

To appear in: E. Salas, C. Bowers & E. Edens (Eds.), *Applying resource management in organizations: A guide for training professionals*. Mahwah, NJ: Lawrence Erlbaum Assoc.

Airline Experiences with Resource Management Programs¹

Deborah A. Boehm-Davis, Robert W. Holt

George Mason University

and

Thomas L. Seamster

Cognitive and Human Factors

This chapter reports on the experience of a research team in developing, implementing and evaluating resource management procedures at two airlines. This project arose from the synthesis of three factors. The Federal Aviation Administration (FAA) wanted to evaluate a proceduralized form of Crew Resource Management (CRM). A regional airline wanted to upgrade its CRM training by customizing it to its operational needs. Our research team was involved in applied aviation research and was interested in developing and testing an innovative approach to CRM training. This chapter will describe the development, implementation, and evaluation of proceduralized additions to CRM at a regional airline, and the development, implementation, and planned evaluation of one aspect of this work at a major airline. The chapter will conclude with the lessons learned from these efforts.

EXPERIENCES WITH A REGIONAL AIRLINE

The goals of the work with our regional airline were to develop and implement a comprehensive, procedurally-based CRM training program and evaluate the effects of this specific CRM training on the performance of flight crews. The effort was unique in that it took the philosophy and principles of CRM that are typically only taught in the classroom and implemented them into standard operating procedures (SOP). The implementation included changes to normal briefings used on a regular basis as well as changes to checklists and procedures in the Quick Reference Handbook (QRH) that pilots use in abnormal or emergency situations.

Development of the Resource Management Training Program

Need for the Training Program

The regional airline identified a need for a program that would supplement the basic CRM training in philosophy and principles of CRM. The goal for this supplemental training was to result in objectively observable changes in crew performance for operations in normal and abnormal/emergency

conditions. The airline felt that general CRM did not fully address the requirements of their operational conditions. The airline specialized in short-haul operations in a four-season environment with both congested airports and difficult terrain at specific airports and wanted to develop new CRM procedures that would help crews cope with these flight conditions. The airline's recent application under the Advanced Qualification Program (AQP) provided an excellent opportunity to introduce these procedures and the improved CRM training.

The initial sponsor of this training was the CRM coordinator at the airline. This person acted as an opinion leader who convinced the training department and corporate management of the necessity and desirability of additional CRM training. Initially, an existing training program (Mudge, 1993) was assessed for use at this airline, but in the end a training program had to be designed to meet the specific operational needs of this airline.

Background Assessment

Organizational Climate Survey. This project began with an organizational climate survey, which included items about job satisfaction, job description, organizational and safety climate, and the functioning of the training center and the effectiveness of the training provided. This survey was distributed to randomly-selected sub-samples of pilots, flight attendants, and dispatch and ground personnel.

Overall, pilots were fairly satisfied with their jobs and were generally positive towards the organization, the safety climate, and the training center. These general results indicated that there were no broad-scale organizational problems that would interfere with developing or implementing additional CRM training. A more specific analysis was required to pinpoint specific problem areas for development of CRM procedures.

Identifying Problem Areas. The second step in the development of this program was to identify the areas of operation that most needed improvements in order to enhance safety. Key issues were determined using the results from the National Transportation Safety Board (NTSB) Safety Study (NTSB, 1994), data from Aviation Safety Reporting System (ASRS) reports, and information from a survey of the

airline's Instructor/Evaluators (I/Es). The data from these sources were used to identify the types of procedural changes that could substantially improve CRM performance in both normal and abnormal/emergency aircraft operations.

The NTSB Safety Study, which was based on accidents and incidents in U.S. airlines from 1978 through 1990, suggested that there was a pattern to problems in the cockpit:

- in 81% of the accidents, the Captain was the pilot flying;
- in 73% of the accidents, it was the crew's first day of flying together; and
- crews who had been awake longer made more errors overall and significantly more procedural and tactical decision errors than crews whose time awake was shorter.

An analysis of incidents contained in the ASRS database also suggested that there was a pattern to crew problems in the cockpit. For example, the ASRS data suggested that crew distractions and problems in information transfer played a role in a large number of reports. Specific incident reports in those areas were reviewed to obtain a more detailed problem definition.

Information about specific airline problem areas was generated from a written survey distributed to I/E pilots. This survey tapped observations made by the evaluators during the previous year.

Specifically, evaluators were asked to:

1. write down typical CRM problems with suggested reasons for these problems;
2. rank order CRM items where crews had the most problems; and
3. specify CRM items for which crews would benefit from additional training.

This survey lead to the identification of three problem areas for this airline: assertiveness, briefings, and decision making. In the area of assertiveness, the I/E pilots felt that crew members were not speaking up with appropriate persistence. Briefings were not done consistently, they were often too general, and there were no guidelines for what constituted an appropriate briefing. Finally, although crews discussed options when making decisions, they did not develop clearly stated plans of action.

Taken together, the data from these three sources lead to the development of three goals for the CRM training program:

1. reduce distractions to the pilot flying (PF) in both normal and abnormal situations;

2. increase structure in briefings to enhance the crew's performance on the first day together and improve information transfer; and
3. design checklists, the Quick Reference Handbook (QRH), and briefings to reduce workload and enhance decision-making skills, especially when crews would be fatigued, running late, or under high workload.

