A Multidimensional Evaluation of Fire Fighter Training for Hazardous Materials Response: First Results from the IAFF Program

Alexander Cohen, PhD*

The International Association of Fire Fighters (IAFF) course on hazardous materials training for first responders is described together with an evaluation plan that includes multiple levels of assessment. Trainee appraisals of the course, shifts in their ratings of task competencies, gains in knowledge quiz scores, and self-reports on actions reflecting lessons learned from the course are among the measures used. Evaluations of courses given in several city fire departments found more than 60% of trainee judgments of course quality and utility to be highly favorable, along with significant post-course improvements in their competency ratings and quiz scores. Follow-up interviews with samples of trainees also suggested more self-protective behaviors and preventive actions being taken with regard to alarms and risks of hazardous materials exposures. However, cross-comparing the results for the various evaluation measures gave only limited support to a popular evaluation model that hypothesized that they would be interdependent. Limitations in appreciating technical course subjects, the value of add-on or refresher instruction, and variable risk experiences are noted in explaining differences in some training results. Am. J. Ind. Med. 34:331–341, 1998. *© 1998 Wiley-Liss, Inc.*

***KEY WORDS:* occupational safety training; fire fighters; hazardous materials exposures**

INTRODUCTION

As first responders to common alarms, fire fighters face an increased risk of being exposed to hazardous materials that may be generated directly by a fire, stored at the fire site, or spilled as a result of accidents during transit. This risk continues to grow with the proliferation of chemical substances and other harmful agents found in our workplaces, communities and homes. Traditional fire fighting practices which stress aggressive actions in responding to fires and rescues may be inappropriate in the presence of these kinds of harmful agents. Indeed, fire fighter units that are first on the scene may find themselves ill-equipped and without knowledge of the special techniques needed to deal effectively with incidents involving hazardous materials. As a consequence, their own safety as well as that of others may be jeopardized. Events that underscore these concerns have been reported. Table I summarizes some of the more dramatic incidents. But countless others exist that lack such notoriety. For example, dumpster fires, residential fires, especially in areas where household chemicals may be stored, similarly present risks for exposures to harmful substances. In light of these concerns, training in responding to hazardous materials incidents has become a priority for fire fighters as well as other emergency responders.

This article describes an evaluation of a training course intended to educate fire fighters about procedures to safeguard their health and safety when emergency calls involve potential exposures to hazardous materials. Its purposes are threefold: first, to describe the course and an evaluation plan designed to yield multiple measures of the course's effect on trainees; second, to summarize data collected on the various measures from courses already given to several city fire departments; and third, related to the second purpose, to compare the evaluation data with expectations based on a popular model of training evaluation criteria and outcomes.

Cincinnati, Ohio Contract grant sponsor: National Institute for Occupational Safety and Health.
*Correspondence to: Alexander Cohen, 6752 East Farmacres Drive, Cincinnati. OH 45237
Accepted 4 April 1998

TABLE I. Select Incidents Where Fire Fighters Were Killed or Injured from Exposures to Hazardous Materials

Fire fighters in Kingman, Arizona, responded to a fire on a 30,000-gallon tank care that was unloading liquid petroleum gas at a storage facility. Though there was no threat to life in the immediate vicinity of the burning tank car, the fire fighters on the scene took up hose positions to apply water on the burning tank car. The volume of water, in impinging on the vapor space of the burning tank car, caused a liquid expanding vapor explosion resulting in the death of 12 fire fighters.

Fire fighters in Auburn, Indiana, were summoned to a metal plating facility where several plant workers had been overcome by hydrogen cyanide fumes while cleaning a metal plating tank without any respiratory protection. The fire fighters, wearing structural fire clothing and self-contained breathing apparatus, removed the victims. The hydrogen cyanide vapor produced in the tank permeated the exposed skin of the fire fighters and contaminated their protective clothing and thus placed them at added risk to this chemical hazard.

Fire fighters from Frederick County, Virginia responded to a road accident involving an overturned tractor trailer. A white granular powder was spilling from the damaged containers in the area around the trailer. Fire department members wore structural fire fighting clothing, three-quarter length boots and demand-type breathing apparatus. Without knowing about the trailer contents, the fighters removed the damaged drums and righted those that were intact. In doing so, they began to complain of a burning sensation around their wrists and ankles. Upon removing their gloves and boots, several fire fighters were found to have first and second degree burns caused by exposure to sodium hydroxide.

Los Angeles, California fire fighters were called to a fire at a warehouse owned by a research corporation. Numerous flammable, corrosive and reactive compounds were stored in the building and drums of chemicals exploded during the blaze. A cloud of toxic smoke developed and spread downwind more than one-third of a mile. Fire fighters as well as police summoned to the scene were subject to both inhalation and direct skin exposure hazards from chemicals present in runoff waters. Three police and six fire fighters were hospitalized for apparent ill effects (irritability, headaches, nose bleeds, nausea and memory loss) from smoke inhalation and toxic chemical exposures related to the incident. Some health problems persisted for more than one year after the fire.

Fire fighters in Genesee Falls, New York, responded to a fire in a barn containing a number of farm chemicals. Forty-one people, mostly fire fighters, were sent to the hospital after breathing contaminated smoke and fumes. Fifty-nine students at a local school were also exposed, and suffered headaches and stomach problems.

