanik[26] provides models for assessing the adequacy of existing transportation networks for hypothetical emergency evacuations. Our distributions of compliance rates can easily be incorporated into these traffic models to predict transportation system performance in an emergency in a way that also considers the effects of securing (or losing) public trust (and compliance).

7.4. A Coda on Katrina

Although we focused on a hypothetical RDD scenario, our approach could be applied to any risk analysis where behavior plays a central role. Our specific results would be most relevant to situations that closely resemble the RDD one. A natural question, as of this writing (September 9, 2005), is how it would have applied to predicting the outcome of Hurricane Katrina. The following reflections are necessarily speculative—perhaps tainted by the wisdom of hindsight, perhaps lessened by the fact that the full story of Katrina has yet to be revealed.

Our academic expert interviews were conducted early in 2005, our local experts' surveys in June and July 2005. Despite the many differences between Katrina and our scenario, our experts' predictions were roughly accurate: about 80% compliance with an evacuation order. Their judgments, doubtless, drew on the research literature on disaster behavior, which, to varying degrees, they had studied—and created. Hurricanes are among the most studied disasters.[47–49] As mentioned, Baker[27] summarizes evidence regarding the impacts of many variables, including risk area, evacuation notices, housing, storm threat information, hurricane probability forecasts, hurricane experience, length of residence, hurricane awareness, crying wolf, and demographics. Based on that research, we anticipate that our academic experts would have guided us to create a Katrina model that included some of the following variables (with their possible effects on predicted evacuation rates in parentheses):

1. timing of warning (weekend timing should require more people to evacuate homes);
2. amount of advance warning (should increase compliance, with the amount depending on citizens' and authorities' response capabilities);
3. condition of roads (actual and anticipated disruption should increase early evacuation, reduce late evacuation)
4. demographics (poor, urban, and elderly population should reduce evacuation and increase public shelter use);
5. degree of interagency coordination (complexity should increase actual and perceived confusion, increasing media and public skepticism, reducing compliance).

Thus, even without benefit of the information on organizational disarray that is emerging in reports on the disaster, one might have predicted problems that required an especially high level of organizational performance. Incorporating behavioral research in a formal model provides guidance on improving such performance.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (SES-0433152 & SBR-9521914) and the MacArthur Foundation. We are grateful for comments from Wändi Bruine de Bruin, Julie Downs, Keith Florig, Granger Morgan, and Victor Weedn. We are also grateful for feedback from Dr. Sara Eggers and Dr. Shalini Vajjhala. We owe a special debt to the 10 social science experts and 32 emergency coordinators who participated in our studies. We are grateful for the assistance provided by Lucille Underwood, Dennis Narey, Dave Johnson, Mike Adametz, and the PA Region 13 Joint Terrorism Task Force. The views expressed are those of the authors.

APPENDIX A Initial Baseline Scenario for Academic Experts (ES2)

