

## **Air Carrier Training Research**

The FAA Office of the Chief Scientist for Human Factors (AAR-100) directs an air carrier training research program centering methods for effective pilot training and assessment. Semi-annual research reviews are conducted by this office. The following report summarizes the February 1998 review.

### **Introduction**

Although basic technical and Crew Resource Management (CRM) concepts are widely accepted, much remains to be learned regarding the appropriate methods for effective training and valid and reliable assessment of training programs. The general research philosophy guiding efforts to improve training and assessment is that research must consider distinct segments of aviation training systems. Individuals comprising the crew, instructors who train and evaluate crews in the classroom, the simulator and on the line as well as the management culture responsible for the safety climate of the carriers should all be considered. Additionally, this research must regard the variables important to Line Oriented Flight Training (LOFT) development, implementation and evaluation. Thus this research centers on (1) crew training and assessment, (2) instructor training, (3) LOFT development strategies (4) and (4) organizational and systematic influences on pilot performance.

Address questions or comments to:

Eleana Edens, Ph.D.  
AAR-100  
(703) 960-1085  
Eedens@mail.FAA.GOV

**Project Title:** *Identification of Hard to Train CRM Skills*

**Performing Agency:** University of New Mexico, Albuquerque

**Background:** Traditionally, CRM training has focused on the high level CRM components or elements of pilot performance. The training and assessment of high-level concepts is fraught with many problems. Among them, the validity of course content is questionable and the reliability and validity of CRM assessment has not been established. Prior research has shown that assessment of observable behaviors, rather than global concepts, especially within simulator scenarios event sets, can improve the reliability of CRM assessment. Therefore, observable behaviors have provided a more meaningful and lower level of detail for the CRM assessment process. However, the research and operational communities have not yet developed a comparable level of detail for the CRM training process. It is impractical to try to train observable behaviors. Thus, at present, it is likely that training would be on the global level but assessment could take place at the observable behavior level. This training paradigm, although improved from all CRM training activities being at the global level, still lack the training vigor employed, for example, in technical training. An approach to solving this dilemma is to attempt to identify the knowledge, skills and abilities underlying CRM concepts and then to train at that level. Observable behaviors could be identified and assessed for skills in situation specific training events.

In addition to the need for refined skill analysis methods, reliable data collection systems must be developed in order that carriers and the FAA can evaluate the effectiveness of training programs.

**Project Description:** This research project has been investigating two distinct segments of air carrier training, skill analysis and reliable performance assessment.

Work has been conducted at major carriers to identify key CRM skills that need to be trained in transition and recurrent training. The main objective of the research is to provide carriers with set of steps to identify and link CRM skills to essential elements of crew training so that those skills can be systematically trained and assessed. This research collected rating and agreement index data to help representatives from different carrier departments form a consensus in identifying CRM skills. It is significant that Advanced Qualification Program (AQP) and Human Factors Subject Matter Experts (SMEs) as well as check airmen participated in the process, providing the CRM skill identification with a substantially broader base than is provided by traditional ISD methods. Using a greater number of SMEs for each of the steps allowed for a broader consensus and better understanding of CRM skills across the organization while distributing the workload over a larger number of individuals. This process reduced some of the bottlenecks encountered in traditional carrier skill analyses where one or two SMEs are asked to perform most of the analysis. This has resulted in a CRM skill listing and a set of methods that carriers can use to identify their own critical CRM skills. An effort is under way to develop a CRM skill Identification Manual that will summarize the process for other air carriers.

The results of this CRM skills identification will be combined with analyses from previous research which focused on LOFT/LOE session debriefings, decision making and situational awareness skills. The sets of CRM skills, derived from four major carriers, will be used to develop a CRM curriculum outline applicable across air carriers. The CRM curriculum outline will include the basic CRM skill requirements, the advanced CRM skill requirements, the key knowledge components of those advanced skills. This will be developed in FY 98 and distributed in FY 99.

Another component of this research project focuses on pilot assessment. Efforts to reliably assess pilot performance, during training, have traditionally been less than successful. During FY96 and FY 97 this research team has been working with an air carrier to improve training assessment by designing rater training and new data collection systems. This work focuses on rater calibration, development of grade sheets, development of videos used in the calibration sessions, presentations to the instructors on the importance of quality data, collection and analysis of the calibration data, reporting results. Several calibration studies have been completed and additional sessions are planned. Thus far, the results of these calibration studies show (1) how minor changes in the wording of observed behaviors listed on gradesheets can have a significant effect on ratings, (2) large individual differences among raters exists, (3) higher inter-referent than inter-rater reliability and 4) a significant increase in reliability when improvements to the calibration sessions were implemented (wording of the gradesheet, quality of video, better briefing of the event sets). The frequency of the higher inter-referent reliability is important to note because there are differences among carriers in their views of whether a referent should be established and instructors trained to the referent or whether instructor's inter-rater reliability is adequate. AAR-100 is investigating this question further by work being done by George Mason University at several other carriers employing the inter-rater reliability methodology.