ESSENCE OF THE TRAINING COURSE

The course ‘‘Hazardous Materials Training for First Responders’’ was developed by the International Association of Fire Fighters (IAFF) Hazardous Materials Training Department. The IAFF is one of several groups offering such training to emergency responders and has done so in two ways. One is by dispensing packages of course manuals and materials to persons already serving as instructors in the field. The second is by providing direct onsite training to fire fighters and others who have emergency responder roles in local fire departments. In this case, specially trained IAFF instructors present the course. This evaluation report deals with first results from the latter type of IAFF training.

Consistent with objectives set forth by the Occupational Safety and Health Administration [1989] and the National Fire Protection Association [1992], the IAFF hazardous materials course for first responders emphasizes hazard awareness, self-protective and preventive measures, and defensive roles. Actions to neutralize or clean up a hazardous materials event are left to HAZMAT teams with more advanced training and equipment. The IAFF course covers the following four units of subject matter:

Unit 1: Common alarms. Explains the need for recognizing material hazards in common alarms; notes material hazards likely to be found at select locations and industry sites

Unit 2: Health and safety. Describes toxic substances, routes of entry into the body, health effects and the importance of medical surveillance

Unit 3: Hazardous materials. Depicts placards, labels, cargo container shapes for identifying hazardous materials; demonstrates how chemical properties can affect reactivity

Unit 4: The planned response. Explains limits of structural fire fighting gear, use of self-contained breathing apparatus (SCBA), other personal protective equipment (PPE), and decontamination methods; defines zones in managing hazardous materials incidents, and pre-incident surveys to identify high-risk sites

The course is designed to be taught in 24-h class hours, usually arranged as three 8-h class days, and led by two IAFF instructors. Class sizes cannot exceed 30 trainees, and thus multiple classes have to be formed in departments to cover the numbers normally scheduled to attend. Given limited means to provide backup to those attending classes, the 3 days of training for any one class may have to be stretched over weeks as opposed to being taken on successive days. Completion of the three training days for classes in some departments supplying data for this evaluation report took as much as 6–8 weeks (Table II). This extended time period can also necessitate using other than the same two IAFF instructors to cover classes.

Details on IAFF trainer qualifications, teaching techniques, course materials, are available from the IAFF (address: 1750 New York Avenue, Washington, DC 20006). Requests for IAFF training courses should be addressed to their offices.

TABLE II. Current Database for IAFF First-Responder Course Evaluation*

State/city (time span for 3 trng days)	Trainee reaction	Competency ratings	Pre/post quiz	Behavioral change	Organization change
Illinois (IL) City (4–6 wk)		X (n = 134)		Y (n = 33) (actions followed)	
Florida (FL) City (3 wk)	X (n = 80)	X (n = 80)	X (n = 88)	Y (n = 30) (actions followed)	
W. Virginia (WVA1) City (3 wk)	X (n = 150)	X (n = 150)	X (n = 131) Y (n = 28) (retest)	Y (n = 28) (lessons noted)	Y (n = 28) (actions reported)
W. Virginia (WVA2) (4 wk)	X (n = 100)	X (n = 83)	X (n = 80) Y (n = 18) (retest)	Y (n = 18) (lessons noted)	Y (n = 18) (actions reported)
Alabama (AL) City (3–4 wk)	X (n = 109)	X (n = 109)	X (n = 90)	Y (n = 27) (actions followed)	
Texas (TX) City (6–8 wk)			X (n = 453)	Y (n = 35) (actions followed)	
			Y (n = 33) (retest)	Y (n = 33) (lessons noted)	Y (n = 33) (actions reported)

*X, Entry reflects actual number of trainees present in classes supplying data on designated measure. n values can vary for measures obtained within same department owing to incomplete responses on some forms, missing data sheets for trainees being called out of class to respond to alarms. Y, Entry is based on a random sample of trainees drawn from total class roster and collected 6–12 months after the course ended. Descriptors in parentheses indicate nature of follow-up data obtained.

THE EVALUATION PLAN

The latest recommendations for evaluating health education and safety training programs suggest using combination techniques and measures to corroborate results as well as obtain information about the processes that may be involved [Israel et al., 1995]. Accordingly, the evaluation plan for the IAFF first responder training program included a variety of measures reflecting four different kinds or levels of course impact. The four types of measures are described below along with the methods used for actual data collection. Copies of the forms to be noted are available from the IAFF.

Level 1: Trainee reactions. The trainees completed a ‘‘Reaction to Course’’ form at the end of the course wherein they rated the various units in terms of the quality of the instruction (either high, satisfactory, or low), the amount of information covered (either too much, okay, or too little), and its utility (either high, medium, or low). In addition, each trainee was asked to rate their competency in nine (9) tasks connected with hazardous materials response as they judged them to be before the training and now afterward. The task statements simply framed the course topics in operational terms. (Examples: Recognize clues that hazardous materials are present in various common alarms. Understand the need and elements of medical surveillance. Know how to use reference sources for hazardous materials information. Apply knowledge of chemical properties in assessing exposure risk). A 4-point scale was used for rating task competency (1 = highest competency; 4 = incompetence). Both the pre- and post-course competency ratings were made at the end of the course to ensure the same level of reference in these judgments. This procedure, referred to by Berger et al. [1996] as a ‘‘retroactive pretest design,’’ was intended to avoid a respond shift problem wherein the rater’s basis for making judgments about their capabilities before taking the instruction changes as a result of the training. This response shift would confound pre- and post-course differences if measured in the conventional way.

Level 2: Learning. Each trainee took a 75-item ‘‘Knowledge Quiz’’ at the outset of the instruction and also at its conclusion, the difference between the two scores yielding a measure of knowledge gain as a result of the training. A repeat administration of this quiz was given to samples of trainees in follow-up sessions, ranging from 6 to 12 months after the course, to determine how much of the original learning was retained.

