The impacts of disaster exercises on participants

Danny M. Peterson and Ronald W. Perry

The value of conducting disaster exercises is underscored in virtually all discussions of disaster planning (Drabek, 1990; Michaels, 1996). Disaster exercises are mandated by legislation and executive rules in connection with a variety of natural and technological threats in England, Australia, Canada and the United States (Selvarajah, 1993; Daines, 1991). In fact, exercises are mandated by the governments of most industrialized countries. This reflects the belief that exercises, sometimes called drills, are a valuable part of the process of emergency management. Yet it remains that there is very little research that demonstrates the precise benefits of exercises. While there is some literature on military applications (Shelton and Sifers, 1994), quantitative studies of the value of exercises in non-military applications have been infrequent (Peterson, 1996). The purpose of this paper is to explore and document one dimension of the value of exercises: the reactions of participants. The research design uses two different disaster exercises, one built around a hazardous materials accident in an industrial firm, and the other developed as a means of testing a medical triage system for mass casualty events.

In general, there are five principal benefits attributed to disaster exercises. First, they permit inferential testing of the adequacy of a disaster plan. That is, since an exercise represents an event response designed using premises and resources described in a plan, the extent to which the exercise is successful demonstrates the efficacy of the plan. Second, exercises allow – again by inference – testing of the adequacy of training of personnel. Exercises involve incident scenarios that are intended to replicate the demands on knowledge, skills and abilities posed by real events, and which an effective training program should create. To the extent that exercise performance uncovers gaps in participant knowledge, skills and abilities, the relevant training programs can be deemed to require revision. Third, exercises may be publicized in the community consequently enhancing the visibility of the agencies involved (demonstrating their readiness and possibly increasing their public credibility), and similarly reassuring the public that emergency authorities are aware of dangers and prepared to take measures to reduce negative impacts. Fourth, exercises provide “hands on” checks of communication systems, equipment, and

The authors
Danny M. Peterson is Associate Director of Research in the Office of Environmental Technology, Arizona State University, Mesa, Arizona, USA.
Ronald W. Perry is Professor of Public Affairs at Arizona State University, Tempe, Arizona, USA.

Keywords
Disaster management, Planning, Fire services

Abstract
There is a prevailing assumption in the research literature that disaster exercises produce a wide variety of benefits that promote effective emergency management. Unfortunately, there are few studies available that confirm this assumption. This paper reviews the role of exercises in disaster management and places them within the context of preparedness activities. Within this context, the links among planning, training and exercising are explicated. The potential benefits of exercises are reviewed and hypotheses generated that link exercise experiences with emergency responders’ perceptions of planning adequacy, training adequacy, teamwork, response network effectiveness, equipment adequacy and job risk. The effects of two exercises – one dealing with hazardous materials and one with medical mass casualties – are examined using a quasi-experimental research design. The subjects were professional firefighters. Results indicated that successful exercises can enhance perceptions of teamwork, training adequacy, response network effectiveness, job risk, and equipment adequacy. The link between exercise participation and perception of planning adequacy was found to be equivocal.

Electronic access
The current issue and full text archive of this journal is available at http://www.emerald-library.com

Disaster Prevention and Management
Volume 8 · Number 4 · 1999 · pp. 241–254
© MCB University Press · ISSN 0965-3562
other materiel. Finally, a well designed exercise tests the viability of the emergency response network relative to the threat exercised.

The issue of impacts on participants deals with a subset of the above mentioned benefits attributed to exercises. At a general level, conducting exercises involves determining if the appropriate agencies are involved in an incident response and if they are able to execute their “plan” functions – accomplish the tasks assigned them – as specified in their training. At another more micro level, exercises impinge directly on the individual responders. In this regard, exercises can be seen to affect the orientation and “mind set” of participants (responders), who execute the tasks for which they have been trained in keeping with the response plan. The emphasis in the research literature is often on exercises as an opportunity for planners, trainers and managers to observe the performance of responders. But the experience of the exercise also shapes the responder’s perceptions of the emergency management process, and the preparations given them to deal with the threat in the exercise scenario. These judgements on the part of responders seriously impact the larger emergency management system. Jurisdictional emergency managers, planners, and trainers must be seen as credible professionals in the eyes of responders to achieve any level of community protection (Drabek, 1987). Similarly, to effectively execute their jobs, emergency responders must perceive that their training has been adequate and that their equipment is appropriate to their assigned tasks. Each time emergency responders participate in an exercise, their opinions and perceptions of these and a host of other related issues are shaped.

The focus of the research presented here is to assess and document the extent that exercise participation affects the perceptions of participants. To accomplish this end, emergency responders who were involved in two disaster exercises were studied using quasi-experimental research designs. In each case, a group of responders – and a matched set of control subjects – was given a questionnaire before and after exercise participation that collected their perceptions on a range of variables representing variables (perceptions) purported to be affected by exercises. As a preface to the presentation of results, two discussions are undertaken. First, we examine the role of exercises in the emergency management process, attempting to place exercising under the rubric of preparedness and linking it with planning and training. Second, we explore in some detail the available literature on exercise impacts on participants to identify relevant perceptual dimensions (variables) that can be reasonably expected to change. Following these reviews, the data for participants in a hazardous materials accident exercise and a major medical event are examined.

**Exercises and disaster management**

At least since the late 1970s, disaster management has been conceptualized as involving four generic types of activity:

1. mitigation;
2. preparedness;
3. response; and
4. recovery (Lindell and Perry, 1992).