Time	Description
10:00 AM	It is midmorning on a spring workday. The temperature is 50°F with partly cloudy skies in Pittsburgh, PA.
10:30 AM	A relatively large, but concentrated explosion occurs at the U.S. Steel Tower in downtown Pittsburgh. People in several buildings are injured from the explosion, windows in neighboring buildings are blown out, and there is a measurable amount of debris and a large cloud of dust in and around the streets surrounding the site of the explosion. The building at the source of the blast has suffered significant damage, but appears to be structurally intact. A large cloud of dust is covering the neighboring streets.
10:37 AM	First responders are beginning to report from the scene of the explosion. They report that fires are burning in the building and neighboring buildings. There are numerous casualties and injuries at the scene.
10:38 AM	The battalion chief from District 2 originating from Squirrel Hill reports a spike on his radiation monitor of 50 mRem/hr at the corner of Ross Street and Center Avenue.
10:40 AM	The local news media begins continual coverage on the incident, breaking into daytime broadcasts. The media are reporting objectively from the scene of the incident and there is some speculation of a bomb or gas leak.
10:41 AM	The battalion chief from District 3 originating from Highland Park reports a spike on his radiation monitor of 100 mRem/hr at the corner of 7th Avenue and Grant Street.
10:52 AM	All emergency personnel have reported to the scene. There are some reports of unusual radiation readings and there are some reports of nothing out of the ordinary.[a] Emergency Services requests the Specialized Intervention Team HAZMAT Response: Green Team to perform an initial radiation atmospheric assessment in and around downtown.
10:53 AM	The on-scene commander of the incident (fire bureau) releases the following public statement: *"At about 10:30 AM this morning an explosion occurred at the US Steel Tower in downtown Pittsburgh. First responders are at the site of the incident and have begun rescue efforts and have set up triage and treatment units. Our focus right now is on fire suppression and the treatment of victims. We have begun a full investigation into the cause of this incident and we have notified the proper state and federal agencies to assist in response to this incident. We will keep the public informed as more information becomes available."*
10:56 AM	A local news station begins reporting about a potential radiological release from the incident. Other news stations quickly follow suit.
11:05 AM	The media continue their coverage with interviews of victims from the blast, family members, and an occasional public official who reiterate the fire bureau's public statement and contend that they are currently unaware of any toxic release from the blast. The media continue to speculate as to the cause of the incident and there is continued speculation of a radiological release.
11:10 AM	Emergency Services and DEP release an initial health alert describing the situation and various abnormal radiation readings around Ground Zero, with limited information from the Green HAZMAT team. A recommendation is given to the mayor's office to protect the public. The mayor consults with EMA, DEP, and information officers at the EOC. The Allegheny County Health Department, PEMA, FEMA, and other respective state and federal agencies are notified that the city will recommend that the public shelter in place.
11:30 AM	The mayor, EMA, and DEP release a joint public statement. The statement provides specific recommendations for how the public should respond: *"We have reason to believe that the incident this morning may be an intentional act. In the interest of public safety and law enforcement, the area around the incident site is being monitored and a barrier is being established around it. Radioactive material may have been released from the incident. The highest levels of radiation would be around the incident site, but some radiation may have been carried downwind of the incident. As a precaution, the public living and working in Pittsburgh city limits is advised to remain in their current locations and do not venture outdoors unless it is necessary. Please turn off ventilation systems with the outside air and monitor local news and radio for official updates. If your current building has broken windows or doors, we recommend that you evacuate to buildings without broken windows or doors."*
11:40 AM	A handful of first responders begin patrolling streets to assist the public with the recommendation. Loud speakers are used to inform the public and the emergency broadcast system is initiated.
11:45 AM	The media begin interviewing experts who are sympathetic to the mayor's viewpoint and recommendation. They speculate about the risk from the toxic release, but the media are generally supportive of the mayor's advice.

[a] Ferguson *et al*. (2003)[50] provides information on commercial radioactive sources that might be used in an RDD event.

APPENDIX B Revised Baseline Scenario for Emergency Coordinators (ES3)