Level 3: Behavior. A set of forms, one for each course unit, was developed, each listing five (5) specific, positive actions capturing the major lessons to be learned in that unit. (The list of action statements for each unit appears in the Appendix.) After each course unit, the trainee chose one of these actions for followthrough in their roles as first responders. They could also draft their own statement if none of those listed applied to their situation. The trainee retained one copy of the actions selected for the 4 units; a second copy was collected by the instructor for forwarding to the IAFF office. The plan was to contact a sample of trainees, 6–12 months after the course, to determine the

TABLE III. IAFF Training Results—Percentage of Trainees Giving High-Quality and High-Utility Ratings to the Course Subject Matter (Ranks in Brackets)

	Florida City		W. Virginia City 1		W. Virginia City 2		Alabama City	
Course unit	Quality	Utility	Quality	Utility	Quality	Utility	Quality	Utility
Common alarms	58% [7]	73% [3]	62% [5]	65% [5]	59% [4]	70% [4]	59% [5]	75% [4]
Health and safety	83% [1]	85% [1]	77% [1]	79% [2]	77% [1]	83% [1]	77% [1]	82% [1.5]
Medical surveillance	65% [3]	73% [2]	47% [6]	52% [7]	47% [7]	61% [6]	54% [6]	66% [6]
Recognition and identification	71% [2]	69% [5]	74% [2]	81% [1]	75% [2]	82% [2]	70% [3]	82% [1.5]
Physical properties	59% [5]	64% [7]	47% [7]	55% [6]	54% [6]	55% [7]	53% [7]	63% [7]
Scene management	64% [4]	68% [6]	66% [4]	77% [3]	57% [5]	68% [5]	69% [4]	69% [5]
Pre-incident planning	61% [5]	70% [4]	70% [3]	75% [4]	66% [3]	77% [3]	72% [2]	77% [3]
Overall average:	66%	72%	63%	69%	62%	71%	65%	73%

extent of followthrough on the actions chosen. This type of data collection was intended to offer behavioral data on the course impact. As an alternative to this measure of behavior change, trainees in followback interviews in some departments were asked to describe incidents where they saw a connection between their actions and course learning. Also, to report on any changes they saw in the behaviors of fellow fire fighters as a result of the course.

Level 4: Institutional change. Trainees were asked in the same followback interviews mentioned above whether they had observed any changes in incident command practices, and/or in department policies with respect to alarms and risks of hazardous materials exposures as a result of the course.

DATABASE FOR THE EVALUATION

Table II depicts the database for use in this report, and lists six city fire departments where the IAFF course was given. While desirable, time pressures and other constraints prevented data being collected on all the proposed measures in every course offering. As shown, data were collected on at least two measures in each department. Double entries are shown in the Pre/Post Course Quiz column for three city fire departments (TX, WVA1, WVA2) to note that a second post-course quiz was given some months after the training to measure how much of the knowledge gained was still retained. The retention data were collected on a random sample of trainees who took part in follow-up interviews. The Behavior Change column also contains two types of entries. One notes where followthrough data were collected on chosen action measures, the other where one's personal experiences or those of peers reflecting lessons learned in the course were described. The numbers of trainees whose data were collected on the various measures is as shown. While admittedly limited, the database noted in Table II offers some first indications of the results of this training program.

RESULTS

Trainee Reactions

To discern course strengths and weaknesses, the percentages of trainees giving "high" quality and "high" utility ratings to different course topics provided the most differential and useful data. Table III shows such data based on IAFF first responder classes in four city fire departments. Much agreement is indicated between and within the department findings. Variations in the overall percentages of high ratings of course quality and quality across the four departments is less than 5%. All are better than 60%, which would indicate a highly favorable reaction to the course overall. As confirmed by the Kendall coefficient of concordance [Siegel, 1956], rank orders of the percentages shown for each topic in terms of quality and utility show significant agreement ($p < 0.01$) from department to department. "Health and Safety" merits the top ranking in almost every department for both quality and utility, and suggests success in meeting a basic course objective. On the other hand, "Physical Properties of Chemicals" is ranked at the bottom on these course attributes which suggests needs for improvement on this topic. More is said about this later.

As already mentioned, the Course Reaction form included a 4-point scale for trainees to rate their competency (1 = high competence, 4 = incompetence) on nine tasks related to hazardous materials response before and now after the course. Table IV gives these before/after course ratings from trainees in five departments when averaged over the nine select tasks. Also shown is the percentage of trainees

TABLE IV. IAFF Training Results: Average Pre/Post-Course Competency Ratings for Nine Tasks Related to Hazardous Materials Response

Department	Pre-course ratings	Post-course ratings	Diff	% with positive shift
Illinois City (n = 134)	2.5	1.7	0.8	65
Florida City (n = 80)	2.5	1.4	1.1	77
Alabama City (n = 109)	2.5	1.6	0.9	72
W. Virginia City 1 (n = 150)	3.0	1.5	1.5	86
W. Virginia City 2 (n = 83)	2.9	1.6	1.3	86

showing upward shifts in their competency ratings as a result of the training.