In this parlance, mitigation is seen as prevention – stopping a negative event before it happens. A common example is to use “breach proof” containment for hazardous substances as a means of preventing accidental environmental releases during the transportation process (Lindell and Perry, 1996). Preparedness represents actions that are undertaken to reduce the negative consequences of events where there is insufficient human control to institute mitigation measures. For example, evacuation planning is a preparedness measure frequently used to protect vulnerable populations from riverine flooding (Zelinsky and Kosinski, 1991). Response refers to actions undertaken immediately before and during impact to reduce primary and secondary negative effects. Ideally, response measures are conceived as preparedness, but some proportion of them seem to be inevitably devised spontaneously (Perry, 1991). Finally, recovery measures encompass what has traditionally been called reconstruction and recovery; ultimately the rebuilding of the disaster-impacted entity (Drabek, 1986).

Initially, these activities were often described as “stages”, implying a time order for action that began with mitigation. A more contemporary view is to acknowledge that there are no clear separations between them, and no absolute time sequence either (Lindell
and Perry, 1992). That is, each of the four activities overlap to a certain extent. A principal difference between mitigation and preparedness can be found in the degree of human understanding of the threat and the level of human control over factors that influence the threat. Otherwise, both lines of action involve taking measures that will reduce negative consequences in some fashion. Also, mitigation actions are often undertaken as part of reconstruction and recovery; a bridge may not only be rebuilt, but it may incorporate structural measures to reduce the probability of its failure in a subsequent earthquake or flood. These issues are not particularly problematic and actually do not arise except to the extent that one falls into the logical error of reifying the four-fold classification (Perry, 1998). The classification does serve as a useful heuristic for representing and generally grouping activities that constitute emergency management. Ultimately, for the purposes of the present research, the goal is only to indicate that exercises fall under the rubric of preparedness. They are designed to be rehearsals or simulations of plans that would be instituted during a response phase to deal with a threat over which there is insufficient human knowledge and control to prevent.

**Disaster preparedness**

Literally volumes have been written on disaster preparedness (Drabek, 1986), as well as several treatises that seek to define the notion itself (Gillespie and Colignon, 1993). The purpose of this discussion is not to exhaustively deal with preparedness, but to provide an overview of the relationships among three preparedness components: planning, training, and exercising. Concern here is with the narrowest sense of planning. It is assumed that broader functions—vulnerability analysis, technical analysis of structural and nonstructural options, and political and technical feasibility reviews—have been completed. A specific threat has been identified and a constellation of measures to be implemented in the response phase has been selected. At this point, planning involves the construction and codification of strategy and tactics into an executable series of actions or tasks. Plans address all aspects of the response, including personnel, equipment, contingency issues, policy issues, and inter-organizational and inter-governmental relations. The plan represents, then, a blueprint for addressing all aspects of a particular environmental threat.

Once a plan is created, its implementation revolves around the logistics and protocol needed to execute the specified series of tasks. Implementation usually begins with a detailed assessment of capacity. Agencies which provide different functions under the plan must be evaluated for their ability to comply. Personnel complements and equipment need to be brought to plan specifications. In connection with or as a product of this assessment, a variety of training needs may be identified. These include both training and education of personnel regarding the threat, the response processes and procedures, and the use of the equipment called for under the plan. In this context, training is the activity that translates information defined as needed by the plan into a coherent program that can be imparted to responders.

Broadly speaking, exercises represent constructed opportunities to test the protocols and equipment specified under a plan and taught in the training phase. Exercises may be seen as a form—perhaps an extension—of training in the sense that individuals are rehearsing response measures. Ultimately, however, exercises provide the forum or context to test the effectiveness of both the training program and the plan, as well as the ability of personnel to execute the plan. Some emergency managers argue that exercises serve as an opportunity to “validate” training and planning (Shapiro, 1995). Historically, exercises and drills are probably best classified as part of the rubric of evaluation research. Certainly the federal regulations that require annual (or periodic) exercises do so to ensure that potential shortcomings in the plan and the training process are identified. At the same time, the “training function” is also important; exercises engage responders to execute their assigned functions and in the process build “experience” in interacting with other responders (potentially from a myriad of agencies with a wide range of expertise).

The creation of exercises is an exacting process that reflects the logic that they test personnel, protocol and equipment (Emergency Management Institute, 1990). Thus, developing an exercise involves a variety of milestones designed to create or simulate an event or incident that approximates the threat addressed in the planning and training phases. A critical initial milestone is to set the
objectives for the exercise. These may be broad or narrow, depending on the scope of testing that is desired. Thus, in exercising a flood evacuation plan, one might devise an exercise built solely around testing the warning phase activity. Consequently, such an exercise might involve tasks associated with:

1. evaluation of threat data by local emergency managers;
2. the decision that evacuation was required in one or more geographic areas;
3. the construction of a warning message; and
4. the notification and mobilization of the agencies involved in disseminating the message to the public.

These considerations ultimately define the “type” of exercise as discussed below.

Once objectives have been set, one must devise an event scenario. Such a scenario must posit an event or series of events that offer an opportunity for plan specifications to be implemented. In some exercises, the creation of the scenario is particularly intricate involving not only simulation of an event, but also victims and physical damage. The management of an exercise is somewhat similar to a major stage production in that there must be realistic actors and props, with detailed stage direction (training). Often the utility of the exercise depends in part upon the participant’s (emergency responder’s) perception of its “believability”, which itself hinges upon the realism of the scenario.

Types of exercises

The process of exercising has evolved over a number of years. In some cases, a single exercise (scenario) will be devised and used multiple times; either on the same responders or across different jurisdictions. More commonly, exercises are generated with specific goals and objectives that are peculiar to a given jurisdiction’s vulnerability pattern and planning activity. Without regard to these variations, three general types or levels of exercises are identified in the literature (Daines, 1991). These types are labelled tabletop, functional, and full-scale exercises.

