Advanced Training for Crisis Decision Making: Simulation, Critiquing, and Immersive Interfaces

Janet A. Sniezek
Department of Psychology and Beckman Institute
University of Illinois
jsniezek@uiuc.edu

David C. Wilkins
Beckman Institute
University of Illinois
dcw@uiuc.edu

Patrick L. Wadlington
Department of Psychology
University of Illinois
wadlingt@uiuc.edu

Abstract

Crises demand swift and effective decision making. Yet crises entail unique characteristics that hinder training of personnel with the process knowledge necessary to achieve these two goals. First, crises are, by definition, rare; thus, it is usually not possible for humans to acquire decision-making expertise directly through experience in natural settings. Second, managing crises often involves dealing with massive uncertainty and complexity under conditions of acute stress. Each of these features poses a unique challenge to training. We present an example of a trainer for ship damage control that addresses these challenges. It consists of a first-principles simulator that generates large numbers of realistic scenarios, an immersive multimedia interface that helps replicate decision-making information overload, and a critiquing expert system that provides real-time and post-session feedback on human decision-making performance. Experimental results are presented that indicate that the described computer-based trainer has psychological realism from the standpoint of allowing a trainee to practice decision making processes while under a high level of stress.

1. Introduction

"One very specific environment in which decision making is critical is a crisis: An unexpected, life-threatening, time-compressed event, such as an engine failure or a frozen flight control that usually occurs once in a lifetime." [16]

Crisis management is a unique skill that is critical, but difficult to acquire. Analyses of disasters often point to mistakes that humans make that worsen, fail to prevent, or cause disasters. For example, in the aviation community an industry-wide analysis has shown that over 70% of aviation accidents can be attributed to human error [18] and in medicine it is claimed that between 70% and 82% of anesthetic incidents in the operating room have been attributed to human error [4]. Some human actions can have widespread disastrous consequences, such as the bombing of an embassy, the shooting down of a civilian airliner, or meltdown in a nuclear power plant.

In some domains, conventional non-computer based training approaches have serious limitations with respect to training personnel for crisis management duties. Many of these limitations can be reduced with the use of advanced training technologies that include three components: an immersive multi-media interface, crisis simulator, and a critiquing system. In this paper we identify some of the theoretical and practical issues in developing training programs for crisis management, and provide an illustration of a training system for crisis management on US Navy ships. In the first section, we present a theoretical framework for understanding the special problems in training for crisis management. Next, we show the advantages of a system with an immersive interface, crisis simulator, and critiquing system compared to "conventional" training (i.e. non-computerized systems with none of these features). Lastly, we describe the case of training for ship damage control and the special benefits of DC-TRAIN, a multi-media immersive trainer with crisis simulation and a critiquing system.

2. Crisis management

Crises come in many forms: a devastating earthquake, hostile takeover of the company, airplane crash, food or fuel shortage, taking of hostages, stock market crash, or enemy attack. Regardless of the domain, crises usually share four features:

2.1. Uncertainty

There are many sources of uncertainty for those who attempt to influence the events of a crisis; the cause of the problem and its extent or duration may be unknown, or unknowable at the time of the event. More importantly, it is difficult to know how to proceed to manage the crisis. To take effective actions, it is necessary to know the likely course of multiple events and how to allocate resources in dealing with them. While some of the uncertainty results from the rapid onset of the crisis, some of it is due to the inherent complexity and unpredictability of the events comprising the crisis.

2.2. Rapid onset

Of course, a crisis may be long in the making, but the point at which there is consensus that a crisis exists is typically characterized by a rapid unfolding of events, and requires timely action in response.

2.3. Imminent or realized severe losses

The goal of the response is to prevent or reduce negative consequences. Surprising events that are difficult to influence are labeled crises only if it is apparent that something of value is being lost, or soon will be.

2.4. A lack of controllability

Despite the urgency, effective action is difficult. Typically the losses are irreversible and, to some extent, unpreventable. Although the events of the crisis cannot be completely controlled, they can be partially influenced. Crisis management is the set of actions taken to exert control over the events of a crisis to minimize losses. It is similar to risk management—except that the events are real, not potential, and action takes precedence over planning. To be effective, crisis management involves making good decisions under severe time pressure and uncertainty.

3. Problems with conventional training

Crisis management is a highly complex skill that is difficult to acquire for many reasons. First, crises are, by definition, rare, making it difficult to acquire direct experience in the management of crises in a given domain. Further, experience is not always a good teacher. The highly uncertain and complex environment of a crisis can be a poor place to try to discern cause-effect relations. Second, when crises do occur, conditions are unfavorable for training [7]. Those with relatively more expertise must act and have no time to attend to the needs of a trainee. It may not even be possible for novices to be in a position to passively observe the decision making of experts during the crisis. Third, crisis management skills may not generalize well across crises. The complexity and uncertainty of the environment in which crises occur mean that each crisis is unique. The scenario of a disaster is rarely the one for which personnel were trained. In addition, the skills and knowledge required need to be updated continuously both because the environment of crises is dynamic, and because the regulations and technology for managing crises continually changes.