Two departments (WVA1, WVA2) display near-identical results in trainee competency ratings when compared to the other three departments. Both show relatively poor competency levels before the course and then greater amounts of positive shift afterward. A higher percentage of trainees in these two departments also display positive shifts than in the other three. In effect, these changes bring the two departments up to the same post-course competency levels as the other three. Statistical treatments of the competency data (mixed analysis of variance [ANOVA] as described by Hatcher and Stepanski [1994]) yielded significant differences between departments for both the percent showing shift and amount of shift values. Post hoc analysis showed the WVA1 and WVA2 departments to be superior to the other three in these comparisons; the IL department was found to be significantly poorer than all others in terms of the percentage showing shift.

Rank-ordering the competency data by the nine tasks within departments revealed no significant agreement across the five departments. However, like the quality and utility results reported above, competency shifts for the task, ‘‘Applying knowledge of chemical properties in assessing exposure risk’’ tended to have the lowest ranks in most departments.

Knowledge Gain

Plotted in Figure 1 are the mean percentages of correct answers on the 75-item test taken just before and at the end of the IAFF course for five departments. Differences between the pre- and post-course scores reflecting the knowledge gained from the instruction were greatest for the TX city department (25.5%), and smallest for the AL city department (11.3%). Nevertheless, statistical tests (i.e., mixed ANOVA treatments with post hoc multiple comparisons as described by Hatcher and Stepanski [1994]) found all such differences to be significant ($p < 0.01$).

Follow-up visits in three departments (WVA1, WVA2, and TX), conducted from 6 to 12 months after the course ended, included retests on the original quiz for samples of trainees randomly drawn from the class rosters. Table V shows the mean quiz scores for the pre-, post-, and follow-up course test times for these trainees and differences between these measures. Most notable is the slight drop (less than 3%) in the follow-up test scores for one department (WVA2) when compared to its post-course results. Similar comparisons for the two other departments (WVA1, TX) show losses that are 3–4 times greater. A mixed analysis of variance and post hoc multiple comparisons of the data [Hatcher and Stepanski, 1994] confirmed significant follow-up losses only for the latter two departments. One probable reason for these differential findings in retention is offered in the discussion.

Behavior Change

The aforementioned follow-up sessions with samples of trainees, also included one-on-one interviews with IAFF staff to determine carryover impacts of the course. Interviews to determine the extent of followthrough on actions chosen by the trainees have been conducted in four departments as of this reporting. During such interviews, each action statement originally chosen by the fire fighter was read, and questions asked about followthrough. An add-on question during each interview was whether the course manual had been used by the trainee after the course, and if so, how frequently. The responses of trainees in four departments are shown in Table VI.

The percentage of trainees in all samples reporting regular adherence to at least one action is 85% or better and to two selected actions is greater than 60%. Followthrough for three or more actions shows a significant drop for all but the TX City trainees. Almost twice as many TX City trainees also report at least a one-time use of the course manual in comparison to the other groups.

The selected actions with the highest followthrough counts in the four department samples revealed both similarities and differences. Three statements, each with high followthrough rates, were common in all four departments:

- Routinely observe placards and other information systems at fixed sites and on transportation cargo carriers to test my accuracy in recognizing hazard markings.
- Keep a record of my responses to alarms where hazardous materials were detected. Learn about the materials and their harmful effects.
- Devise my own ways for thinking about or reinforcing defensive actions in responding to alarms that involve

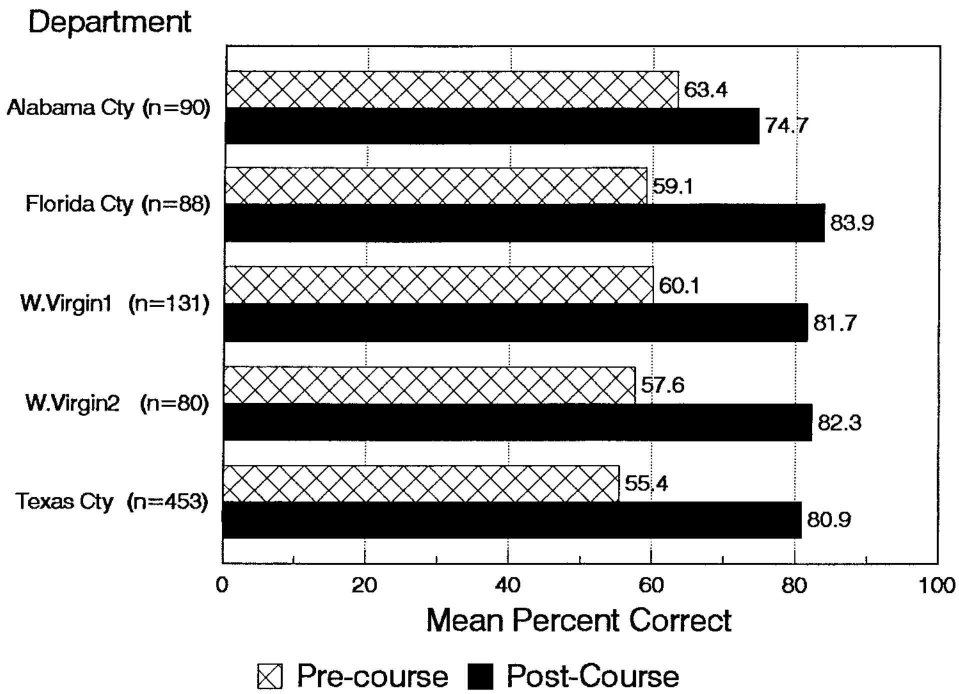

Fig. 1. Pre-/post-quiz scores by department on International Association of Fire Fighters (IAFF) Training Course.