A tabletop exercise is the least complex of all exercises, and focuses upon a primarily verbal “walk through” of the plan. Participants in these types of exercises are usually gathered in a single location, sometimes in the same conference room. Most “tabletops” begin with a simulated event (disaster) described to participants in narrative form. Each participant (or group of participants) plays their assigned role in the relevant emergency response system. Typically, they verbally respond to event demands by describing the actions they would initiate. Contacts with other responders or agencies are usually simulated. “Controllers” carefully implement the exercise protocol and monitor the responses of participants, sometimes injecting event variations or “crises” into the exercise process to test specific exercise objectives. In these types of exercise evaluation and self-critique may be conducted after the exercise is completed or progressively through the event using “time outs” for assessment. Tabletop exercises are rarely conducted in “real time” and can be characterized as the least formal type of exercise. The types of objectives that can be adequately tested are concomitantly general; the actions involved are “action intentions” not operational execution in the field.

A functional exercise represents a higher level of complexity in testing planning and training. In keeping with their name, these exercises select one or a small number of functions under a disaster response plan and conduct a test. For example, in the context of a plan to respond to a terrorist generated explosion that dispersed sarin gas, one might conduct a functional exercise to test the Emergency Medical Services (EMS) segment – victim extraction, agent identification, patient decontamination, triage, scene treatment, and transportation to hospital treatment – of the plan. Similarly, a functional exercise could be focused on the Law Enforcement segment dealing with objectives related to scene isolation and control, evidence gathering, perpetrator identification, and investigation. Depending upon the nature of the plan tested, a functional exercise may involve a single responder agency or many. Usually, functional exercises are conducted in real time, in the field, with operational personnel executing their functions using appropriate equipment. This normally demands that the exercise staff include actors – such as victims in an EMS exercise – and that the scenario be supported with appropriate props consistent with the threat. Unlike the tabletop exercise, realism is important in functional exercises.
Finally, the most complex form is the full-scale exercise. The purpose of the full-scale exercise is to test all or a major portion of the functions specified in an emergency response plan. To accomplish tests of multiple functions by definition entails the statement of many exercise goals (objectives) and sub-goals, the participation of multiple responder agencies, and the requirement for a high level of realism in the scenario. Thus, full-scale exercises are major enterprises that demand many resources, a full staff of evaluators and controllers, a complement of actors (victims and other event-impacted personnel), and props representing the physical damage consequences of the event. Participants at all levels must literally execute their tasks under the disaster plan on the operational field in real time.

In terms of exercising a jurisdiction’s emergency management capacity, these types of exercises would be used progressively. That is, one would undertake initial plan evaluations in terms of the tabletop exercise, identifying and developing remedies for issues uncovered at that level. Work would then progress to functional exercises of major plan functions, eventually graduating to the full-scale exercise as a comprehensive test of plan viability and execution. It is important to point out, however, that the present research interest – impacts of exercise participation – will be realized at all three levels of exercise. To the extent that such impacts are positive then, there is an opportunity to create them at all levels of exercise complexity.

**Impacts of exercise participation**

We initially described several “generic” benefits that are attributed to the exercising process. At this point, we want to turn to characterizing some specific benefits of exercises that accrue to those emergency responders who are directly involved. In particular, the goal is to focus on perceptions of exercise participants that are related to preparedness matters. Thus, attention will focus on the social psychological impacts of exercise participation; an issue of changing perceptions.

It is possible to identify several specific responder impacts or perceptions that might change as a function of participation in a successful exercise. Expository papers on the effects of exercises suggest that responders involved in the event should be able to observe first hand levels of teamwork between agencies and gain confidence that collections of personnel from diverse agencies can work together in the field. Teamwork is defined in this context as the belief that individual responders can successfully “partner” or “co-work” with responders from other agencies to achieve a common goal. In particular, Kartez (1988) argues that since exercises demand teamwork across agencies, a successful outcome should increase participants’ perceptions that teamwork can be achieved. Similarly, the United States National Response Team (1990) manual on exercising identifies team building as a major consequence of exercise participation.

Along a related dimension, it has also been argued that exercise participation enhances participants’ perceptions of response network effectiveness. The notion here is that as responders simulate a disaster event they witness first-hand the support for the response provided by the predefined network of agencies with specialized skills and material. It is different from the concept of teamwork – which focuses on individual interactions – in that the issue is the ability of diverse agencies to deliver needed supplies and services to those who need them at an appropriate time. Response network effectiveness focuses on collateral support, perhaps afforded through mutual aid agreements or as a matter of responsibilities defined by jurisdictional rules or legislative mandates. In connection with the Gulf War (Operation Desert Shield), Lenorivitz (1990) reported that the logistical support network was identified as having major deficiencies; an experience which reduced participant beliefs in an adequate response network.

There is much argument that exercises serve as both an extension and reinforcement of training, and a test of planning adequacy (Hildreth, 1989; Wilson, 1989; Belardo et al., 1983). Relative to training, the exercise simulation provides a series of structured challenges for the responder. These challenges, assuming the planning-training-exercising process is correctly executed, should have been among those directly addressed either as part of the responder’s general education or as part of a specific training program. Responders in the field face the challenges and make judgements...
regarding the adequacy of their training in terms of how effectively the techniques they have been taught address the challenge. Thus, in a successful exercise, participants’ perceptions of the adequacy of training should be increased.

While training typically addresses the techniques responders employ in the field to solve problems, planning is the process that identifies the kinds of problems which are likely to be confronted. Responders experience the adequacy of planning in terms of the extent to which the problems encountered were the same as those at which training (under the plan) was directed. When problems arise that were “unanticipated”, it serves as evidence that the planning process was less than ideally effective. In a successful exercise, responders will encounter largely problems that they were warned about and trained for, consequently increasing their belief in planning adequacy.

Functional and full scale exercises involve testing response functions in “full dress”, such that any equipment associated with the threat response is present. Consequently, participants experience several dimensions of equipment adequacy. As conceived here, this is not a question of learning to use equipment, but one of participant assessment of:

1. the extent to which equipment performs as intended; and
2. the extent to which it is possible for a trained responder to actually use the equipment under response circumstances.