3.1. Finding a trainer

One means by which organizations can acquire expertise in crisis management is to hire those who are experts and replace them with new and better experts as requirements change. This means is rarely feasible due to a shortage of those with experience. Another problem with trying to hire crisis management expertise is the specificity of skills needed in any domain. The army general who brilliantly responded to an enemy attack may be ineffective in managing a natural disaster.

An alternative way for an organization to acquire expertise in crisis management is to train existing personnel. One of the first major problems to be addressed is deciding who does the training. An organization can hire an outside professional trainer who is regarded as an expert both in the domain and in teaching. A major disadvantage is cost--which usually translates into a finite and insufficient period of time for training. Further, an outside trainer may not customize the program to fit the needs of the specific organization and may be viewed as an outsider by the trainees. Often, outside experts do not exist, or if they do, it is not easy to verify their level of competence in either the domain or in training. Another option is to use someone within the organization to do the training. This resolves some of the major problems associated with outside consultants; the inside trainer is likely to be less costly, more available, and better accepted. However, an experienced person who learned his or her skills from within the organization may lack exposure to alternative ideas and methods and thus perpetuate one organization’s limited approach to crisis management.

3.2. Determining training content

The second problem in establishing a crisis management training program is conducting a needs assessment to identify what needs to be taught. Those with experience with past crises are one source of information. Of course, any one person’s direct experience with crises will be limited. But even if multiple persons have input on the content of the training, there is still the problem that none of them can predict the exact nature of a future crisis. This may lessen respect for the training, promoting an attitude that “they cannot teach you what you need to know.” Although no one can know exactly what needs to be taught, it is likely that both domain knowledge and procedural knowledge as well as general problem solving strategies are crucial to crisis management. The latter requires a wide range of skills (e.g., communication, social, perceptual, and decision making) that further complicates the process by widening the scope of training.

3.3. Performance evaluation

Once the content of a training program is determined, there is the problem of assessing how well the trainee has mastered it. There are two specific issues. One is measuring performance for purposes of personnel decisions, e.g., deciding if the trainee has successfully completed the training, needs remedial training, or is the best of the training group. Second, performance assessment data on all training is important input for utility analyses to determine the value of a training program.

3.4. Feedback

Another purpose of obtaining information about performance is to give the trainee feedback that can be used to improve performance during training. To be most effective, feedback needs to be prompt and specific. The information given to the trainee should not only identify the error, but also provide sufficient explanation for the trainee to develop an improved understanding of the crisis event. Generic instructions that repeat rules are generally not as beneficial as instructions directed at the particular deficiency displayed by the trainee.

3.5. Interactions

Delivery of feedback requires some form of communication between the trainer and trainee. It its simplest form, this communication is unidirectional, from the trainer to the trainee, and occurs only after execution of actions in a crisis scenario. Cost considerations often confine interactions to a single trainer communicating with multiple trainees. However, communication to instruct and improve trainee performance is most effective if is bi-directional as well as individualized. Two-way interactions that allow either party to initiate communication have particular advantages; the expert can choose to deliver advice or feedback or the novice can request such assistance. There are occasions in which it is more efficient for the trainee to request advice and then practice executing it than to practice with faulty actions and subsequently try to correct them. Expert tutoring often involves the use of hints and suggestions. Conventional "teacher-centered" training rarely allows for consultation with expert on demand.

3.6. Scheduling

At a more practical level is the scheduling problem of determining where and when to do the training. Training for crisis management will be competing with other organizational activities for time and space as well as other resources. As long as the trainer must be physically present with the trainees, there is major scheduling problem and a risk that training time will be limited to insufficient levels. Effective training time per person may be reduced also by having a single trainer concurrently work with multiple trainees.

3.7. Cost

Another training issue concerns the level of expenditures for materials, equipment, and possibly support personnel. This can be difficult because there is not always a clear payoff from greater investments, but insufficient material support can render all other training efforts worthless. This category of training cost is closely linked to training method. Ideally, the method is chosen for its fit to the skills requirement, but in practice, cost is a major factor in decisions about training method. For employees to truly have a grasp of new skills required for crisis management, they need to be exposed to training numerous times over long period of time, which greatly increases training costs.

3.8. Realism

The realism of training is one of many issues pertaining to the transfer of training. But given its special significance for crisis management training, we are treating it as a separate topic. Some of the forms of training that are least expensive, such as classroom training and written materials, have the least realism. The most realistic training takes place in contexts that represent the context of an actual crisis. If done correctly, realistic training is more beneficial than classroom training due to the advantages of active training over passive training [1]. A lack of realism can prevent transfer of training (e.g., for performance under time pressure) [28]. The exact requirements for realistic training vary across domains, but usually involve some combination of natural and simulated features. For financial and ethical reasons, it is not possible to involve real events that lead to real losses. Thus, training contexts often make extensive use of simulations that approximate crisis events. Yet, even very expensive simulations that look like crisis may not be sufficiently realistic.