TABLE V. IAFF Training Results: Comparison of Pre-, Post-, and Follow-up Course Quiz Results in Three Departments

Department (n = size of sample)	Pre-course	Post-course	Follow-up	Post vs pre-course (knowledge gained)	Post vs follow-up (amount lost)
TX City (N = 33)	59.3	83.6	71.0	24.3	12.6
WVA City 1 (N = 28)	58.1	81.6	71.8	23.5	9.8
WVA City 2 (N = 18)	57.3	82.7	79.9	25.4	2.8

strange odors, unknown spilled materials, or other potential chemical releases.

Those specific to the different departments were:

- Conducting site visits for pre-planning for potential incidents (AL City).
- Checking information on chemicals stored at sites in ‘‘first due’’ area (i.e., primary area for a fire station’s response in case of an alarm) to ensure accuracy, completeness (IL City)
- Discussing with others on shift, responses to calls involving hazardous materials and lessons learned (FL City, AL City, TX City)

Giving further substance to these actions were trainee answers to questions describing if and how course lessons had affected their behaviors or that of their peers in responding to alarms. Responses to these questions gained from interviews in three departments (WVA1, WVA2, TX) are among those summarized in Table VII. Positive answers in two departments (WVA2, TX) appear to outnumber those shown for a third department (WVA1), especially with regards to linking personal actions in the field with course learning. Excerpts from all three include greater use of SCBAs in responding to common alarms, checking reference guides to identify potentially hazardous chemicals when responding to calls, limiting actions to defensive control measures while awaiting the HAZMAT team, and cleaning personal protective equipment more frequently.

Institutional Change

In the post-course interviews described above, trainees were also asked:

> What effect, if any, has this first responders training had on your department’s standard operating practices (SOPs)?
>
> What effect, if any, has this first responders training had on officers and command staff within your department?

Responses to these questions for three departments are found in Table VII.

Interview data of trainees in two departments (WVA2, TX) suggest a greater effect than that found in the third

TABLE VI. IAFF Training Results: Summary Data on Followthrough Efforts on Actions Selected by Fire Fighter Trainees

Illinois City **sample size = 36** **response rate = 33/36 (92%)**	**Florida City** **sample size = 40** **response rate = 30/40 (75%)**
28 (85%) regularly following through on at least one action	29 (96%) regularly following through on at least one action
20 (61%) regularly following through on two or more	21 (70%) regularly following through on two or more
11 (33%) regularly following through on three or more	12 (40%) regularly following through on three or more
6 (18%) regularly following through on four	3 (10%) regularly following through on four
Manual use after course: none = 15 (45%); once = 3 (9%); more than once = 15 (45%)	Manual use after course: none = 13 (43%); once = 8 (27%); more than once = 9 (30%)
Alabama City **sample size = 35** **response rate = 27/35 (77%)**	**Texas City** **sample size = 35** **response rate = 35/35 (100%)**
27 (100%) regularly following through on at least one action	31 (89%) regularly following through on at least one action
23 (85%) regularly following through on two or more	27 (77%) regularly following through on two or more
13 (11%) regularly following through on three or more	22 (63%) regularly following through on three or more
3 (11%) regularly following through on four	7 (20%) regularly following through on four
Manual use after course: none = 12 (44%); once = 1 (4%); more than once = 14 (52%)	Manual use after course: none = 8 (23%); once 15 (43%); more than once = 12 (34%)

(WVA1). There appears to be more attention to HAZMAT-related procedures, and more thoughtful incident commander actions in responding to calls having the potential for harmful exposures. Regarding the latter, two of the most telling personal accounts of how the course changed command behaviors are noted below:

Report: ‘‘Rail tank car left track and was spilling contents. I assessed situation from a distance with binoculars and determined that product in car was chlorine. A defensive perimeter was set up and a request made for the HAZMAT team. Before the course I would have just rushed up to the tank car to see the situation up close and probably exposed myself and others needlessly to this harmful chemical.’’

Report: ‘‘Before course, our Captain had little use for reference sources such as the DOT ERG. (DOT ERG refers to the Department of Transportation Emergency Response Guidebook, which lists hazardous chemicals along with basic information about their physical properties. Phone numbers to call for assistance are included along with evacuation distances). After the course, we responded to an alarm at a residence where our Captain spotted several containers on site. Viewing the placards and labels together with the ERG, a determination was made to pull back the company and call in the HAZMAT team. It was learned that one of the container contents would have reacted with water if we had proceeded and started hose lines. In the past, the Captain said he would have been the last person to pick up and use the ERG.’’

Incidental information obtained during followback contacts with departments suggested other supportive administrative actions being taken since the course. Among those noted were:

- Computerized information on chemical materials and properties as found at sites with people occupants; reference texts in vehicles now carry notations of whether they have equipment onboard to respond safely if called to the premises
- Arrangements with local hospitals to pick up, sterilize, and return fire fighter clothing suspected of being contaminated with biologic material
- More decontamination sinks installed in fire stations

RELATIONSHIPS AMONG THE VARIOUS MEASURES

Implicit in a popular model of training evaluation criteria is the assumption that positive attitudes about the quality of the instruction or course content, gains in skills or knowledge, behavior change and end goals of such actions are intercorrelated if not causally linked [Kirkpatrick, 1967; Alliger and Janak, 1989; Kirkpatrick, 1994]. Some suggest an ascending order of these measures; that is, favorable reactions to training lead to learning which, in turn, leads to changes in job behaviors, which then leads to related changes in an organization [Newstrom, 1978; Clement, 1982]. Data on the multiple measures from the course evaluations just described offered a means for verifying these assertions. Specifically, it was assumed that more positive reactions to the quality/utility of the course would be correlated with improved competencies and knowledge gain, and these measures, in turn, would be correlated with more behavior change. Pearson product-moment correlations computed between these different measures [Hatcher and Stepanski, 1994] for trainee data obtained in each department are shown in Table VIII.