Assuming technically correctly functioning equipment and that adequate planning identified appropriate equipment choices for the threat environment, participation in an exercise should increase responder perceptions of equipment adequacy.

Finally, over the years there has been much concern on the part of emergency managers with levels of stress associated with disaster response and consequences of this stress for responders themselves (Lystad, 1988). Efforts to deal with such stresses “after the fact” have been developed in the form of Critical Incident Stress Debriefing (CISD), where responders are counselled and provided psychological support either individually or in groups. Taylor (1990) has argued that providing employees with experiences consistent with job demands reduces their perceptions of danger associated with job execution. By extension, one can argue that an exercise – as a realistic simulation of the demands associated with disaster response – constitutes an experience of the threats responders are likely to experience in a genuine response environment. Thus, assuming that an exercise is successful, participants should appreciate that the combination of planning, training and personal ability they possess enables them to survive and perform in a threat setting. Consequently, one would expect that exercise participants would perceive that the risks (dangers) associated with the response task are lower (that is, that they are manageable).

If this reasoning is correct, one should also acknowledge that exercises form a path to stress reduction for emergency responders. This logic assumes that personal stress is reduced when individuals believe that their level of direct control over their own welfare is high.

The preceding discussion has identified six areas of perception that should be affected as a function of participating in an exercise. Assuming that the exercise is successful, these effects can be stated as a series of expectations:

1. exercise participants should perceive higher levels of teamwork;
2. exercise participants should perceive higher levels of response network effectiveness;
3. exercise participants should perceive higher levels of training adequacy;
4. exercise participants should perceive higher levels of planning adequacy;
5. exercise participants should perceive higher levels of equipment adequacy;
6. exercise participants should perceive lower levels of job risk.

Of course, the present research is only concerned that these six perceptual dimensions are affected by exercise participation. In an unsuccessful exercise, one would expect that participant confidence levels in each dimension would decline. We have elected to test each expectation in the context of the two exercises that are described below. The exercises were conducted serially, with several months between them, enabling the measurement scheme to be adjusted for the second (major medical) exercise based on experiences in the hazardous materials exercise. Each exercise was conducted as a functional exercise whose primary purpose
was to test functions in an emergency response plan. The participant satisfaction component dealt with exclusively here was a small part each overall exercise activity. In the sections that follow, each exercise is described, measurement schemes are reported, and the results provided. In the closing section, we attempt to integrate and interpret the findings from each exercise.

Hazardous materials incident exercise

Methodology
To examine participant reports of the impacts of an exercise built around a hazardous materials incident, a simple quasi-experimental research design was employed. Arrangements were made with a private industrial firm to collect data in connection with their mandatory annual United States Occupational Safety and Health Administration training (required under OSHA regulation 1910.120[q]). An exercise scenario was developed in conjunction with the organization’s management which involved two classes of emergency responders: professional firefighters from jurisdictions that would respond to accidents at the plant, and specially trained company employees who serve on the company emergency response team. Only the professional emergency responders – the fire department personnel – were included as subjects in the analyses reported here. The 25 professional firefighter responders were defined as the experimental (or treatment) group who would participate in the exercise. A control group was formed by recruiting 18 comparably trained firefighters from the same departments but working shifts not affected by the exercise.

As part of the before-after comparison group design (Campbell and Stanley, 1979), all members of the experimental and control groups were asked to complete a questionnaire that included measures of the selected perceptions before the exercise. For both groups, this measurement took place one week before the event. After the exercise, each participant completed a questionnaire with the same items. At about the same time, members of the control group, who were not exposed to the exercise, were asked to complete a second questionnaire. This research design enables us to examine initial levels of each impact perception for experimental and control groups and to record the levels of the same perceptions after the exercise. The before-after score (called gain score) differences were calculated for each group. Thus, the gain score indicates how much change occurred for each perceptual dimension for the treatment (experimental) group, to enable comparison with the control group. The statistical reliability of differences in gain scores was examined using simple (one way) analysis of variance (ANOVA).

Six variables, representing the areas of perception identified above, were measured using single questions on a self-administered questionnaire. The operational definitions used for each variable are shown in Table I. Thus each variable was measured as the emergency responder’s degree of agreement with a simple, declarative statement regarding the reference perception. A Likert-type response format was developed and used for each statement (Edwards, 1957). The response format was composed of six assessments of the “truth” of the statement and an assigned score:

1. absolutely true;
2. largely true;
3. moderately true;
4. sometimes true;
5. rarely true; and
6. never true.

Thus, emergency responder ratings for each perceptual area varied from 1 through 6, with smaller numbers representing higher levels of each variable.

Exercise overview
The scenario for the exercise placed a hazardous materials accident at a local manufacturing facility. A collision was simulated

<table>
<thead>
<tr>
<th>Table I</th>
<th>Measures for hazardous materials exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Statement</td>
</tr>
<tr>
<td>Teamwork</td>
<td>There is a high level of teamwork in my area(^a)</td>
</tr>
<tr>
<td>Response network</td>
<td>Our response network is highly effective</td>
</tr>
<tr>
<td>Training adequacy</td>
<td>Our training program is completely adequate</td>
</tr>
<tr>
<td>Planning adequacy</td>
<td>Adequate planning for Hazmat incidents or accidents has been accomplished in my department</td>
</tr>
<tr>
<td>Equipment adequacy</td>
<td>Our Hazmat team is adequately equipped</td>
</tr>
<tr>
<td>Job risk</td>
<td>There is a high level of risk in my job</td>
</tr>
</tbody>
</table>

Note: "The term "area" in this context is a colloquial reference to the workplace that includes the individuals who compose the task group"
between a large tank truck carrying sulfuric acid and a smaller delivery truck also carrying hazardous chemicals. A breach of containment was simulated that produced a slow leak from the sulfuric acid truck, and similar leaks were simulated from the smaller vehicle. Both drivers and one passenger were moulaged (that is, had injury-simulating make up applied) and assigned to behave as if they were overcome by vapor. The exercise participants included members of the hazardous materials (Hazmat) Response Team maintained by the company, but the focal group for this study was the professional firefighters – Hazmat teams – who were dispatched from three nearby fire departments. The emergency response task required that the hazardous materials be contained, the leaks stopped, and the victims be extracted, decontaminated, treated at the scene and transported to a local hospital. The exercise required about one hour to complete.