10:00 AM	It is midmorning on a spring workday. The temperature is 50°F with partly cloudy skies in Pittsburgh. The U.S. Department of Homeland Security's national terror alert level has been at Yellow (Elevated) for many months.
10:30 AM	A large explosion occurs outside the 70-story U.S. Steel Tower in downtown Pittsburgh. There are numerous casualties and injuries at the scene. There is a large amount of debris in the plaza in front of the building and on surrounding streets. Fires are burning in the U.S. Steel Tower and several neighboring buildings, although all buildings seem structurally intact. A thick cloud of dust covers a three-block area.
10:37 AM	First responders begin to report to the scene.
10:38 AM	A battalion fire chief, on the way to the scene, reports a spike on his radiation monitor of 50 mRem/hr two blocks east of the blast site.
10:40 AM	Local news media begin continual coverage on the incident, breaking into daytime broadcasts. Local TV networks show live reports by reporters and news helicopters. Reporters interview bystanders near the incident site. There is some speculation about the cause (intentional bombing or accidental gas leak).
10:52 AM	Many local emergency personnel are on the scene. A one-block perimeter around the site is established and cordoned off to block access by media, the public, and others. Superintendents of buildings within the perimeter are ordered to begin evacuation. Evacuees are told to leave the cordoned area. Rescue efforts begin in the U.S. Steel Tower's first floors. Fire companies work on fires within the perimeter. Some responders near the scene report unusual radiation readings. Emergency Services (ES) requests a HAZMAT team to perform a radiation assessment around the scene.
10:53 AM	The on-scene commander (fire bureau) releases the following public statement: *"At about 10:30 AM this morning, an explosion occurred at the US Steel Tower in downtown Pittsburgh. First responders are at the scene and have begun rescue efforts. They have set up triage and treatment units for the injured. Our focus right now is on controlling fires and attending to the victims. We have begun a full investigation into the cause of this incident and have notified the proper state and federal authorities. We will keep the public informed as soon as more information becomes available."*
10:55 AM	Several national TV networks interrupt their broadcasts with news of the incident.
10:56 AM	A TV news reporter at the scene reports the possibility of a radiological release.
11:05 AM	The media interview public officials, who repeat the fire bureau's statement denying any knowledge of a toxic release. The media speculate about a possible radiological release. The media also continue to interview victims near the incident site.
11:10 AM	The Mayor, in consultation with ES, the Department of Environmental Protection (DEP), and emergency managers, drafts an announcement, recommending that the public shelter in place. County, state, and federal agencies are notified.
11:30 AM	The Mayor, ES, and DEP release a joint public statement: *"We are continuing to investigate the cause of this morning's explosion at the US Steel Tower. In the interest of public safety and law enforcement, the area around the incident site is being monitored and a barrier is being established around it. At this time, there are indications that some radioactive material may have been released. If this is confirmed, then the highest radiation levels would be in the immediate vicinity of the explosion, but some radiation could be carried several miles downwind. As a precaution, people living and working in Pittsburgh city limits are strongly advised to remain inside at their current locations and not to go outdoors unless it is an emergency situation and absolutely necessary."*
11:35 AM	The media begin interviewing experts, who support the Mayor's recommendations. Radio and television reports speculate about the risk from the toxic release, while supporting the Mayor's recommendation to shelter in place.
11:40 AM	Loud speakers on emergency vehicles inform the public of the Mayor's recommendation. The emergency broadcast system is also initiated.

REFERENCES

1. Ikle, F. C., & Kincaid, H. V. (1956). *Social Aspects of Wartime Evacuation of American Cities*. Disaster Study Number 4—Publication 393, Committee on Disaster Studies. Washington, DC: National Academy of Sciences—National Research Council.
2. Wolfenstein, M. (1957). *Disaster: A Psychological Essay*. Glencoe, IL: Free Press.
3. Ikle, F. C. (1958). *The Social Impact of Bomb Destruction*. Norman, OK: University of Oklahoma Press.
4. Baker, G. W., & Rohrer, J. H. (Eds.). (1960). *Symposium on Human Problems in the Utilization of Fallout Shelters*. Disaster Study Number 12—Publication 800, Disaster Research Group. Washington, DC: National Academy of Sciences—National Research Council.
5. Mack, R. W., & Baker, G. W. (1961). *The Occasion Instant: The Structure of Social Responses to Unanticipated Air Raid Warnings*. Disaster Study Number 15—Publication 945, Disaster Research Group, Washington, DC: National Academy of Sciences—National Research Council.
6. Baker, G. W., & Cottrell, L. S. Jr. (Eds.). (1962). *Behavioral Science and Civil Defense*. Disaster Study Number 16—Publication 997, Disaster Research Group. Washington, DC: National Academy of Sciences—National Research Council.
7. Canter, D. (Ed.). (1980). *Fires and Human Behavior*. Chichester, UK: John Wiley & Sons.
8. Houts, P. S., Cleary, P. D., & Hu, T. (1989). *Three Mile Island Crisis: Psychological, Social, and Economic Impacts on the*