For training purposes, a simulation is realistic if it induces the same *psychological processes* in the training context that are experienced during an actual crisis [12]. Perhaps the greatest challenge in training for crisis management is inducing the psychological processes associated with the acute stress experienced in actual crises. Acute stress can be defined as a state that occurs in situations of potential harm, time pressure, and arousal [19]—all characteristic of crisis management. Although there has been a longstanding interest in the problem of preventing decrements in human performance under stress, research on human performance under acute stress is limited [19]. It is difficult to conduct research on human performance in crises for the same fore mentioned reasons it is difficult to train. However, there has been some study of some components of acute stress—arousal, time pressure, and anxiety. Each of these factors has been found to hurt performance. Salas, Driskell, and Hughes propose that the manner in which individuals evaluate the level of potential harm, time pressure, and demands of the task plays a key role in how well the individuals perform. According to Driskell and Salas [5], the "training and simulation" intervention strategy has the goal of changing the appraisal process in the face of a crisis. The specific recommendation is to "provide the individual with the knowledge and confidence that will lead to the development of positive performance expectations, and result in more effective task performance under stress" (p. 187). It follows that one way to promote effective crisis management under acute stress is to train individuals to perform under acute stress, or at a minimum, some combination of arousal, time pressure, and anxiety.

3.9. Transfer of training

A final problem that needs to be addressed with any training program is transfer of training. The motivation for realistic training is to allow for the skills acquired in training to be applied in the context of an actual crisis. There are other factors that can prevent transfer of training for crisis management. One is that the nature of crises can change over time due to changes in the physical, social, or economic environment. In addition, crisis management technologies and procedures change. In short, one problem is that today's training may not help tomorrow's performance.

4. Training with immersive interfaces, simulators, and critiquing systems

One alternative solution to the problems of conventional training methods is the use of advanced computer-based training technologies. This approach involves use of a training program that is initially based on established specific doctrine used by an organization [14]. Computer-based systems to facilitate human learning exist in many forms, but here we are interested in those that combine a simulator for generation of training examples, an immersive interface to assist with replicating the information overload associated with many crisis situations, and a critiquing system for generation of the feedback that is normally given by a teacher. We shall describe the importance of these capabilities for crisis management training.

Table 1 summarizes the major issues in training for crisis management, and the ways in which an expert critiquing and simulation system can address the problems associated with conventional training approaches. An obvious advantage of an expert critiquing system is that it can alleviate the problem of a shortage of human experts in training. Subject matter experts are still essential; they provide the necessary knowledge for development of the critiquing system and crisis simulator. The human experts are valuable also in testing and validating the system. The real contribution of the expert critiquing and simulation system to the problem of a shortage of human experts is realized once the system has been perfected. Multiple versions of the computer-based trainer allow many students to be trained simultaneously. Further, this replication provides standardization that is usually not present with training from multiple experts in a complex domain. As is true with the adoption of many forms of technological trainers, a shift in emphasis occurs in who and what is at the center of the endeavor. Whereas conventional training is instructor-centric and emphasizes teaching, training with an expert critiquing and simulation system is trainee-centric and emphasizes learning.

Whether training in crisis management is conducted by the experts themselves or with a computer-based system, it is subject matter experts who provide the domain and procedural knowledge necessary that will be the context of training. But in building a computer-based system, expert knowledge is elicited from multiple subject matter experts, thereby combining their respective inputs into what can be a more comprehensive knowledge set than any individual expert possesses. This helps designers create a system that can simulate a large number of crisis scenarios that require the trainee to engage in actions requiring a wide variety of skills, such as recognizing patterns, detecting signals, coordinating with team members, communicating effectively, predicting outcomes of events, making decisions, and remembering protocol.

4.1.Critiquing systems and performance feedback

The fact that the expert critiquing system can analyze a simulated crisis and the actions taken to control it means that it can also assess the trainee's performance according to well-defined criteria. This can be done with conventional training techniques also, but it requires an extra effort that is rarely undertaken with the same care as is put into design of the expert reasoning system. Conventional training makes great use of subjective methods of performance assessment, for example, by having trainers rate trainee performance. In other cases, it tends to be focus on outcomes—which are easily observed—rather than process. In crisis management, outcomes are a misleading indicator of performance. Losses are almost always guaranteed, but do not necessarily reflect poor crisis management. Performance assessment with an expert critiquing system concentrates on the quality of the reasoning process used in managing the crisis.