The highest coefficients are found between measures of course reaction and final competency level. All are in the expected direction, i.e., more positive reactions to the course correspond to better final ratings of competency, and are statistically significant. Other significant correlations, however, are few and scattered among the remaining measures.

TABLE VII. Trainee Responses to Questions on IAFF Course Impact

Questionnaire item	Texas City	W. Virginia City 1	W. Virginia City 2
Effect of first responder training on department SOPs	57%: efforts to update, increase knowledge of SOPs; more systematic approaches to HAZMAT situations 43%: don't know	18%: added interest in procedures but nothing formalized 82%: noted no specific impact on SOPs	33%: efforts made to review, update materials 66%: don't know
Effect of first responder training on officer/command staff	77%: greater awareness of chemical hazards in calls; in sizing up scene before taking action; more subordinate input on training needs 23%: no effect	39%: caution in sizing up scene before taking action; check for agents that may pose risks 39%: no difference 12%: couldn't respond	83%: more caution in sizing up scene before taking action; more use of SCBA and PPE washing/cleaning 17%: no effect
Effect of first responder training on fellow officers	94%: more use of SCBA; cargo placards and references checked for agent properties before taking action 6%: no effect	82%: more use of SCBA: site visits to check on hazardous materials; equipment needs noted 18%: none noted	89%: greater awareness/talks about calls with suspect materials; more use of SCBA; more frequent gear decontamination 11%: no effect
Connection of first responder training to field experience	71%: more donning of SCBA in common alarms; reference sources checked on agent hazards; actions taken to limit agent spread until HAZMAT team arrives 29%: none recalled	50%: greater use of SCBAs in common alarms; concern for limitations of PPE in alarms at a chemical storage site; defensive measures in handling roadway spills 50%: none recalled	76%: cargo carrier shapes/placards noted in calls to highway spills; diking and damming used to control spread of chemical agents with SCBA fully deployed, and washed after use 24%: none recalled

Even measures of competency and knowledge show significant association in only a few instances and correlations between these measures and behavior are fewer still.

DISCUSSION

First results from evaluating the IAFF hazardous materials training course for first responders appear positive. Substantial percentages of trainees rate the instruction as high in quality and utility, and judge themselves more competent in handling tasks when alarms involve risks of exposure to hazardous materials. Significant gains are noted too in the level of knowledge post training, and trainees report more self-protective and preventive actions based on course learning. Lastly, and though still fragmented, there are indications of institutional changes which seem due in part to the trainees' learning experience.

While gratifying overall, the aforementioned results also reveal variable effects in some cases which deserve comment or explanation. In doing so, issues are raised not only to this IAFF course but to others having similar objectives. For example, ''Physical Properties of Chemicals,'' were ranked lowest in the quality and utility ratings given to various course topics, and also showed lesser shifts in related task competency. This subject is more technical than others, and course time devoted to this area may not be sufficient to ensure comprehension. In interchanging ideas for dealing with this topic, IAFF instructors have developed many novel laboratory-type demonstrations, enabling the trainees to witness examples of various chemical reactions. But the trainees may be viewing these effects as no more than exotic demonstrations. Further use of exercises to emphasize how such reactions could and have affected the outcomes of hazardous materials events would appear indicated.

While competency shifts for all departments suggested improved capabilities, the shifts shown for some departments were greater than found in others. As already mentioned, these results may simply be a case of those with the poorer ratings at the outset of the course catching up with the rest of the group. The IL department whose trainees show the least gain in competency scores may be the result of a problematic training schedule. Indeed, a 6-week period was necessary for each trainee class in this department to receive 3 days of instruction, with some classes having 3- to 4-week intervals between the units of instruction, and taught by different instructors. By contrast, other department training schedules were more compact and used the same instructors thoughout.

TABLE VIII. IAFF Course Results: Pearson Correlations (r) Between Different Evaluation Measures

Evaluation measure	Competency ratings		Knowledge quiz		Behavior index[c,i]
	Shift	Final	Gain	Final	
Reaction index[d]	0.50[a] (AL n = 47)	−0.59[a] (AL)	0.10 (AL)	0.14 (AL)	0.54[a] (AL n = 14)
	0.11 (WV1 n = 38)	−0.60[b] (WV1)	0.08 (WV1)	0.31 (WV1)	
	0.26 (FL n = 55)	−0.27[a] (FL)	0.02 (FL)	0.12 (FL)	−0.06 (FL n = 20)
	0.20 (WV2 n = 75)	−0.39[b] (WV2)	0.02 (WV2)	0.25[a] (WV2)	
Competency rating					
Shift[e]			−0.13 (AL n = 47)	0.29[a] (AL)	0.60[a] (AL n = 14)
			−0.01 (WV1 n = 38)	−0.02 (WV1)	
			0.32[a] (FL n = 55)	−0.24 (FL)	−0.24 (FL n = 20)
			0.12 (WV2 n = 75)	0.07 (WV2)	
					0.20 (IL n = 27)
Final[c]			−0.18 (AL n = 47)	−0.05 (AL)	−0.44 (AL n = 14)
			0.03 (WV1 n = 38)	−0.36[a] (WV1)	
			0.17 (FL n = 55)	−0.13 (FL)	−0.28 (FL n = 20)
			0.11 (WV2 n = 75)	−0.25[a] (WV2)	
					−0.17 (IL n = 27)
Knowledge quiz					
Gain[g]					−0.01 (AL n = 14)
					−0.58[a] (FL n = 20)
					−0.23 (TX n = 35)
Final[h]					−0.12 (AL n = 14)
					−0.01 (FL n = 20)
					0.03 (TX n = 35)