Results
Table II shows the analysis results for the firefighter ratings of each of the six perceptual dimensions. Each of the perceptual dimensions is discussed in turn. The experimental design shows that participation in the exercise did not have a statistically significant impact on perceptions of teamwork \( (F = 0.09, p > 0.05) \). Indeed, the gain scores indicate that both experimental and control group mean scores changed by less than one-tenth of a scale point. This outcome may be a function of the type of subjects who participated: professional firefighters who responded with the crew members who formed their daily work team. These individuals routinely work together and both the control and experimental groups showed very high levels of perceived teamwork on the pretest. In fact, at 1.89 (controls) and 1.88 (experimental), the mean ratings of teamwork before the exercise approached the upper anchor point of the scale (1.0). Post exercise measures changed little, with the experimental score increasing slightly and the control score decreasing slightly. The conclusion of no measurable exercise impact on perceptions of teamwork is best qualified to refer to audiences of professional firefighters at this point. It is possible that exercises may impact teamwork perceptions among participants who are not professional emergency responders, or who do not routinely work with one another in teams on a daily basis.

Global perceptions of the adequacy of the response network were also only slightly impacted by participation in the exercise. The control group mean score was 2.39 at the pretest, and remained stable at 2.33 for the post-test, creating a gain score of -0.06. The experimental group began with a mean rating on this dimension of 2.64, and following the exercise, response network adequacy measured at 2.75. This change represents a slight decrease in level of perceived network adequacy, which is not statistically significant \( (F = 0.04, p > 0.05) \). Two issues arise in connection with this finding. First, it is probable that the statement used to measure

### Table II  Results for hazardous materials exercise

<table>
<thead>
<tr>
<th>Perceptual dimension</th>
<th>Control pretest meana</th>
<th>Control post-test mean</th>
<th>Treatment pretest mean</th>
<th>Treatment post-test mean</th>
<th>Control gain score</th>
<th>Treatment gain score</th>
<th>F ratiob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teamwork</td>
<td>1.89</td>
<td>1.83</td>
<td>1.88</td>
<td>1.92</td>
<td>-0.06</td>
<td>0.04</td>
<td>0.09</td>
</tr>
<tr>
<td>Response network</td>
<td>2.39</td>
<td>2.33</td>
<td>2.64</td>
<td>2.75</td>
<td>-0.06</td>
<td>0.11</td>
<td>0.04</td>
</tr>
<tr>
<td>Training adequacy</td>
<td>2.44</td>
<td>2.50</td>
<td>3.04</td>
<td>2.84</td>
<td>0.06</td>
<td>-0.20</td>
<td>0.95</td>
</tr>
<tr>
<td>Planning adequacy</td>
<td>2.24</td>
<td>2.50</td>
<td>2.76</td>
<td>2.96</td>
<td>0.26</td>
<td>0.20</td>
<td>0.33</td>
</tr>
<tr>
<td>Equipment adequacy</td>
<td>2.61</td>
<td>2.83</td>
<td>3.08</td>
<td>2.72</td>
<td>0.22</td>
<td>-0.36</td>
<td>4.56c</td>
</tr>
<tr>
<td>Job risk</td>
<td>1.50</td>
<td>1.61</td>
<td>2.09</td>
<td>2.42</td>
<td>0.11</td>
<td>0.33</td>
<td>4.56c</td>
</tr>
</tbody>
</table>

Notes: aAll means are based on a scale where 1 represents the highest level of confidence and 6 represents the lowest; bOne way analysis of variance for gain score by membership in treatment or control group; between groups degrees of freedom = 1, within groups degrees of freedom = 41; cF ratio statistically significant at \( p \leqslant 0.05 \).
Response network adequacy was too broad and ambiguous, thereby not creating a clear referent in the minds of firefighters relative to which a change in perception could be measured. Second, it may be that these responders tended to assume that network adequacy exists: the before and after means for treatment and control groups are near the top of the measurement scale.

Perceptions of the adequacy of training registered more substantial change among exercise participants. Both experimental and control groups rated training adequacy at about the mid-point of the measurement scale in the pretest (controls = 2.44 and experimental = 3.04). The control group remained at a similar level when post-tested (2.50), but the experimental group changed to 2.84, registering a more positive assessment of training adequacy after exercise participation. Although the gain score for the experimental group is much larger than that for the control group, the difference is not large enough to be statistically reliable ($F = 0.95, p > 0.05$).

With respect to perceptions of planning adequacy, only a small difference was observed between the gain scores of the experimental and control groups. The experimental group showed a net gain of 0.20, compared with the control group’s net gain of 0.26. This difference is not statistically significant ($F = 0.02, p > 0.05$). In the context of the exercise studied here, both groups rated planning adequacy on the higher end of the measurement scale at the pretest. It is possible that this rating represents a threshold effect similar to that seen with perceptions of teamwork. That is, the degree of planning adequacy was initially seen as high, and the nature of the exercise involved straightforward application of standard operating procedures.