These goals were translated into actual procedures by a design team consisting of the airline CRM coordinator (a pilot), a pilot from a major airline acting as a design consultant, an instructional designer, and researchers specializing in aviation, cognitive human factors and team research. The diversity of background on the design team resulted in more time required for team interaction, but was important to prevent narrowness of viewpoint in the development process.

The initial design of the procedures required two working days. Procedure development was iterative and cycled between the CRM philosophy and principles at the abstract level and operational goals and procedures at the detail level. Procedures could be proposed by any team member but a proposed procedure had to 1) address one of the major operational goals, and 2) derive directly from one or more CRM principles. All team members were consulted for each suggested procedure and modifications were made until consensus was achieved.

Further cycles of development involved the fleet management, flight standards, and training personnel of the airline. These cycles were generally critiques of written versions of the proposed procedures, and this part of the process required several months. In retrospect, having a fleet representative as a member of the initial design team might have shortened the time required for fleet revisions and acceptance.

Implementing the Resource Management Training Program

Discussions with fleet representatives clarified that the implementation of the CRM procedures would require changing the full range of operational documentation to be consistent with the proposed procedures. That is, to avoid confusing pilots, it was essential that the Flight Operations Manual (FOM) and Quick Reference Handbook for operating this aircraft be consistent with the developed CRM procedures. To this end, sections of the FOM and QRH were re-written by two fleet representatives

working with the design team. The QRH was completely revised to include the procedures wherever appropriate and to apply human factors principles to QRH indexing, formatting, font and layout. The abnormal and emergency sections of the FOM were revised to be identical to the QRH and other appropriate changes were made in the FOM as necessary. These extensive revisions, and the steps required to obtain FAA approval for these revisions, required several more months. The necessity of these collateral changes and the time required for changing official documentation were critical lessons that we learned during implementation.

The next step in the implementation process was training the I/Es at the airline to understand and teach the CRM procedures. Examples of the procedures and methods for the effective instruction of the procedures were introduced to the I/Es. Based on I/E feedback, the training materials were revised to a form suitable for training pilots. While pilots were being trained in small groups over the course of a year, appropriate evaluation materials were developed and implemented as the first step in the systematic evaluation of this proceduralized CRM training program. Intermediate steps in the implementation process were instituting a Line Operation Evaluation (LOE) assessment process and training evaluators for reliable and valid evaluations. The last step in the implementation process was the formal implementation of the CRM procedures as required company SOP for that fleet.

Formal implementation of the CRM procedures required that the supplemental training for all fleet pilots be completed successfully and that the FAA-approved versions of the revised FOM and QRH be duplicated and distributed to all pilots and aircraft. In addition, formal notice was sent to all pilots about the implementation date for the CRM procedures together with an explanation in the form of a refresher videotape that illustrated these procedures. The refresher videotape was developed because the first pilot groups had received CRM procedures training up to a year prior to the formal implementation date. In retrospect, concentrating the pilot training closer to the formal implementation date would have helped to prevent memory loss and degradation of acquired CRM skills.

Evaluation of the Resource Management Training Program

Evaluation Design

The evaluation design selected for this study was a quasi-experimental design comparing the fleet that had normal CRM training plus proceduralized CRM against a control fleet that had just the normal CRM training. This design was quasi-experimental because 1) the pilots were assigned to fleets by normal airline rules and seniority rather than by random assignment, and 2) some pilots transitioned between the two fleets during the evaluation, resulting in some knowledge of the CRM procedures in both fleets.

The focal performance measures were crew performance in each fleet on the yearly recurrent LOEs and on random Line Checks. Since the changes in observed performance were expected to lag behind the formal implementation of these procedures, the design included three years of performance evaluations. Performance evaluations covered the year prior to formal implementation of the CRM procedures as a baseline, the implementation year, and the year following implementation to check for slowly-emerging effects of the training.

The research team recognized that a structured evaluation process was essential in order to achieve systematic and reliable observations and ratings of performance. Focal performance evaluation measures were selected to provide as complete a set of performance data as could be collected within the constraints of the regional operational environment.

LOE Evaluations

The LOE evaluation, conducted during the pilots' annual evaluation for flight certification, was a work sample performance evaluation with crews performing under a set of scripted conditions. The evaluator adhered to an LOE script to consistently introduce specific problems and distracting conditions. In this way, crew responses to abnormal and emergency situations could be assessed. The evaluation forms emphasized specific crew behaviors for these events, including both technical and CRM performance items. Since the LOE evaluations were on pre-announced months, as mandated by FAA requirements, the crew could prepare for the evaluation ahead of time and potentially be on their best behavior for the test situation.

LOE event sets. The administration of the actual LOE session was guided by a LOE script that specified all actions and conditions for the evaluator to set for the simulated flight. The script for the entire LOE session was divided into segments called “event sets”. Each event set specified the onset of a focal problem for the crew plus other auxiliary events.

The event set was developed as a refinement of the AQP concept of "event" (FAA, 1991) and was an integral part of training and evaluation. An event set is made up of an event trigger, supporting conditions, and distracters. The event trigger was the condition under which the event was fully activated and presented the crew with some problem or issue in a natural context. Supporting conditions were other events taking place within the event set designed to further CRM and technical training objectives and to increase event set realism. Finally, distracters were conditions inserted within the event set time frame that were designed to divert the crew's attention from other events or to increase workload. Most commonly, the LOE event sets corresponded to the usual segmentation of phases of flight. Because of evidence (Seamster, Edens, & Holt, 1995) that assessments made at the event set level result in more reliable assessments compared with the overall session assessments that are traditionally used at regional airlines, crew evaluations in this project were based on event sets.