Although an expert critiquing system yields optimal decision strategies, this does not by itself guarantee that students will learn from it [8]. The real benefit of a critiquing system can be seen in performance feedback supplied to the trainee, a key feature of many intelligent tutoring systems [20]. It is important that trainees understand their deficiencies in training exercises [6]. Human experts can tailor feedback to individual trainees, but it is a very labor-intensive process. The expert must understand the reasoning of the trainee as well as the events of the crisis and the actions taken by the trainee in response. An expert system can do this analysis automatically, in real time, allowing for swift feedback. Speed, of course, is not enough. One important feature of this feedback is that it emphasizes the trainee's process, and not just the outcomes. It is well-known in human decision making research that process feedback is superior to outcome feedback [2]. A potential problem with this feedback is that if it is counter-intuitive, trainees may reject it and not improve in performance. Explanations are part of the critiques; thus the feedback is more likely to be understood. Explanations also increase trust [21], which in term promotes acceptance of feedback.

A recent model of the effects of feedback suggests that different types of feedback are required at different levels of expertise [13]. The authors demonstrate in a meta-analysis that the traditional explanations of the effects of feedback are not enough to account for the variance in the findings over studies on the effects of feedback. In their model, feedback focuses the trainee's attention to levels in a hierarchy of goals. Feedback that focuses attention on lower levels of the hierarchy is best for learning on complex tasks. As the individual's skill increases, feedback focusing on overall task performance will be more useful. An expert critiquing system can provide feedback about performance at any level at the time it is most valuable.

The instructional process is further enhanced if there are various means for communication of feedback and other performance information. An expert critiquing system is very flexible in terms of the timing and nature of interactions with the trainee. Because the expert critiquing system can analyze events and actions as they occur, there is no need for any delay in providing information to the trainee. The trainee can even request advice prior to acting. Or, the system can interrupt the trainee to correct a poor action before it becomes a habit. The ability to freely interact with the trainer for advice prior to an action or feedback after one, provides features of human tutoring that have not yet been implemented in training technologies. Notice that the advice or feedback can be at the level of a particular belief or action, or at the level of the trainee's model of the crisis and strategy for action.

3.9. Training timing, cost, and transfer

Another collection of problems has to do with the simple questions of the when and where to do the training plus what materials are needed for good training. If written correctly, the expert critiquing and simulation system can be a self-contained training system. This characteristic allows training to occur at any place and any time. The only material needed is a moderately high-functioning computer. In essence especially in today's day and age, this makes the training system flexible [17]. For example, if so desired, an employee could train in the comfort of home whenever desired. Any feature that increases flexibility removing barriers to practice—all else being equal—will increase actual training time. Opportunities for practice allow the trainee to realize the benefits of critiquing. Understanding one's deficiencies with a skill and then being able to practice the skill in context is necessary for effective tra ining [6].

As for the new problems that arise with the high-tech approach, cost is definitely the largest. The initial research and development costs of an expert training system with extensive simulation capabilities is prohibitive in many domains. It can be unaffordable and unwise, to modify a system from another domain. If the initial costs of system development can be met, there is the advantage that implementation and maintenance can be predicted with reasonable accuracy. Finally, although cost is a definite obstacle in the present, it may prove to be trivial in the near future. Cost is likely to be reduced with the further advancement of technology, both in terms of hardware and software, and in development costs for expert training systems.

A powerful simulator that presents scenarios to the trainee through an immersive interface greatly improves the chances for transfer of training. The simulator creates an active learning environment that is preferred to passive learning environments [14,15, 25]. The more closely the simulator can stimulate the psychological processes of crisis management in actual crises, the more realistic is the training. In particular, simulations can create acute stress through multi-media interfaces, time pressure, information overload, and graphic representation of the most disturbing losses of a crisis. As a trainee's exposure to total number of scenarios and number of different types of scenarios increases, the trainee's ability to cope with uncertainty and transfer training to a real crisis improves. Experience managing many kinds of crisis scenarios reduces the chance that an actual crisis will be so novel as to render all training experience irrelevant.

5. The DC-TRAIN system and ship damage control

The US Navy's strategy for achieving effective damage control operations on fleet begins with training. Naval officers are trained for the position of Damage Control Assistant (DCA) at the Sea Warfare Officers School in Newport, RI. Our research group at the University of Illinois is developing a research prototype of an expert critiquing and simulation system for ship damage control training for this training program. The design of this system, named DC-TRAIN, is shown in Figure 1.