[a]Correlation (r) significant $p < 0.05$.
[b]r significant $p < 0.01$.
[c]Behavioral data are based on samples only; hence, n values are smaller than those shown for other measures.
[d]Reaction index for each trainee defined as the number of high ratings given to both quality and relevance for the seven course topics noted in Table III.
[e]Competency shift defined as pre/post-course differences in scaled ratings for 9 tasks in exercise.
[f]Final competency level defined as average of post-course scaled ratings for the 9 tasks.
[g]Knowledge gain based on pre/post course differences in quiz scores.
[h]Final knowledge quiz score was the post-course quiz score.
[i]Behavior index computed for follow through on action statements only. A value of 3 given to each statement for which trainees indicated regular followthrough efforts, 1 for sometimes efforts, −1 for no effort, and 0 for no opportunity. A bonus of ½ point was given to trainees indicating one-time use of course, and 1 point for using it more than once. The sum of these measures for each trainee defined his/her behavior index.

The need for extended time schedules to cover the 3-day course requirements may be even more detrimental to scores on the knowledge quiz. One could argue that the quiz results for IAFF training could be better if the course were given on 3 successive days and with the same instructors. At the minimum, this would alleviate forgetting factors especially for material presented earlier in the course. The greatest gains in quiz scores (36%) for the IAFF first responder course have been posted by cadet classes at the fire academy where training conditions favor 3 successive days of instruction by the same instructors.

Recommendations for resolving the time schedule problems just described are being considered by the IAFF. One possibility is to arrange for and pay backup personnel to cover the duties of those attending classes so that the trainees can complete the courses in minimal time. Another is to videotape or produce CD-ROM versions of the course so that those unable to meet class schedules, or called out on emergencies, could still keep up with the course work. The latter is also being considered for refresher training. On this point, extra training may be one explanation of why trainees in one department (WVA2) showed little loss in knowledge quiz scores when retested 6 months after the end of the IAFF course. Roughly two months after completing the IAFF course, the WVA2 trainees received an added day of hands-on instruction in techniques for managing hazardous materials incidents from another organization. Besides better retention, the original course learning reinforced by this added instruction may also be responsible for the other strong carryover effects found for trainees in this department

(Table VII). Recommending added training days or refresher courses should have some merit in light of these findings. IAFF has given consideration to adding an extra day to their first responder course for staging mock drills and practical exercises. It bears mention that issues of refresher training are currently under review as part of the OSHA proposal for an occupational safety and health program standard [OSHA, 1996].

Positive actions from the course were most evident in the responses of the TX city department trainees. A greater number of TX city fire fighters followed through on more of their chosen action statements, made use of the course manual, and reported similar actions by their peers. Two reasons for this result come to mind. The first is that the TX city in question was a major petrochemical hub where hazardous materials calls are near daily occurrences. This increased risk presented numerous opportunities to put the course lessons to use. The second is that the IAFF instructors of the TX city course were themselves local fire fighters who used the local experiences and incidents to give the course and its objectives added cogency. Fire fighters in the other cities included in this report were not subject to the same risks, and thus could not be expected to have the same motivation.

Correlations found between the various evaluation measures give only limited support to the assumption that they would be interdependent. The evidence for association was most obvious for reaction index and competency measures and not for the others. One possible explanation is that the data on these two measures were collected from the same form which was administered at the same time (i.e., at the end of the course). Measures of knowledge gain reflected two different times of administration, and data on behavior change were obtained well after the course ended. Hence, temporal factors may enhance associations in one instance and obscure them in another.

Admittedly, restrictions in the variance for the different measures used in this analysis probably hampered prospects for obtaining larger correlations. For example, competency ratings could vary only across a 4-point scale, and reaction and behavior indices were likewise limited in their range of numerical scores. On the other hand, and as reported by Alliger and Janak [1989], attempts to show interdependency among different levels of evaluation for other training courses have similarly yielded uncertain results. It is believed that uncontrolled factors such as differences in trainee motivation and attitudes, and situation-specific conditions affecting transfer of training may be responsible for the failure to show greater linkages.

Ways to improve the content and delivery of the IAFF first responder training course and its evaluation continue to be the subject of internal review. Adding extra days to the course for practical hands-on exercises, and computerized versions for refresher training and other uses have already been mentioned. Refinements to the evaluation plan are also being addressed, especially easier ways to collect and process data on the various measures to gauge course benefits. However, the plan remains built on largely subjective or self-report measures and, as we have seen, positive changes in some do not correspond to similar changes in others. Whether such training results in fewer occurrences of the sort reported at the outset of this paper will require more verifiable evidence. More objective study of first responder actions and their consequences before and after training is needed.

ACKNOWLEDGMENTS

The author thanks the staff of the Hazardous Materials Training Department of the IAFF for their support in developing the evaluation plan, overseeing its implementation, and the collection of data. Thanks are also due the IAFF training instructors and trainees who furnished the critical data for assessment. Materials for the IAFF first responder training course were developed via grants from the National Institute of Environmental Health Sciences. The National Institute for Occupational Safety and Health (NIOSH) is currently funding IAFF direct training of this course and its evaluation. The opinions, findings, and conclusions expressed in this document are those of the author and not necessarily those of the IAFF or NIOSH.