LOE worksheets. Crew performance during the LOEs was assessed using structured worksheets. The evaluation of each LOE event set was guided by a corresponding LOE worksheet page. The LOE worksheet simplified what could be a relatively complex evaluation process, provided instructors with a tool for more reliable ratings, and helped the instructor deliver a more balanced debrief that covered the CRM as well as the technical elements of each event set.

The worksheets helped evaluators standardize the assessment of LOE sessions (ATA, 1994, Hamman, Seamster, Smith, & Lofaro, 1993). The LOE Worksheet had three critical elements that provided a clear structure to the assessment of both CRM and technical crew performance: specific observable behaviors or skills, a standard rating scale, and specific examples of performance for each level of rating. Using worksheets based on each event set allowed the Instructor/Evaluator (I/E) to concentrate on a limited range of observable behaviors and on specific CRM and technical training objectives for each

flight segment. Regular Instructor/Evaluator calibration training (discussed below) ensured the correct use of these worksheets.

Observable behaviors. Properly specified observable behaviors were the basis for the evaluation worksheets that guided the I/E's assessment and provided information for the debriefing process. There are a number of different approaches that can be used in the identification of observable behaviors. In this project, we used I/Es rankings to indicate where crews showed the most CRM problems in the recurrent training environment. We also identified observable behaviors and skills from the fleet-specific task analysis conducted under the Advanced Qualification Program.

The specification of observable behaviors was an important step because it affected the implementation of the LOE Worksheet and the ultimate outcome of CRM assessment (Seamster, Hamman, & Edens, 1995). These observable behaviors or skills were carefully identified and validated as being central to successful performance on each specific event set. These behaviors or skills provided a point of focus for the I/Es during the observation of the LOE and also during the debriefing. These focused observations were a basis for evaluating higher-level components of CRM and technical performance. Each evaluator synthesized the observed behavior and rated performance components on a standard rating scale.

Standard rating scale. The second critical element of the LOE Worksheet was the use of a standard rating scale. In this project, standardization was addressed in two different and equally important ways. First, we wanted the rating scale to be standard in that the same scale would be used throughout the crew assessment process. That is, we wanted a scale that could be used across the full range of evaluation environments from simulator sessions to line checks, and across a full range of types of items. Currently, many airlines use binary pass/fail scales to assess individual or crew performance, but use rating scales with five or more values to rate CRM performance. The research team felt that a standard scale can and should be used for as many of the different assessment instruments as possible. Using a standard rating scale can reduce the training time required to familiarize I/Es with different assessment instruments. The use of a standardized scale can also increase the amount of practice that I/Es have with that scale. Increased practice with a single scale should also lead to better assessment skills and ultimately

better rater reliability. At this airline, a four-point scale was initially developed for the LOE and later extended to Maneuver Validation and Line Check evaluations.

The second way a rating scale may be standard is through each scale point being based on a set of specified standards. Based on a job analysis, airlines should be able to specify the criteria for "standard performance" for a range of technical and CRM areas (also called Qualification Standards). For this airline, both technical and CRM performance ratings were based on a four-point scale covering the full range of possible crew performance: *unsatisfactory*, *satisfactory*, *standard*, and *above standard*. The labels and precise meanings of each scale point were the result of several cycles of discussion between the I/Es on the one hand, and the researchers on the other. These discussions illuminated the need for both good labels and definitions for each scale point. For example, the label *average* was discussed but discarded as being inherently ambiguous and not clearly tied to the qualification standards for a given item.

At the end of this development process, each of the four scale points had a qualitatively unique definition. An *unsatisfactory* performance rating meant that observed crew behavior did not meet minimal requirements. A *satisfactory* performance rating meant that crew performance met FAA standards but not airline standards. A *standard* performance rating meant that crew performance met airline standards. An *above standard* rating meant that observed crew behavior was markedly better than the *standard* performance in some important way.

Specific examples. During the first two years of evaluation, the I/Es discovered problems in their calibration during I/E recurrent training and requested more guidance in the evaluation of *unsatisfactory*, *satisfactory*, *standard*, and *above standard* performance. Ultimately, a variation of a Behaviorally-Anchored Rating Scale (BARS) was developed in which examples of each level of performance for each skill item were defined by training personnel (see Smith & Kendall, 1963). These performance levels included short descriptions of the qualification standard for *standard* performance as well as brief examples of *unsatisfactory*, *satisfactory*, and *above standard* performance. For ease of reference, these performance levels were printed on the back of the preceding page of the LOE Worksheets so that they would be immediately available for reference during the rating process.

Line Check Evaluations

The initial data collection plan included evaluation of Line Check data to see if the performance differences between the fleets would be observable in normal operating conditions. In contrast to the LOE evaluations, the Line Check evaluations were collected on randomly selected legs of routine commercial flights. There was no prior notification for line checks and the evaluation situation was natural. However, the limitation in this type of evaluation is that routine commercial flights almost always contain only normal operational conditions. The evaluation forms were therefore structured to emphasize normal conditions occurring on routine flights. Similar to the structuring of LOE items by Event Sets, the evaluation items on the Line Checks were structured by phase of flight to reduce evaluator workload.