5.1. System Design

This system includes a multiple scenario generator for immersive multi-media training for Damage Control Assistants (DCAs). The multiple scenario generator creates the capacity for instructors or users to select scenarios designed to cover various teaching points, or to construct any desired combination of events producing ship damage. The scenario generator creates scenarios that vary in level of difficulty, as do actual ship crisis situations. The range of difficulty covers problems

Table 1. Problems of training for crisis management and s solutions based on an expert critiquing and simulation system

Training Issue	Problem for Crisis Management Training	Immersive Interface, Expert Critiquing and Simulation System Solution
Expert Selection and Recruitment	Few experts, available experts have limited experience with actual crises; Lack of agreement among experts; Instructor-centric training	Critiquing system makes use of expertise of multiple experts; Standardization of training; trainee-centered training
Training Content	Difficult to define precise domain and procedural knowledge; Wide variety of skills is needed	More comprehensive knowledge base; Simulation capability allows training multiple skills
Effectiveness Assessment	Extra effort to specify and measure performance, use of subjective measures; Emphasis on outcomes over process	Effective performance can be well-defined and measured by critiquing system; Emphasis on reasoning process
Feedback	Multiple acceptable courses of action; Need to tailor feedback to trainee's mental model; Time delays	Can assess trainee actions automatically; Generate individualized critiques
Interactions with Trainer	Cost encourages many trainees per trainer, and limits 2-way interaction	Efficient 2-way one-on-one interaction during performance; Advising to encourage development of mental model
When and Where to Train	Repeated periods of substantial time needed for learning of complex skills	Flexible scheduling for massive practice, training anytime and anywhere computer access is available
Cost	Costs increase with number of trainers, length of training, and as training context approaches that of actual crises	High initial development costs; Predictable implementation and maintenance costs
Realism	Actual crises impossible for training; Difficult to create acute stress	Simulation allows for representation of actual crises; immersive interface contributes to acute stress
Transfer of Training	Expert knowledge and experience may not generalize to future crises; Difficult to prepare for uncertainty of crises	Multiple scenarios aid in the reduction of uncertainty

suitable for training, i.e., those that are sufficiently difficult to pose a challenge but not so difficult that any human action is irrelevant to the outcome. The Illinois training system is expected to improve training effectiveness by making it feasible for naval officer students at SWOS to obtain extended practice with an immersive trainer over a wide variety of simulated crises.

The knowledge within this system was created from Navy doctrine that specifies how naval personnel should respond to events that create damage on their ship while at sea. The training system has the abilities to expose trainees to multiple scenarios, record their actions, evaluate the trainees' performance, and then give them constructive feedback. The encompassing abilities of such a complex system make it very apt at teaching each trainee in an exhaustive, effective, and efficient manner [11].

DC-Train provides DCA trainees with feedback regarding their selection of actions in damage control scenarios. Several areas of the psychological literature on feedback and performance were consulted to determine the form of this feedback. DCA trainees often have to infer the state of the ship and likely future states of the ship using probabilistic information. That is, the information available to the DCA trainee (e.g., cues such likely fuel sources and temperature readings) is imperfectly related to the state of ship (e.g., whether a

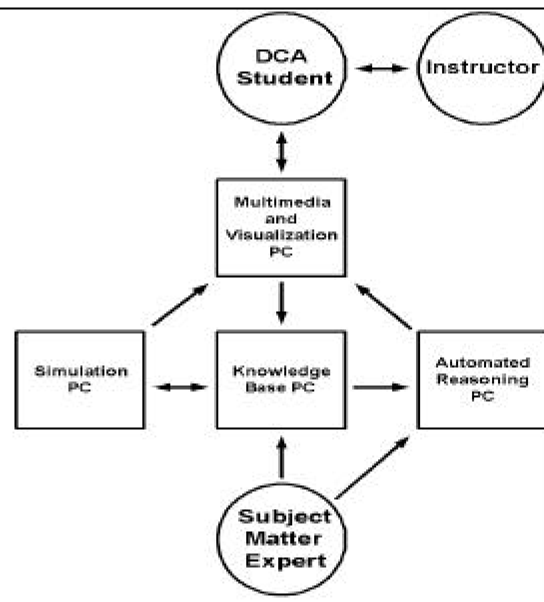

Figure 1. Architecture of the Computer-Based Training System

certain compartment is on fire or not). In problems of this sort, psychologists have found that feedback regarding the underlying relationships between the cues and various outcome states is effective in improving performance [1, 9, 11, [9, 21], while simple outcome feedback (e.g., whether the compartment really was on fire or not) is not effective [2, 26].