APPENDIX: ACTION STATEMENTS

[Lead sentence to each unit]: "I intend to do things that I have not done up to now or will do differently. Specifically, I will.

Unit 1: Common Alarms

1. Review department's standard operating procedures (SOPs) and guidelines to see whether they cover risks of hazardous materials in responding to common alarms. Will suggest ideas for improving SOPs or guidelines.
2. Routinely discuss with others on my shift our responses to calls involving hazardous materials, reviewing conditions and any particular risks to department personnel, civilians, or the environment. Discuss lessons learned from these incidents.
3. Request to attend other hazardous materials training courses.
4. Drive or walk through my first due area to note places likely to be occupied by people, transportation corridors, and other sites where hazardous materials may be stored, carried, or disposed of.
5. Devise my own ways for thinking about or reinforcing defensive actions in responding to alarms that involve strange odors, unknown spilled materials, other potential chemical releases.
6. [Use this space to write your own statement.]

Unit 2: Health and Safety

1. Avoid contact with persons or equipment that may have been contaminated in a hazardous materials incident until they have gone through a decontamination procedure.
2. Review my department's procedures for reporting exposure to ensure that reports are handled efficiently and confidentially.
3. Report any unusual signs of irritation, dizziness, breathing difficulty, hand/feet tingling, nausea that I might feel following responses to alarms where toxic materials were present.
4. Keep a record of my responses to alarms where hazardous materials were detected. Learn about the hazardous materials and their possible harmful effects.
5. Take steps to decontaminate my own clothing and equipment in cases where actions taken on the fireground exposed me to toxic materials.
6. [Use this space to write your own action statement.]

Unit 3: Hazardous Materials

1. Make up a chart showing container/tank/vessel shapes and label them for likely contents of hazardous materials.
2. Compare recommended actions in the DOT Emergency Response Guidebook for major chemicals found at sites in my first due areas with department SOPs, resolve any significant differences.
3. Refer to hazardous materials references, materials safety data sheets, or shipping papers to learn more about the properties of chemicals commonly used or stored in large amounts in first due area and actions to be taken in cases of overexposure.
4. Routinely observe placards and other information signs at fixed sites and on transportation cargo carriers to test my accuracy in recognizing hazard markings.
5. Engage in company preplanning for response to sites in the first due area where there are known water reactive chemicals, chemicals with high vapor densities, liquids in closed vessels, radioactive materials, other unique hazards.
6. [Use this space to write in your own action statement.]

Unit 4: The Planned Response

1. Check SOPs in my department on plans and procedures for routine and emergency decontamination, suggesting needed changes to my supervisor based on information gained in the course, personal knowledge, or experience.
2. Check SOPs in my department dealing with cleaning, inspection, and storage of personal protective clothing and respirators, suggesting changes to my supervisor based on information gained in the course, personal knowledge or experience.
3. Conduct pre-incident surveys of occupant sites, visit warehouses, factories, fabrication shops in my first due area to learn if large amounts of chemicals or other substances are in use or stored, ensuring that chemical information is complete and accurate.
4. Make up a list of the ‘‘top 10’’ sites in my first due area posing the greatest risk of exposure to hazardous chemicals. Analyze and discuss plans or actions in responding to alarms at these sites with other companies.
5. Use the course manual to make up a checklist of actions appropriate for first responders at hazardous materials incidents, and grade myself for several incidents, noting the circumstances where I was unable to follow proper actions.
6. [Use this space to write in your own action statement.]

REFERENCES

Allerger GM, Janak EA (1989): Kirkpatrick's levels of training criteria: Thirty years later. Personnel Psychol 42:331–342.

Berger PK, Gunto SJ, Rice, C, Haley JV (1996): Estimating the impact of health and safety training using a retrospective pre-test design. Appl Occup Environ Hyg 11:1198–1203.

Clement RW (1982): Testing the hierarchy theory of training evaluation: An expanded role for training reactions. Public Personnel Mgmt J 11:176–184.

Hatcher L, Stepanski EJ (1994): ‘‘A Step-by-Step Approach to Using the SAS System for Univariate and Multivariate Statistics.’’ Cary, NC: SAS Institute.

Israel BA, Cummings KM, Dignan MB, Heaney CA, Perales DP, Simons-Morton BG, Zimmerman MA (1995): Evaluation of health education programs: Current assessment and future directions. Health Educ Q 22:364–389.

Kirkpatrick DL (1967): Evaluation of training. In Craig R, Binel L (eds): ‘‘Training and Development Handbook.’’ New York: McGraw-Hill, pp 87–112.

Kirkpatrick DL (1994): ‘‘Evaluating Training Programs: The Four Levels.’’ San Francisco: Berrett-Koehler.

National Fire Protection Association (1992): ‘‘Standard for Professional Competence of Responders to Hazardous Materials Incidents.’’ Quincy, MA: NFPA, p 472.

Newstrom JW (1978): Catch 22: The problems of incomplete evaluation of training. Training Dev J 32:22–24.

Occupational Safety and Health Administration (1989): Hazardous waste operations and emergency response: Final rule. Federal Register, 29 CFR 1910.120 (q)(6)(ii), March 6, 1989, p 9329.

Occupational Safety and Health Administration (1996): Draft proposed safety and health program standard. Office of Safety Standards. Washington, DC:US Department of Labor.

Siegel S (1956): ‘‘Non-Parametric Statistics.’’ New York: McGraw-Hill, pp 229–238.