Surrounding Population. University Park, PA: Pennsylvania State University Press.
9. Ursano, R. J., McCaughey, B. G., & Fullerton, C. S. (Eds.). (1994). *Individual and Community Responses to Trauma and Disaster: The Structure of Human Chaos*. Cambridge, UK: Cambridge University Press.
10. Silver, R. C., Holman, E. A., McIntosh, D. N., Poulin, M., & Gil-Rivas, V. (2002). Nationwide longitudinal study of psychological responses to September 11. *Journal of the American Medical Association*, *288*, 1235–1244.
11. Tierney, K. (2003). Disaster beliefs and institutional interests: Recycling disaster myths in the aftermath of 9-11. In L. Clarke (Ed.), *Research in Social Problems and Public Policy, Volume 11, Terrorism and Disaster: New Threats, New Ideas*. St. Louis, MO: Elsevier Science Ltd.
12. Sorensen, J. H. (1991). When shall we leave? Factors affecting the timing of evacuation departures. *International Journal of Mass Emergencies and Disasters*, *9*, 153–165.
13. Mileti, D. S., & Sorensen, J. H. (1990). *Communication of Emergency Public Warnings: A Social Science Perspective and State-of-the-Art Assessment*. Technical Report ORNL-6609, Prepared for the Federal Emergency Management Agency by the Oak Ridge National Laboratory, Oak Ridge, TN.
14. Tierney, K. J., Lindell, M., & Perry, R. W. (2001). *Facing the Unexpected: Disaster Preparedness and Response in the United States*. Washington, DC: Joseph Henry Press.
15. Sorensen, J. H. (2000). Hazard warning systems: Review of 20 years of progress. *Natural Hazards Review*, *1*, 119–125.
16. Lindell, M. K., & Perry, R. W. (1992). *Behavioral Foundations of Community Emergency Management*. Washington, DC: Hemisphere Publishing Corp.
17. Sorensen, J. H., & Rogers, G. O. (1988). Local preparedness for chemical accidents: A survey of U.S. communities. *Industrial Crisis Quarterly*, *2*, 89–108.
18. Scott, W. A. (1955). *Public Reaction to a Surprise Civil Defense Alert*. Ann Arbor, MI: Survey Research Center.
19. Baker, G. W. (1959). *Operation 4:30: A Survey of the Responses to the Washington, DC False Air Raid Warning*. Washington, DC: Office of Civil & Defense Mobilization.
20. Katz, E. (1960). *Joy in Mudville: Public Reactions to the Surprise Sounding of Chicago's Air Raid Sirens*. Chicago, IL: National Opinion Research Center.
21. Berry, L., Jones, A., Powers, T., & Savada, A. M. (1999). *Media Interaction with the Public in Emergency Situations: Four Case Studies*. Washington, DC: Federal Research Division, Library of Congress.
22. Mawson, A. R. (2005). Understanding mass panic and other collective responses to threat and disaster. *Psychiatry*, *68*(2), 95–113.
23. Rogers, G. O., & Sorensen, J. H. (1989). Warning and response in two hazardous materials transportation accidents in the U.S. *Journal of Hazardous Materials*, *22*, 57–74.
24. Janis, I. L., & Mann, L. (1976). Coping with decisional conflict: An analysis of how stress affects decision-making suggests interventions to improve the process. *American Scientist*, *64*, 657–667.
25. Urbanik, T. (2000). Evacuation time estimates for nuclear power plants. *Journal of Hazardous Materials*, *75*, 165–180.
26. Ziegler, D. J., Brunn, S. D., & Johnson, J. H. Jr. (1981). Evacuation from a nuclear technological disaster. *Geographical Review*, *71*(1), 1–16.
27. Baker, E. J. (1991). Hurricane evacuation behavior. *International Journal of Mass Emergencies and Disasters*, *9*, 287–310.
28. United States Army Corps of Engineers. (2004). *Hurricane Evacuation Study Program*, Mobile District, AL. Available at http://www.sam.usace.army.mil/hesdata/.