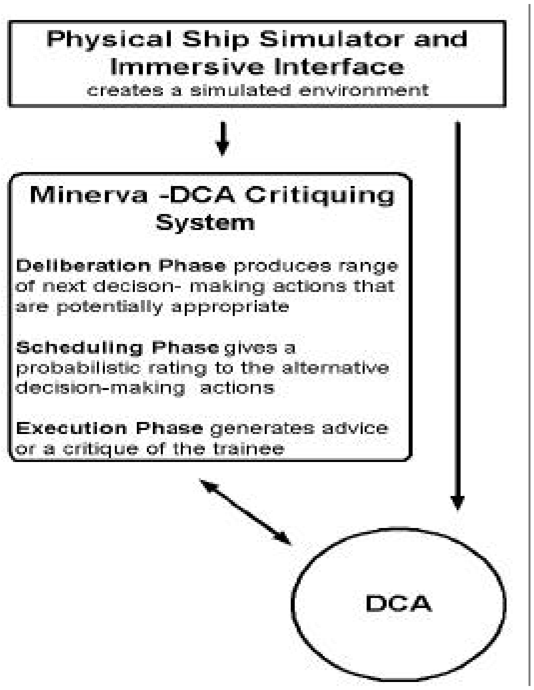

Figure 2 Interaction of System and Environment with DCA

As mentioned in the previous section, recommendations for the level and complexity of the feedback also exist in the psychological literature. Early in training on a complex task, the feedback should regard more basic strategies [13]. As DCA trainees improve, more complex feedback can be given to direct their attention to higher-level goals and strategies. Although we have yet to conduct evaluation experiments comparing different forms of critiquing, we have data that allows us to evaluate other features of the system.

6. Empirical evaluation of DC-TRAIN

In the first study to evaluate DC-TRAIN version 1.0, 58 DCA trainees completed 3 training trials with this system, as well as a single trial with IDCTT 1.0, a computer-based system that is has a high level of realism but can generate only one fire scenario. DC-TRAIN is a generalization of IDCTT that adds a first-principle simulator to generate many scenarios and a critiquing system to provide feedback. Participants were scheduled such that approximately ¼ of the teams did IDCTT first, ¼ did IDCTT after one trial of DC-Train, ¼ after 2 trials of DC-Train, and ¼ after 3 trials. The design was a 4 (DC-Train timing) by 4 (trial) with the first factor between-subjects and the second within-subjects. The participants completed various measures (described below) after each trial in the role of DCA as well as at the end of the simulator data collection.

6.1. Psychological realism

On all four trials, subjective difficulty, subjective time pressure, and effort expended were assessed via ratings. Subjective difficulty was measured by asking participants how difficult the task was on a scale from 1= not at all to 9= extremely. Subjective time pressure and effort data came from the relevant subscales of the TLX measure of subjective workload [10], where 10-point scales ranged from "low" to "high." Mean ratings for each of these measures fell significantly above the midpoint of their scales (p<.02) and did not decline over trials (p>.48).

Theory on the role of anxiety in performance in crises [3] indicates that the presence of substantial anxiety is a desirable feature of a simulator for crisis management training. State anxiety as well as trait anxiety were measured with the STAI [22]. Compared to the norms for working adults aged 19-39 [23], this sample had state and trait anxiety scores equivalent to those at the 80th and 43rd percentiles, respectively. State anxiety illustrated in Figure 3, was on average 20% higher than the baseline given by trait anxiety scores, and significantly higher than predicted by trait anxiety on trials 1, 2, and 4 (all p<.02) and marginally greater on trial 3 (p<.06). Furthermore, anxiety did not decrease over trials (p>.18). This is especially important given the significant decrease in anxiety that occurs with IDCTT after only one trial. Taken together, this suggests that DC-TRAIN 1.0 is psychologically realistic even after a number of trials. In other words, it creates and maintains a psychologically stressful environment that is necessary to prepare DCAs for what they would experience in an actual crisis. Although anxiety data are not available for the second study, it is worth noting that subjective workload ratings are comparable to those in the earlier study.

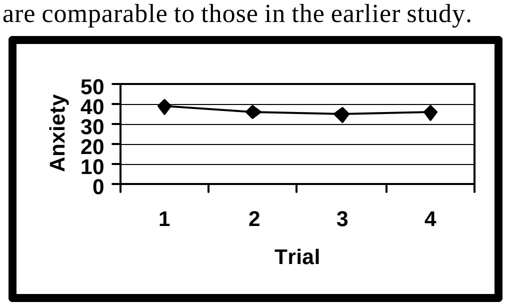

Figure 3. State anxiety scores over trials

6.2. User evaluations

Past evaluations of expert-based computerized simulations for Navy training have made use of various standards for success (e.g., successful course completion, [27]). By far the most prevalent measures for all forms of computer aids in Navy instructional settings are those of user satisfaction [24]. The elicitation of user reactions is an ongoing part of the development of DC-Train. Most of the user evaluations are informal and qualitative, and aimed at providing feedback to the developers so that they can fix “bugs” or make the simulator actions consistent with course material. Formal data on user reactions to the system, illustrated in Figure 4, show that students rate DC-Train as over a 6.5 on a 7-point scale (anchored at 1 = completely useless and 7 = extremely useful). While user satisfaction is necessary to continued use of training technologies, it is a contaminated and deficient measure of effectiveness. User reactions can reflect pre-existing positive or negative biases about the technology while they fail to capture aspects of performance critical to success in the context for which users are being trained. Thus, we turn to the true proof of the success of any training system: the performance of trainees.