Initially, an extremely detailed Line Check form was developed with many items for both Pilot Flying and Pilot Not Flying for each phase of flight. However, analyses showed a systematic pattern of missing data for this form that indicated high evaluator workload. The form was subsequently revised with fewer, more general questions, and missing data decreased to a more acceptable level.

The more general Line Check items could not precisely target the execution of specific CRM procedures. This illustrates an inherent tension in applied research that can occur between operational and research goals. The goal of the research team was to have precise evaluation of the training intervention whereas the goal of the airline was to have a global evaluation of the safety of line flights. These two goals dictated different sets of items that could not all be accommodated on the Line Check form without overloading the evaluator. Therefore, an auxiliary non-jeopardy evaluation (jumpseat observations) was designed to be conducted by an independent cadre of evaluators targeting the CRM procedures and expected effects of those procedures.

Jumpseat Observations

The jumpseat observation form was designed to target the occurrence and effects of proceduralized CRM in normal operations. These forms were used in non-jeopardy flight observations carried out by three grant team members and two airline training center staff, all of whom were pilots. All evaluators clarified the non-jeopardy nature of the event and requested voluntary cooperation from the crew before making observations.

The jumpseat observation form was organized by three global phases of flight: Departure, Cruise, and Arrival. The departure phase of flight included all performance from preparations at the gate to arriving at cruise altitude. The arrival phase of flight included all performance from the beginning of descent to engine shut down at the end of the flight. For each phase of flight, components of normal briefs required with proceduralized CRM were checked off, and the quality of crew performance on specific CRM procedures and related items was evaluated on a 5-point Likert scale. Finally, each evaluator made a summary evaluation of crew effectiveness for each phase of flight.

Instructor/Evaluator Survey

During an evaluator training session, the I/Es maintained that the differences in pilots from the proceduralized CRM fleet compared to the normal CRM fleet were most noticeable when training First Officers for an upgrade to Captain or when training pilots who were transitioning between fleets. This source of information about the effects of proceduralized CRM had not previously been considered. Consequently, an evaluation form that directly compared crews from these two fleets was developed and administered to all I/Es for both fleets. Results from the 19 I/Es who had trained pilots from both fleets indicated that the effects of proceduralized CRM centered on the three basic performance dimensions of communication, workload management, and planning and decision-making. In retrospect, the potential domain of information for comparatively evaluating pilots with and without the procedural CRM training should have been explored more thoroughly.

Pilot Survey

When we expanded our quest for more information on the effects of the training, the evaluations of the pilots themselves became an obvious source. Subsequently, a questionnaire was designed for pilots which targeted multiple possible effects of training as discussed by Kraiger, Ford, & Salas (1993). All pilots reported their attitudes toward CRM in general and proceduralized CRM in particular. In addition, the ACRM-trained pilots reported on their knowledge of the CRM procedures, practice of the CRM procedures in normal operations, and perceived effectiveness of the CRM procedures. The questionnaire was designed in a voluntary, mail-back format and was distributed to all airline pilots in the last year of the study.

Instructor/Evaluator Training

Instructor/Evaluator training was found to be the key to combining the LOE worksheets and other elements of the evaluation process into a working, structured assessment system. For this airline, we focused on both I/E introductory training and continued standardization training.

Initial Instructor/Evaluator Training

The initial training for the I/Es included detailed training on the specific LOE to be used in the evaluation process. This training included the observable behaviors assigned in the LOE, the forms that were used to identify the specific event sets, the assessment criteria for the event sets, and video examples of line crews (not scripted) flying the LOE. This practice supported the identification of the assigned behaviors and began the calibration process for the I/E. We found it critical that the I/E training provide the opportunity to watch video segments of event sets and discuss their ratings of the crew. We used a technique of assigning instructors a de-identified Pilot Identification Number (PIN). After the instructors watched a video segment of the LOE and made their assessments, their ratings were processed by a data collection program and the results immediately returned so that they could compare their performance against the other instructors. This technique provided real numbers for comparison and took the subjectivity out of the instructor calibration process.

In addition to the specific training for the LOE assessment, the I/E's training also included the skills required to brief, administer, and debrief the LOE. Each I/E was required to 1) fly the LOE 2) observe the administration of the LOE by a previously trained I/E, and 3) practice giving the LOE as an evaluator. The I/Es switched pilot flying and instructor roles during the session to allow them the opportunity to experience the various roles in this type of training. The I/E and the flying instructor pilot performed a facilitated debrief after each segment was completed. The non-flying I/E pilot initially began these debriefings with observations of the CRM skills that were exhibited or that should have been used during each event set. This experience was designed to enhance observation and identification skills, give practice using the LOE worksheets for evaluation, enhance the I/E's facilitation techniques by the sharing of concepts and techniques by the three pilots, and give practice in the type of crew-centered briefing/debriefing they would be conducting with normal line crews.

Continued Standardization Training

After the initial training, refresher courses and calibration sessions were established. The frequency of this training depended on calibration results, I/E turnover in the training center, perceived issues or problems with the evaluation process, and so forth. Standardization training took place at least twice a year, and more frequently when the reliability data suggested it was needed. This training day included feedback on the forms the I/E's were using, specific problems with consistently administering or evaluating the LOE, video samples with assessment exercises of event sets, and immediate instructor feedback on their ratings in comparison with other instructors. The regional airline working with this grant placed great emphasis on these improved training techniques to ensure a systematic and standardized evaluation process.