29. Mileti, D. S., Sorensen, J. H., & P. W. O'Brien (1992). Toward an explanation of mass care shelter use in evacuations. *International Journal of Mass Emergencies and Disasters*, *10*(1), 25–42.
30. Fischhoff, B., Gonzalez, R. M., Small, D. A., & Lerner, J. S. (2003). Evaluating the success of terror risk communications. *Biosecurity and Bioterrorism: Biodefence Strategy, Practice, and Science*, *1*(4), 255–258.
31. Levi, M. A., & Kelly, H. C. (2002). Weapons of mass disruption. *Scientific American*, 76–81.
32. Janis, I. L. (1951). *Air War and Emotional Stress*. Westport, CT: Greenwood Press.
33. Quarantelli, E. L., & Dynes, R. R. (1972). When disaster strikes (it isn't much like what you've heard and read about). *Psychology Today*, *5*, 66–70.
34. Glass, T. A. (2001). Understanding public response to disasters. *Public Health Reports*, *2*(116 Supplement), 69–73.
35. Wessely, S. (2005). Don't panic! Short and long term psychological reactions to the new terrorism: The role of information and authorities. *Journal of Mental Health*, *14*(1), 1–6.
36. Fischhoff, B. (2005). A hero in every aisle seat. *New York Times*, August 7, Sec. 4, pg 13, col. 1.
37. Dombroski, M., & Fischbeck, P. S. (2006). An integrated physical dispersion and behavioral response model for risk assessment of radiological dispersion device (RDD) events. *Risk Analysis*, *26*(2), 501–514.
38. Dombroski, M. (2006). *Behaviorally Realistic Risk Assessments of Radiological Dispersion Device Events*. Working paper, Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA.
39. Howard, R. A., & Matheson, J. E. (1984). Influence diagrams. In R. A. Howard & J. E. Matheson (Eds.), *The Principles and Applications of Decision Analysis, Vol. II*. Menlo Park, CA: Strategic Decisions Group.
40. Clemen, R. T., & Reilly, T. (2001). *Making Hard Decisions*. Pacific Grove, CA: Duxbury Press.
41. Rogers, G. O., Watson, A. P., Sorensen, J. H., Sharp, R. D., & Carnes, S. A. (1990). *Evaluating Protective Actions for Chemical Agent Emergencies*. Technical Report ORNL-6615, Prepared for the Department of the Army by the Oak Ridge National Laboratory.
42. Pomerol, J. C. (2001). Scenario development and practical decision making under uncertainty. *Decision Support Systems*, *31*, 197–204.
43. National Council on Radiation Protection and Measurements. (2001). *Management of Terrorist Events Involving Radioactive Material*. NCRP Report No. 138.
44. Morgan, M. G., Fischhoff, B., Bostrom, A., & Atman, C. J. (2002). *Risk Communication: A Mental Models Approach*. Cambridge, UK: Cambridge University Press.
45. Devore, J. L. (1995). *Probability and Statistics for Engineering and the Sciences, 4th ed*. Pacific Grove, CA: Duxbury Press.
46. Arsham, H. (1996). *Statistical Thinking for Managerial Decisions, 9th ed*. National Science Foundation. Available at: http://home.ubalt.edu/ntsbarsh/Business-stat/opre504.htm.
47. Dow, K., & Cutter, S. L. (1998). Crying wolf: Repeat responses to hurricane evacuation orders. *Coastal Management*, *26*, 237–252.
48. Dow, K., & Cutter, S. L. (2002). Emerging hurricane evacuation issues: Hurricane Floyd and South Carolina. *Natural Hazards Review*, *3*, 12–18.
49. Lindell, M. K., Lu, J. C., & Prater, C. S. (2005). Household evacuation decision making in response to Hurricane Lili. *Natural Hazards Review*, *4*, 176–185.
50. Ferguson, C. D., Kazi, T., & Perera, J. (2003). *Commercial Radioactive Sources: Surveying the Security Risks*. Occasional Paper No. 11, Center for Nonproliferation Studies, Monterey Institute of International Studies.