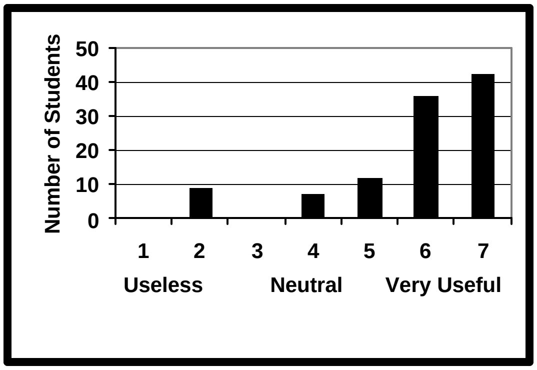

Figure 4. Subjective evaluations by trainees

6.3. Improvements in Performance

In the first evaluation of DC-TRAIN, the low number of training trials per student (4 trials at about 30 min. each) was not expected to be sufficient to affect expert performance. Because the critiquing capability and scoring system were not yet available, it was not possible to measure damage control performance in its entirety. Nevertheless, these data, illustrated in Figure 5, show some indication that firefighting errors were starting to decline by the fourth trial. That is, those students with 3 trials of practice on DC-TRAIN 1.0 made fewer firefighting errors on average on their IDCTT trial than students with less or no practice on DC-TRAIN 1.0. Although encouraging, the difference between conditions was not statistically significant. A second evaluation study with 19 DCA candidates was conducted to evaluate the effect of practice on performance. Nine DCA candidates received 3 practice trials using DC-Train, and then attempted to solve the criterion trial. Each trial used a different scenario (the criterion trial involved a different scenario than any of the three training trials). The other 10 candidates did not receive practice. After controlling for military experience, the group receiving practice had significantly higher scores on average on the criterion scenario than the group who did not receive practice ($p<.05$). Means are shown in Figure 6. Eight of the nine trainees improved, and overall, practice accounted for 37% of the variance in performance scores. In other words, practicing on one set of scenarios was followed by improved performance on a different scenario. Because the criterion trial involved a different scenario than the practice trials, this improvement cannot have been due to participants merely learning how to beat a particular scenario. While promising, this evaluation work has been limited by small sample sizes, scenarios restricted to events implemented in the most recent version of DC-Train, and low numbers of practice trials. As the availability of DC-TRAIN at SWOS increases, it will be possible to obtain data from more trainees on more scenarios with greater amounts of practice.

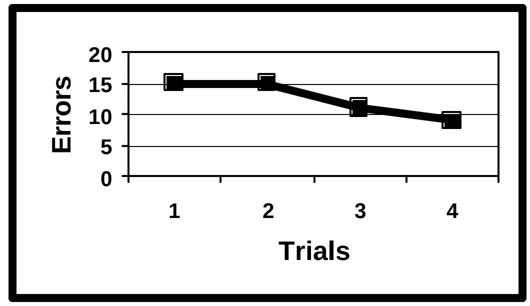

Figure 5. Error scores over practice trials

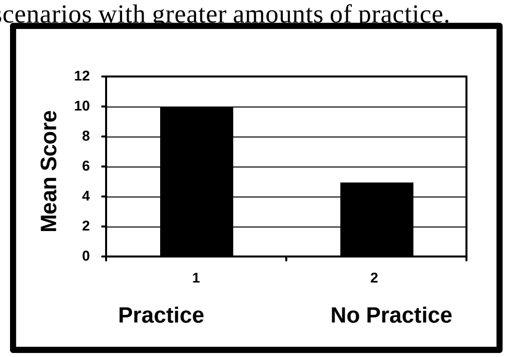

Figure 6. Average of student scores across scenarios

7. Summary

Although high-level effective human performance during crises is vital in a vast array of situations, it can be difficult to prepare personnel adequately. Crises are, by definition, rare. And when they do occur, the first priority is to handle them quickly to gain control of the situation and to minimize loss. It is not to use these hazardous, destructive, and unpredictable phenomena for training purposes. For similar reasons, it is difficult to conduct research on human performance under the acute stress of an emergency. The non-conventional approach of using systems such as DC-TRAIN that have an immersive interface, crisis simulation, and critiquing system offers one solution to both problems. It creates the opportunity for immersive practice so trainees can develop the skills necessary for dealing effectively and efficiently with the whole-task decision-making overload associated with real-life crises. Simultaneously, crisis simulations provide the opportunity for data collection under conditions reflecting the majority of a crisis' characteristics. Methods for evaluating the realism of the system for training and effectiveness in improving performance are an essential part of system development and implementation.

8. Acknowledgements

This research was supported in part by ONR grant N00014-95-1-0749. We gratefully acknowledge comments on earlier drafts of this paper provided by Michael Baumann.

9. References

[1] Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. Psychological Bulletin, 106 (3), 410-433.