Summary of Experiences with the Regional Airline

The AQP that the airline had adopted for this fleet required the assessment of CRM with technical skills in the operational environment. The assessment process was initially focused on the LOE, which contained observable behaviors and technical skills, organized by event sets. This organization allowed for the development of tools (LOE Worksheets) to support the instructor in making fair and accurate assessments of crew performance. Assessment tools were accompanied by a comprehensive I/E training program that included:

1. training with video examples of line crews flying the LOE scenarios;
2. clear definitions of rating scales and standardization;
3. computerized methods for instructors to receive objective (data supported) feedback on their evaluation performance;
4. I/E simulator training programs for practice in flying and administering the scenarios; and
5. continuing training during the AQP cycle to monitor and maintain I/E calibration.

Although the initial focus was on the LOE, several auxiliary evaluation methods were developed to assess the effects of proceduralized CRM during the course of this project. Since each type of evaluator (I/Es, pilots, jumpseat observers) had a different viewpoint and each form of evaluation had a different set of strengths and weaknesses, it was important to assess the effects of any training program with a variety

of evaluators and methods. In our project, this approach led to a more complete evaluation of the resource management procedures which were developed, implemented, and evaluated at the regional airline.

EXPERIENCES WITH A MAJOR AIRLINE

At the time we developed the resource management program for the regional airline, a major airline expressed interest in evaluating some kind of procedural CRM as a part of their resource management program. However, this airline had a number of large fleets, and it did not seem feasible to completely re-write the abnormal procedures as we had done with the regional airline. Instead, we took some specific aspects of the CRM procedures developed in working with the regional and adapted them for a specific application at the major airline.

Program Development

Need for the Training Program

Although no formal needs assessment had been done, the major airline was generally dissatisfied with their current stand-alone CRM training and perceived a need for more effective tools to help pilots, especially in abnormal/emergency situations. This effort was lead by the quality assurance administrator for the training center. At this airline, the quality assurance administrator can take an active role in working with fleets to improve the effectiveness of training. This person acted as a liaison to the fleet which was most interested in improving pilot performance using a proceduralized CRM approach.

Background Assessment

Given the limited scope of this part of the project and limited resources available, no organizational climate survey was performed. For a major airline, fleets can be relatively independent and the climate within each individual fleet may differ from the overall organizational climate.

Identifying problem areas. Based on detailed discussion of critical incidents (NTSB, 1994) and ASRS reports, the airline had previously identified crew performance during abnormal or emergency situations as potential problem areas. The airline's initial goal was to reduce the reliance of the crew on memory for the correct execution of actions during the first two minutes of an abnormal or emergency by developing the QRC, which listed the actions to be taken in a checklist fashion. Subsequently, the research

team proposed augmenting the QRC with some aspects of proceduralized CRM to study its effectiveness with the pilots in the cooperating fleet.

One important lesson we learned right away at this airline was the critical nature of cooperation with the local FAA inspector in making any training changes. Our experience at the regional airline was that the local FAA inspector had been briefed on the project goals and was supportive of the training changes required to achieve these goals. In contrast, the local inspector for the major airline required objective evidence of the effectiveness of the QRC before considering its approval. This necessitated a controlled study in which 120 crews were randomly assigned to QRC or normal (non-QRC) conditions and evaluated on a non-jeopardy LOE. Although the results ultimately supported the use of the QRC in reducing crew errors during these situations, and the use of the QRC was finally approved, there was a delay of several months.

Development of Procedures for Abnormal/Emergency Situations

In developing the program for the major airline, the goal was to assess the effectiveness of developing a *general framework* that captured the resource management practices that had been incorporated into the regional's QRH procedures for abnormal or emergency situations. The general framework that was developed after an iterative cycle of development with the airline and the research team emphasized five basic steps or phases of crew activity for each abnormal or emergency situation. These steps captured the general approach to decision-making that had been embedded in the FOM and QRH at the regional airline. The first step was clearly assigning pilot flying/pilot not flying (PF/PNF) duties. The second step was a careful and complete diagnosis of the problem. The third step was a complete situation assessment. The fourth step was planning and decision making. The fifth step was plan execution and monitoring.

At a detailed level, each general step had specific sub-elements that should be evaluated or performed. For example, in the planning and decision making step, the crew was advised to develop "Bottom Lines" and "Backup Plans". A Bottom Line specifies the lower limits for some aspect of the situation, at which point the current plan is modified or discarded. A Backup Plan is a pre-arranged alternate version of the current plan that is enacted when a Bottom Line is reached. For example, a

Bottom Line may specify 30 minutes of fuel remaining before diverting to an alternate airport. The exact alternate airport and necessary approach procedures constitute the Backup Plan.

To ensure cross-briefing, situation awareness, and full use of the viewpoints of each crew member, we developed a “Brief, Advocate, and Resolve” (BAR) cycle. In the BAR cycle, the pilot not flying (PNF) was assigned to the main duties in the problem diagnosis, situation assessment, and planning steps. At the end of each major step, the PNF was required to brief the pilot flying (PF). The PF was then required to critically evaluate the briefing contents and advocate any different ideas or views about the problem, situation, or plan. Finally, these different ideas or views were to be resolved so that all crewmembers were in agreement before proceeding to the next step.