In exercises where the plan addresses events that are more unique (for example, plans for responding to terrorist threats involving chemical, nuclear or biological agents), it is possible that the perception of planning adequacy would be more influenced by exercise participation. In general, one may speculate that in dealing with familiar threats, responders visualize standard operating procedures as guiding operations. In addition, the experience of responding to events which they routinely face in the workplace creates the psychological sense that “experience” drives their actions more than what they often view as more abstract “plans”. It is when responders are faced with unfamiliar threats that they must eschew their experience and standard procedures to learn – and use – techniques specified in a plan. It is in these situations that one would expect that what takes place on the exercise field would be perceived by responders as directly linked to planning adequacy. This conclusion argues strongly for the critical importance of plans (in the minds of responders) in the exercise process for threats that are “new” to a given jurisdiction.

Perceptions of equipment adequacy changed substantially as a function of exercise participation. The control group pretest score registered 2.61, with a post-test value of 2.83, yielding a gain score of 0.22, indicating a slight decrease in perceived adequacy. Conversely, the experimental group pretest value was 3.08, and changed to a more positive perception value of 2.72 at post-test, a difference of -0.36. This difference was statistically significant ($F = 2.75, p < 0.05$).

Finally, perceptions of risk associated with the job also differentially changed as a function of participation in the exercise. Both groups rated the job as less dangerous after the exercise, but the increase for the control group was smaller (0.11) than the increase for the experimental group (0.33). This difference is statistically significant ($F = 4.56, p < 0.05$). This change in perceptions was in the direction predicted by the general literature on exercising.

To summarize, the results of the first experiment indicate that participation in the exercise had the largest impacts on perceptions of job risk, equipment adequacy and training adequacy. Only small impacts were registered for perceptions of teamwork, response network effectiveness and planning adequacy.

**Medical mass casualty exercise**

**Methodology**
The same basic quasi-experimental design was used to examine participant impacts for the medical mass casualty exercise. The exercise used here represented one series of exercises that was conducted as part of a larger effort to establish and implement a regional system for medical triage by firefighter paramedics. The nature and purpose
of triage is to sort individuals with medical injuries in accord with their condition to determine priority. Professional firefighter-paramedics from 11 regional fire departments were initially given training on the principles by which triage would be conducted and with the triage equipment. In this case, equipment refers to a small fabric “holster” (triage kit carried on a belt) containing designated triage tags, patient attachments, and larger adhesive stickers to locate the highest priority patients. A total of 36 firefighters who participated in the triage exercise were designated as the “experimental” group. A group of 32 comparably trained firefighters from the same departments who did not participate in the exercise formed the control group. At the time of triage training, approximately three to five days prior to exercise participation, firefighters in the experimental group completed a short questionnaire that included the selected perceptual measures. A second questionnaire was administered to this group immediately after completing the exercise. The control group members were given an initial questionnaire, followed by the same questions (with order of presentation randomized) approximately one week later.

Both research expectations and measurement technique were altered slightly as a function of the results obtained from the hazardous materials exercise. In all cases, the measurement changes were designed to make the focus of questions more specific than the global questions used in the first experiment. Two perceptual areas were dropped, on substantive grounds, from this analysis. One deleted item referred to job risk. While this is intuitively an important dimension of hazardous materials response, professional firefighters conducting medical operations tend not to perceive this part of the job as more risk prone than other aspects of the job. Planning adequacy in the case of mass casualty medical response is embodied in the fire services Incident Management System (IMS), which underlies all other types of fire operations as well. Thus, references to planning adequacy in this context would produce ambiguous responses in that no precise incident-unique plan is involved for evaluation.

The remaining four perceptual areas were measured using different statements in the medical mass casualty exercise. In conducting the hazardous materials experiment, our measurement was guided by the available research literature which principally addressed general dimensions of the impact of exercise participation. The analysis of those data, and post-exercise debriefings with participants, suggested that to more precisely measure exercise impacts would require specific rather than global questions. Thus, the measurement strategy adopted for the mass casualty exercise focused directly on the medical context of the activity and directly on specific procedures and outcomes. The revised questions for measuring teamwork, response network adequacy, training adequacy, and equipment adequacy are presented in Table III. The response choices given participants for these statements were the same as those used in the hazardous materials exercise, ranging from 1 through 6, where low numbers affirm the correctness of the statement.

**Exercise overview**

The data reported here were part of a much larger series of exercises designed to compare the field effectiveness of different triage systems as a step in establishing a regional triage system. The exercise analyzed as part of the participant impact perception study involved the use of a single triage system. The exercise scenario involved fire department medical crews responding to an explosion inside a multi-storied city building that generated 20 injured civilians. The victims were moulaged to represent different types and severity of injury, and were located both inside and outside the building. Within the

<table>
<thead>
<tr>
<th>Variable</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teamwork</td>
<td>In mass casualty medical incidents, my crew works effectively with crews from other departments and from other stations in my department</td>
</tr>
<tr>
<td>Response</td>
<td>In a mass casualty medical incident, I am completely confident that transportation sector, extraction sector, police, and scene support will effectively handle their responsibilities</td>
</tr>
<tr>
<td>Network</td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>Our training in mass casualty triage and management prepares me to operate effectively in the field</td>
</tr>
<tr>
<td>Adequacy</td>
<td>My triage kit constitutes an effective piece of equipment for accomplishing field triage in mass casualty incidents</td>
</tr>
</tbody>
</table>

*Note:* For the departments studied, a “crew” is composed of one officer, one fire engineer and two firefighters who train and respond together to fire, medical and hazardous materials incidents.
local fire department protocol, the response to the incident is assigned the status of “2 and 1 medical”, which requires the dispatch of two engine (pumper) crews and one ladder crew. The overall exercise was structured such that all apparatus approached from an assigned direction, and that engine and ladder crews from different fire departments were mixed in the same “2 and 1” call. Upon arrival at the scene and after police had ensured control of the scene, the Incident Commander made crew assignments. In overview, the exercise required that fire crews initiate standardized scene stabilization, locate victims, request extraction (from rubble) of any trapped victims, triage victims, administer any needed scene treatment, and transfer patients for transportation to hospital care. The police sector maintained security and traffic direction throughout the incident. The time to complete the exercise varied between different “2 and 1” calls (a total of nine crews participated in three staged incidents), but averaged 32.8 minutes.