[2] Balzer, W. K., Sulsky, L. M., Hammer, L. B., & Sumner, K. E. (1992). Task information, cognitive information, or functional validity information: Which components of cognitive feedback affect performance? Organizational Behavior & Human Decision Processes, 53 (1), 35-54.

[3] Baumann, M.R., Sniezek, J. A. & Buerkle, C.A. (2000). Self-evaluation, stress, and performance: A model of decision making under acute stress. In E. Salas and G. Klein (eds), *Research, methods and applications of naturalistic decision making principles*. Hillsdale, NJ: Lawrence Erlbaum Associates. In press.

[4] Cook, R.I. & Woods, D.D. (1994). Operating at the sharp end: The complexity of human error. In M. S. Bogner (Ed.), Human Error in Medicine. (pp. 255-310). Hillsdale, NJ: Lawrence Erlbaum Associates.

[5] Driskell, J. E., & Salas, E. (1996). *Stress and Human Performance*. Mahwah, NJ: Lawrence Erlbaum Associates.

[6] Dodani, M. (1999). The art of skill building without teaching. The Journal of Object-Oriented Programming, 12, 49-50.

[7] Donovan, M. A, Sniezek, J. A., & Baumann, M. (1996). Learning from Unusual Events: Decision Making in Crises. Symposium sponsored by Managerial and Organizational Cognition and Organizational Behavior divisions of the Academy of Management. Cincinnati, Ohio. August 11-14, 1996.

[8] Duchastel, P. (1991). Towards methodologies for building Knowledge-Based Instructional Systems. Instructional Science, 20, 349-358.

[9] Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. Annual Review of Psychology, 32, 53-88.

[10] Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (task load index): Results and empirical and theoretical research. In P.A. Hancock and Meshkati (Eds.) Human mental workload. Amsterdam: North Holland Press.

[11] Hays, R.T., & Singer, M.J. (1989). Simulation fidelity in training system design. New York: Springer-Verlag.

[12] Kantowitz, B. H. (1992). Selecting measures for human factors research. Human Factors, 34 (4), 387-398.

[13] Kluger, A. N,. & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin, 119 (2), 254-284.

[14] Koschmann, T.D., Myers, A.C., Feltovich, P.J., & Barrows, H.S. (1994). Using technology to assist in realizing effective learning and instruction: A principle approach to the use of computers in collaborative learning. The Journal of the Learning Sciences, 3 (3), 227-264.

[15] Majchrzak, A. (1988). The Human Side of Factory Automation. San Francisco: Jossey-Bass.

[16] McKinney, E.H. (1993). Flight leads and crisis decision - making. Aviation, Space and Environmental Medicine, 64 (5), 359–362.

[17] Mouloua, M., & Koonce, J.M. (1997). Human-Automation Interaction. New Jersey: Lawrence Erlbaum..

[18] Orasanu, J. M. (1993). Decision-making in the cockpit. In E. Wiener (Ed.), Cockpit resource management (pp. 137-172). San Diego: Academic Press.

[19] Salas, E., Driskell, J. E., & Hughes, S. (1996). Introduction: The study of stress and human performance. In J. E. Driskell and E. Salas (Eds.), Stress and Human Performance (pp 1-45). Mahwah, NJ: Lawrence Erlbaum Associates.

[20] Sleeman, D. & Brown, J. S. (1982). Intelligent tutoring systems. Orlando, FL: Academic Press.

[21] Sniezek, J.A., & Van Swol, L. M. (2000). Trust and expertise in a Judge Advisor System. Organizational Behavior and Human Decision Processes. In press.

[22] Spielberger, C.D., Gorsuch, R.L., & Lushene, R.E. (1970). Manual for the state-trait anxiety inventory. Palo Alto, California: Consulting Psychologist Press.

[23] Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., & Jacobs, G.A. (1983). State-trait anxiety inventory for adults: Sampler set, manual, test, scoring key. Palo Alto, California: Mind Garden.

[24] Sticht, T., Ellis, J., Montague, Wi., Quellmalz, E., & Slappy, J. (1993). Combining environmental design and computer programs to enhance learning in Navy technical training. Military Psychology, 5 (1), 63-75.

[25] Su, Y., & Lin, D.M. (1998). The impact of expert-system-based training on calibration in emergency management. Computers in Human Behavior, 14 (1), 181-194.

[26] Thompson, L., & DeHarpport, T. (1994). Social judgment, feedback, and interpersonal learning in negotiation. Organizational Behavior & Human Decision Processes, 58 (3), 327-345.

[27]Van Matre, N., & Robinson, C. R. (1986). The automated maneuvering board training system: An in progress implementation evaluation. Journal of Computer-Based Instruction, 13 (2), 35-38.

[28] Zakay, D., & Wooler, S. (1984). Time pressure, training and decision effectiveness. Ergonomics, 27 (3), 273-284.