To ensure consistency with airline operations and pilot expectations, the general framework from the regional was revised into a CRM checklist and integrated with the major airline’s training philosophy, CRM principles, and nomenclature. Each step in the cycle became a checklist item. Detailed points for each step were designed to give maximum guidance in performing each step and to provide best practices. For example, in the situation assessment step, crews were advised to consult all relevant sources of information such as ATC, dispatch, maintenance, flight attendants, etc. Clearly, consulting all possible sources of information may take an unrealistically long time. Therefore, the PNF was instructed to complete the five basic steps with as many of the detailed sub-elements as they could complete in the available time.

Evaluation of Procedures for Abnormal/Emergency situations

Evaluation Design

This phase of the project is still in progress. The current design for evaluating the abnormal/emergency CRM procedures is a completely randomized field experiment. Crews in training will volunteer for participation and be randomly assigned to one of three groups. The control group will be trained in the current airline materials and procedures for emergency/abnormal situations. The second group will be trained in generic CRM procedures for these situations and use a corresponding stand-alone CRM checklist. The third group will be trained in generic CRM procedures for these situations and use a set of CRM procedures integrated into the abnormal/emergency sections of the FOM.

Development of Evaluation Materials

Crews will be evaluated on one of three different LOEs in this field experiment. The LOEs represent three completely different abnormal/emergency problems during the cruise phase of the scenario, each requiring use of a different section of the FOM. The advantage of using LOEs with different content for different crews is that they cannot anticipate the content of the LOE with any certainty. This should result in more natural crew reactions to the problem situations. However, there is a disadvantage associated with using LOEs with different content; it is possible that the content of the LOE may interact with the training manipulation in unique ways and make it more difficult to determine if there is a general effect of the training and procedures.

Crews will be evaluated on their LOEs using the normal event set worksheets used by this airline. These sheets will be supplemented by an additional evaluation sheet focusing on whether the crew performed specific CRM procedural elements when the critical trigger event occurred. The additional evaluation page will have major items focusing on 1) division of PF/PNF duties, 2) problem diagnosis, 3) situation assessment, 4) planning and decision making, and 5) plan execution and monitoring. Appropriate subsidiary items will be included in each area. For example, in the planning and decision making section, detail items will evaluate whether pilots established Bottom Lines and developed explicit Backup Plans.

Subjective reactions to the training program will also be collected. A brief mail-back questionnaire will be given to pilots to evaluate the effectiveness of their training materials. The pilots' evaluations of the materials in the control and procedural CRM groups will be compared to assess which type of training is perceived as more effective.

Summary of Experiences with the Major Airline

The cooperating fleet at this airline is also an AQP fleet and has implemented LOE evaluation, LOE worksheets, and rater calibration training in a similar fashion to the regional airline. The scope of the proceduralized CRM intervention is much narrower at this airline, targeting only crew teamwork performance after an abnormal or emergency situation has occurred. Subsequent airline implementation of

any form of proceduralized CRM will depend on the results of this evaluation as well as the local FAA and other major stakeholders.

LESSONS LEARNED IN DEVELOPING RESOURCE MANAGEMENT PROGRAMS IN AVIATION

From our experiences with these two airlines, we have learned a number of lessons that are critical to the success of a resource management program within the airline environment.

Lesson 1: Assess Organizational Commitment

From experiences with these airlines, it is evident that a resource management program must be implemented as an ongoing process that requires the involvement of the entire organization. This means that a commitment must be made throughout different levels and different sections within the organization.

Corporate Commitment

This commitment must start with the senior management. They must support the changes, and make this clear to the organization through actual behavior. If senior management says that it values the application of good resource management, then they must honor this commitment. As an example, single-engine taxiing saves money, but results in multiple tasks and added distractions during the taxi phase. These add to the crew's workload and make safety-critical CRM procedures such as scanning the area for traffic and thoroughly briefing the conditions for the departure more difficult. If the organization says it values good CRM but requires single-engine taxiing to save money, it is not honoring its commitment.

In practice, management policies are not always implemented using consistent rewards and punishments. However, this is critical to the success of resource management training. Without appropriate rewards for participating in the resource management program, it will be difficult to convince the "rank and file" to adopt the procedures and practices trained in the program.

The perceived need for resource management training often originates from the training department, so it should not be difficult to get a commitment from the training department. However, at a large airline, the training department can have diverse sections with different goals and agendas. In such a

case, all sections of the training department must cooperate to make the training work. Further, links should be established and maintained between the training development team and groups representing the customer fleets, the flight standards or flight safety department, and the quality assurance or evaluation department.

Importance of a Champion

Due to the financial considerations involved in establishing and maintaining a resource management training program, it is important that the program have an internal champion who will fight for necessary resources for the program. It is helpful if this individual is well-positioned in the company, both in terms of having financial resources available and in terms of commanding respect from personnel within the company.

Resource Commitment

Relevant resources always include money and time of relevant personnel, but may also include scarce resources such as simulator time. The procedures development process itself can be accomplished with a relatively small budget, but associated elements such as changes to operations documents and training can result in substantial expenses. The management structure of the organization must be considered in obtaining resources. Obtaining resources in a very hierarchical or balkanized organization will require more time and effort than a flat, consolidated organization.