**Results**
The results from the mass casualty medical exercise are presented in Table IV. On the teamwork dimension, the experimental group \( (x = 2.44) \) and the control group \( (x = 2.53) \) pretested at about the center of the scale, indicating “moderate” confidence that crews from different organizations could effectively work together. After the exercise, the control group mean remained similar (2.81), registering a slight drop in average confidence. The experimental group mean (1.19), however, showed a dramatic increase in confidence after exercise participation. An analysis of variance of the gain scores (differences) for the two groups yielded an \( F \) ratio of 112.9, which is statistically significant at an alpha level of 0.05. Thus, for these firefighters the experience of working through a field exercise resulted in much higher levels of confidence in teamwork ability across different crews. This finding is both consistent with the claims in the literature on exercising and with the results of the hazardous materials exercise.

On the second perceptual dimension, response network effectiveness, similar results are observed. The treatment group \( (x = 3.13) \) and control group \( (x = 3.23) \) began with similar levels of confidence – again located at the “moderate” level of the scale – that supporting response organizations could effectively handle their responsibilities. The post-test mean for the control group (3.53) remained similar to the pretest mean, showing a slight decline in confidence (0.30 scale points). The treatment post-test mean showed a much larger increase in confidence, rising to the level of 1.30 (indicating a belief level between “absolutely” and “largely” true). An analysis of variance of the difference in gain scores produced a statistically significant \( F \) ratio of 182.0 \( (p < 0.05) \).

Initially, training adequacy was rated somewhat lower on the six-point scale than either perceived teamwork or response network effectiveness. These lower initial ratings may be a function of the fact that all crews had little experience and only a single training session for the triage system. Furthermore, none of the fire crews involved in the study had used a systematic triage system prior to the training; for the most part

<table>
<thead>
<tr>
<th>Perceptual dimension</th>
<th>Control pretest mean</th>
<th>Control post-test mean</th>
<th>Treatment pretest mean</th>
<th>Treatment post-test mean</th>
<th>Control gain score</th>
<th>Treatment gain score</th>
<th>( F ) ratio(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teamwork</td>
<td>2.53</td>
<td>2.81</td>
<td>2.44</td>
<td>1.19</td>
<td>0.28</td>
<td>-1.25</td>
<td>112.9(^c)</td>
</tr>
<tr>
<td>Response network</td>
<td>3.23</td>
<td>3.53</td>
<td>3.13</td>
<td>1.30</td>
<td>0.30</td>
<td>-1.83</td>
<td>182.0(^c)</td>
</tr>
<tr>
<td>Training adequacy</td>
<td>3.59</td>
<td>3.37</td>
<td>3.75</td>
<td>1.45</td>
<td>-0.21</td>
<td>-2.3</td>
<td>151.9(^c)</td>
</tr>
<tr>
<td>Equipment adequacy</td>
<td>4.06</td>
<td>3.71</td>
<td>3.94</td>
<td>1.44</td>
<td>-0.21</td>
<td>-2.5</td>
<td>210.7(^c)</td>
</tr>
</tbody>
</table>

*Notes:* \(^a\)All means are based on a scale where 1 represents the highest level of confidence and 6 represents the lowest; \(^b\)One-way analysis of variance for gain score by membership in treatment or control group; between groups degrees of freedom = 1, within groups degrees of freedom = 66; \(^c\)\( F \) ratio statistically significant at \( p \leq 0.05 \).
in all participating departments crews triaged verbally on somewhat idiosyncratic factors determined by individual paramedics. Consequently, it is not surprising that a single training session inspired ratings of only “sometimes” confident in training adequacy (control group mean = 3.59; experimental group mean = 3.75). Participation in the exercises strongly impacted perceptions of training adequacy. While the control group post-test mean remained stable at 3.37 (representing a slight increase in confidence over the pretest), the experimental group post-test mean changed to 1.44 (representing a substantial increase in confidence). An analysis of the gain scores yielded an $F$ ratio of 151.9, statistically significant at an alpha level of 0.05. Clearly, participation in the exercise provided convincing evidence to the crew members that their triage training covered the contingencies that they would face in a field incident.

Finally, initial perceptions of equipment adequacy were also low among both the control group members ($x = 4.06$) and the experimental group members ($x = 3.94$). Again, it is likely that such low initial ratings stemmed from the fact that none of the crews had used “triage kits” in the past and the kits constituted “untried” equipment to the firefighters. The experience of using the equipment in the field again seems to have radically altered perceptions. The level of confidence in equipment expressed by the control group ($x = 3.71$) remained stable in the post-test, although it does represent an increase in perceived confidence. The experimental group mean for the post-test changed to 1.44, however, representing another dramatic increase in the level of perceived confidence in equipment adequacy. The $F$ ratio for the differences in gain scores ($F = 210.7$) is statistically significant ($p < 0.05$), documenting that the observed increase in confidence in training adequacy is a reliable difference.

**Findings and implications**

The results of the present study provide empirical evidence that disaster exercises have impacts on the perceptions of the emergency responders who participate. It is important to emphasize, however, that all participants in this research were not only professional emergency responders, but trained fire fighters. Larger or smaller magnitude effects, or different effects altogether, might be detected in different populations, especially those composed of people who are not professional emergency personnel. In general, it was found that exercises have the ability to change participant perceptions of teamwork, response network effectiveness, training adequacy, equipment adequacy, and job risk. Furthermore, the use of two disaster exercises permits identification of some qualifications or conditions associated with these findings.