More resources may be available at larger airlines, but the costs of implementation may also be correspondingly greater. For the major airlines, one change in a procedures can result in the need to re-print one or more pages for 40,000 manuals. This can become a very large cost, and those that do not fully support or understand the need for CRM may use that as an argument to limit or terminate the program. A strong case can be made that procedural changes can be integrated into a scheduled manual update for little or no additional cost. Thus, with careful planning, CRM procedures can be developed and established efficiently by reducing or eliminating some of the larger expenses.

Additional considerations for effective use of personnel resources include integrating some of the team members' development time with their normal responsibilities. However, new activities should not be assigned without providing the appropriate hours to do them. The best solution is to look for ways to

replace some of the existing activities with the CRM procedures development work. It is important to keep the development process on schedule and anticipate and reduce delays that can increase costs in time and resources.

Lesson 2: Background Assessment

Organizational Climate Survey

Conducting an organizational survey during the early stages of training development can help identify organizational elements that need to be involved in the process. For this project, for example, early inclusion of fleet representatives may have facilitated later development steps. The survey can show which departments or fleets understand and want to be involved with the training program and which will require more communication and information. In addition, the organizational survey can be used to increase the entire organization's awareness of the new approach to resource management.

The survey development and administration process can be used as a way to get essential parts of the organization introduced to the new program. Asking key individuals from different departments to provide material for the survey or having them review specific survey items will improve the quality of the survey and make those departments aware of the new program. The survey should address issues and concerns of the flight attendants, dispatchers, maintenance, and other key parts of the operation.

Some of the items that should be considered for inclusion in an organization survey are: airline safety climate; communication and cooperation; departmental management and structure; job responsibilities and standards; organizational management; professionalism and job performance; and quality and frequency of training. Any severe problems indicated by the organizational survey may have to be addressed before training innovations can be attempted with a reasonable expectation of acceptance and cooperation. For example, if one fleet or section of the training department is dead set against the principles and approaches of the proposed training, that issue must be addressed before proceeding with development.

Lesson 3: Program Development

Needs Assessment

A good resource management program will be based on a needs analysis based on data from the other airlines and the airline itself. The needs analysis is used to identify the most important crew performance problems and then determine what form of intervention will best correct those problems. For airlines that do not have data pointing to specific problems, it is useful to review industry-wide reports to learn where other airlines have crew performance problems. Finally, informal or qualitative data from the training or evaluation departments within the airline can also be used to identify airline-specific problems, as was the case at the major airline in this project. The preceding activities will provide sufficient data to pinpoint the CRM problem areas and provide possible causes.

Prioritizing Training Goals

Based on the needs analysis, the development team should choose those problems that point to the need for one or more CRM procedures. The primary determination to be made is whether the performance problem would be resolved best by training, equipment redesign, new or modified procedures, or some combination of these. For example, at the major airline the CRM checklist was implemented in part by a new or modified checklist, and in part by appropriate training in best practices for each major CRM step. A airline's philosophy and policy must be accurately understood to determine whether a procedural solution is the right one for that specific operation and organizational climate. The development team must always be sensitive to the fact that different departments may have different local versions of the airline's philosophy and that these differences may have to be reconciled before development work can continue.

Lesson 4: Evaluating the Training

As the training is being developed, the appropriate evaluation materials should also be developed and tested. At each airline in this project, the development of the best format for the LOE worksheet required several cycles. Similarly, the LOE content and script may need to be refined so that the evaluation situation requires specific resource management skills and behavior. One of the most difficult parts of completing the LOE worksheet is correctly specifying observable behaviors or skills.

Identifying Observable Behaviors

Specifications that have more than one behavior per observation (such as global observations previously used in research) can make that observable behavior more difficult to check than a single behavior. Concise and simple wording is recommended, and the verb should refer to a clearly observable behavior. Ultimately, the observable behavior has to be worded so that the items can be understood and used by I/Es, who often experience high workload in the assessment process. For example, "Captain coordinated flightdeck activities to establish proper balance between command authority and crew member participation, and acts decisively when the situation requires" which was used on one assessment form, can confuse the evaluator because there are so many elements to consider. If the airline wants to concentrate on the assessment of Captain's ability to assign cockpit activities, a single observable behavior might be stated as, "Captain assigns tasks effectively." Therefore, it is important to use I/Es to review and validate the observable behaviors. Recommendations for specifying observable behaviors include:

(Seamster, Hamman, & Edens, 1995):

- careful examination of the existing task analyses and related flight and training manuals to ensure that they include all the key performance elements, especially CRM elements;
- use of I/Es to identify training issues, to determine possible relevant behaviors, and to validate the observable behaviors;
- specification of observable behaviors as simple statements of actions that can be clearly detected by I/Es who may be under a relatively heavy workload;
- specification of the relevant interpersonal or technical activities or actions so that I/Es are not asked to observe mental activities;
- organization of observable behaviors on worksheets for each event set to produce more reliable assessments and more systematic debriefings; and
- validation of the assigned observable behaviors by having the instructors fly the event sets and assess the CRM behaviors chosen.

Associated Documentation

In addition to the procedures developed, there may be a need to develop documentation that helps the pilots learn and implement the new procedures. For example, in this project a refresher videotape was

designed to help pilots recall the correct implementation of the new procedures. Similarly, official documentation such as the QRH and FOM may have to be revised for consistency with the new procedures. Modifying official documentation requires FAA approval that may, in turn, require additional time or revisions.