The results from the hazardous materials exercise show that exercise participation caused responders to report a decrease in the perceived level of danger or risk associated with their job. The measured effect was substantial and statistically reliable. Since Hazmat work is widely accepted to be rife with real danger, it can be inferred that the combination of training and experience obtained in the course of exercise participation combined to convince responders that careful adherence to protocol can reduce levels of risk. It should be emphasized here that this result would be expected only as a consequence of a “successful” exercise; one in which the procedures and protocols produced positive outcomes. One should remember that in exercises where it is demonstrated that protective measures and protocols are inadequate to the task, participant perceptions are likely to change in a way that reflects reduced confidence. This possibility underscores the need for careful planning and execution of exercises; an event done poorly carries as many negative effects as a well executed event carries positive effects. Furthermore, one would not expect that this outcome (reduced perception of danger) would generalize to exercise participation where the threat managed was not acutely and obviously hazardous.

The hazardous materials exercise did not produce significant effects on participant’s perceptions of the adequacy of planning. This outcome may have been an artifact of the subjects studied and the nature of the problem, and much caution should be observed in drawing conclusions. Claims in the literature that exercises should test planning adequacy and consequently that participant perceptions of planning adequacy likewise should be affected depend upon two factors. First, it is assumed that the functions tested by the exercise are themselves a product of
specific plan elements. Second, and more critically, it assumes that participants are sufficiently aware of the plan and its contents to connect it to their training and connect their training to the exercise scenario. First responder expertise tends to be a product of the training process and of past experience. The job demands of such positions and the types of personalities drawn to them mitigate against any consuming concern with a plan; instead the focus remains on the task, the training, and incidents. In this context, the conclusion that firefighters participating in an exercise did not perceive plans differently appears more reasonable; especially since first responders tend to deal with the specifics of response rather than develop an awareness of more general plans. It is possible, however, that emergency managers working at a more general level – perhaps in the jurisdictional emergency operations center – would perceive their work to be more closely plan-connected and consequently register changes in perceptions of plan adequacy following exercises.

The perception of teamwork was measured in two different fashions. A more global assessment of teamwork was used in the hazardous materials exercise. It was found that both experimental and control subjects initially perceived teamwork to be high and continued to see it as high following the exercise. It is likely that in the hazardous materials exercise teamwork was defined by participants as referring to coordinated action among crew members. This is an appropriate and expected perception since the crews studied here routinely train and work together on a daily basis. For these exercise participants the impact of exercise participation on perception of teamwork was very slight. Conversely, in the medical mass casualty exercise, this measurement was refined to refer to perceived teamwork or cooperation between crews from different parts of the same organization or from different organizations. In this case, a successful exercise produced a large and statistically significant positive impact on perceived levels of teamwork. Thus, in a more general sense, one can conclude that to the extent that exercises bring together people who do not routinely work in concert, a successful exercise can enhance perceptions of teamwork.

Data from both of the exercise experiences examined here show that exercise participation has effects upon perceptions of response network effectiveness. The magnitude of the effect was smallest in the hazardous materials exercise, where measurement was most general and allowed the participant to subjectively decide what constituted the “response network”. In the medical mass casualty exercise, response network elements were defined for the participant – or made explicit in the question – so that it was clear that the object was to assess functions performed by different agencies. With this clarification, subjects showed a large statistically reliable increase in perception of response network effectiveness. It can be inferred that the experience of watching successful operations mounted by outside agencies, and successful interface of agency personnel, increases confidence that the response network is appropriately configured and capable of delivering the support required to sustain the constellation of responder agencies.

Another perceptual dimension which was impacted in both exercises was training adequacy. In the first exercise, training adequacy was operationalized by a general reference to hazardous materials incident training. That is, neither a particular training experience nor specific training content was referenced in the question asked of participants. In this case, a small increase in perceived adequacy was registered; despite the ambiguity of the reference, subjects connected exercise performance with their training experiences. For the medical mass casualty exercise, the reference to training was made much more specific. Thus, participants were asked about a particular triage training experience that had been recently conducted. These participants showed a substantial statistically significant increase in perception of training adequacy. In this context, it appears that the closer the training is to the successful exercise scenario, the more precisely the training content is tied to the exercise experience, and the more precisely the training content is tied to the exercise outcome. Of course, one would expect that either an unsuccessful exercise, or poor training, would produce a decrease in the perception of training adequacy.

The greatest magnitude of change in perceptions took place in connection with equipment adequacy. In both successful exercises, perceptions of the adequacy of equipment increased dramatically (and statistically significantly) as a function of exercise
participation. This effect was observed in the hazardous materials exercise where the referent of the term equipment was general and included a range of devices (from respiratory protection to protective clothing). The magnitude of change was even larger in the medical mass casualty exercise where the referent for equipment was more narrow and specifically indicated triage kit. These outcomes strongly support the contention that the performance of equipment during exercises is strongly connected with participant’s perceptions of the adequacy of equipment for emergency response.

In closing, in addition to specific findings about the impact of exercises on the perceptions of those who participate in them, some general conclusions may also be offered. The present research underscores the importance of the links between planning, training and exercising to achieve disaster preparedness. Each component is critically connected to the others such that a failure in a single link can hamper the achievement of jurisdictional preparedness. Furthermore, exercises produce measurable effects upon the responder’s perceptions of both the adequacy of the strategy for addressing hazards and the tactics for managing response. Seen in this light, exercises are both an important part of emergency management and an avenue to continuous quality improvement.

References


Peterson, D.M. (1996), The Effects of Participating in a Hazmat Exercise on Key Perceptual Variables, Arizona State University School of Public Affairs, Tempe, AZ.