Lesson 5: Report the Program Assessment/ Performance Data

Reporting data can either be focused on the effects of a specific training intervention or more broadly focused on an overall picture of crew performance. For this project, the research team goals of evaluating the proceduralized CRM training required a focused data analysis and report. In contrast, the airline goals for analyzing the data included the main goal of a broad picture of crew performance on different assessments as well as the more minor goal of evaluation of training effectiveness. A more accurate picture of crew performance can be an extremely valuable result of improving the quality of evaluation.

Timely, accurate, and relevant data reporting is one of the most important byproducts of this kind of program. Using these improved assessment methods, it is possible to collect a large amount of accurate, reliable, and valid performance data concerning both the technical and teamwork aspects of crews. In addition, tools (see Holt, Boehm-Davis, & Beaubien, this volume) allow airlines to analyze different dimensions of I/E reliability and ensure calibrated evaluations. However, the large amount of data combined with different possible forms of analysis can result in an overwhelming amount of information, so the challenge is to report the essential data in a usable format to the appropriate organizational elements.

There are four main organizational elements to consider for data reporting, and each of these stakeholders may have different information needs:

1. Crews and union representatives
2. Instructor/evaluators
3. Training department management and quality assurance
4. Fleet managers and other relevant fleet personnel

For example, a report designed for fleet managers might emphasize some of the following within-fleet data: distribution of overall performance by position (Capt., F/O, S/O); distribution of technical performance by position; distribution of CRM performance by position; maneuver validation ratings by type of maneuver; distribution of event set performance where there are problems; and overall crew performance trends (quarterly and annual).

The training department management, on the other hand, may also be interested in comparisons across fleets. Across fleets, interesting data might include: distribution of overall performance by fleet; distribution of technical performance by fleet; distribution of CRM performance by fleet; maneuver validation ratings by type of maneuver and by fleet; event set performance by fleet where there are identical event sets; and overall crew performance trends by fleet.

CONCLUSIONS

Our experiences with two airlines suggest that an organization can direct improvements in resource management through the development of and training in management procedures customized to its own operation. Over the past 20 years, resource management has evolved from an emphasis on attitudes and personalities to a focus on specific resource management skills. The procedural approach outlined in this chapter complements that skill-based approach allowing organizations to move beyond the elaboration of general principles and policies to specific procedures designed to improve worker performance. This procedural approach to resource management facilitates a more focused training and assessment program that allows organizations to analyze the effect of the new procedures on individual and team performance. When wisely used, these tools can produce a large quantity of high-quality performance information, which can be used by the organization in many different ways. The challenge remains to isolate the essential data and represent it in the most meaningful ways.

Acknowledgments

This research was supported by the Office of the Chief Scientific and Technical Advisor for Human Factors (AAR-100) at the Federal Aviation Administration through grant 94-G-034 to George Mason University. We thank Dr. Eleana Edens at AAR-100 and Dr. Thomas Longridge at AFS-230 for their continuing support of this work.

REFERENCES

- ATA (1994). *Line Operational Simulations: LOFT Scenario Design, Conduct and Validation*. LOFT Design Focus Group, AQP Subcommittee Report, Air Transport Association, November 2, 1994.
- FAA (1991). *Advisory Circular 120-54: Advanced Qualification Program*. Washington, DC: Federal Aviation Administration.
- GMU (in preparation). *Final Report on Evaluating Resource Management Training at a Regional Air Carrier*. Technical Report, George Mason University.
- Hamman, W. R., Seamster, T. L., Smith, K. M., & Lofaro, R. J. (1993). The future of LOFT scenario design and validation. *Proceedings of the Seventh International Symposium on Aviation Psychology*, 589-593. Columbus, OH: The Ohio State University.
- Holt, R. W., Boehm-Davis, D. A., & Beaubien, J. M. (this volume). Evaluating resource management training. In E. Salas, C. Bowers & E. Edens (Ed.), *Applying resource management in organizations: A guide for training professionals*. Mahwah, NJ: Lawrence Erlbaum Assoc.
- Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based and affective theories of learning outcomes to new methods in training evaluation. *Journal of Applied Psychology*, 78, 311-328.
- Mudge, R. W. (1993). Pilot judgement - and the management system. *Proceedings of the Seventh International Symposium on Aviation Psychology*, 216-220. Columbus, OH: The Ohio State University.
- NTSB (1994). *Safety Study: A Review of Flightcrew-Involved, Major Accidents of U.S. Air Carriers, 1978 through 1990* (PB94-917001 NTSB.SS-94/01). Washington DC: National Transportation Safety Board.
- Seamster, T. L., Edens, E. S., & Holt, R. W. (1995). Scenario event sets and the reliability of CRM assessment. *Proceedings of the Eighth International Symposium on Aviation Psychology*, 613-618. Columbus, OH: The Ohio State University.
- Seamster, T. L., Hamman, W. R., & Edens, E. S. (1995). Specification of observable behaviors within LOE/LOFT event sets. In R. S. Jensen (Ed.), *Proceedings of the Eighth International Symposium on Aviation Psychology*, 663-668. Columbus: The Ohio State University.

Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 1963, 47, 149-155.