Based on the observations of this process at the major carrier, this team of researchers has developed a software package to be used in group calibration sessions. This software will allow carriers to more effectively train and assess evaluators in their ability to judge Line Oriented Evaluations (LOE) performance. This software facilitates the data collection, statistical data analysis, and report generation for evaluator LOE calibration sessions. A group of evaluators individually rate the performance of a crew shown on a video flying an LOE. After the ratings are entered on a standard gradesheet they are collected and the scores are entered into a PC running Access software. After entering the evaluators ratings a complete set of statistical analyses of the data is provided. These reports summarize each evaluators performance along with various group performance statistics. These reports then serve as the basis for the calibration feedback, debriefing, and further training of the evaluators.

Although carriers conduct regular and recurrent rater calibration sessions, using the methodologies developed by this project, it is often necessary to recalibrate one instructor. The cost of bringing together a group of instructors in order to conduct a group session to calibrate a single instructor is prohibitive. This research team has developed a prototype PC-based system that effectively presents an individual calibration session.

This software tool is designed to allow individual evaluators to undergo the same type of training and calibration as discussed above in the group sessions but to self-calibrate individually at their own convenience. The system provides the prebriefing instructions, regarding the specific flight scenarios that would be presented and some general instructions regarding the use of the software. An audio visual presentation of a crew flying the LOE is presented along with the gradesheet. The evaluator enters ratings by simply clicking the appropriate buttons displayed. After entering all of the ratings for all events, the data are analyzed and the results presented in a form that is meaningful to the evaluator (e.g., how close the ratings were to referent grads, how the ratings compared with other evaluators, which ratings were most deviant from a standard, etc.) This tool can be used either as a training or evaluation tool. Its most obvious strength as a training system is that it allows evaluators to obtain training at their convenience, receive immediate feedback on their performance, and to have their grades linked to fleet qualification standards. Its greatest strength as an evaluation tool is that it allows personalized assessment. Assuming that evaluator grading history is stored in an Evaluator Performance Proficiency

Database, it is possible to determine whether an evaluator had specific weaknesses allowing the carrier to address this issue.

This research team has also produced a document describing what it means to have quality data, why quality data are necessary for training under AQP, and the methods for achieving quality in performance assessment. This document is intended for evaluators, and instructor/evaluator supervisors. It is currently under AAR-100 review and will be distributed 6/98.

Additional work being initiated by this research team will focus on developing a Performance Proficiency Database for Instructors and Evaluators as required in full implementation of AQP. It is necessary to have data that reflect the effectiveness as instructors (i.e., the performance of their former students on various measures, extending beyond the specific classroom evaluation conducted by the instructor). It is necessary to track calibration performance over time, as well as any trends in LOE and first-look performance evaluations. This team will develop the capability of linking pilot performance data to the Instructor Proficiency Data base. It is expected that this segment of this research will be completed early in FY99.

**Products:** Skill Analysis methods, instructor/evaluator calibration software, performance data analysis tools.

**Title:** *Pilot training and Evaluation: Airplane Simulation Human Factors*

**Performing agency:** Volpe Center, Cambridge

**Background:** Airplane simulators are critical for the training and evaluation of pilots because they eliminate training accidents and enable scenario-based training. In addition, the recently enacted requirement that regional airlines operate under the same rules as major airlines as well as the Advanced Qualification Program (AQP) training approach nearly obliges operators to use simulators. However, for regional airlines, cost continues to be an obstacle to flight simulator access, particularly for recurrent training.

Currently, airplane FTD and simulator qualification criteria are contained in Advisory Circulars AC 120-40B and 45B/C and related documents. Changes both in the air carrier training environment and in the field of simulator technology demand a reassessment of these criteria. The shift in training philosophy from a time-based list of specific training events to the Advanced Qualification Program (AQP) training approach based on individual training needs requires that much of the training formerly performed in aircraft will move to the simulator. Considering the initiative that Part 135 carriers be held to the same safety standards as Part 121 carriers, the success of AQP depends on whether regional carriers will be able to afford simulators that are appropriate for training 121 proficiency using an AQP training approach.

In the past, the goal in simulator manufacturing was to produce simulators that were engineered as similarly as possible to the airplane, resulting in the highest possible physical fidelity. Today, however, engineering capabilities have progressed in such a manner that not always the highest technical capabilities need be applied to all systems of the simulator, but high fidelity in one system may compensate for lower fidelity in another. The goal, thus, has shifted from highest possible physical fidelity to highest possible perceptual fidelity. That is, a good simulator should feel like the airplane and most important, elicit the same performance and behavior from pilots as the airplane. The purpose of this project is to increase the availability and affordability of airplane simulators by determining exactly what level of perceptual fidelity, applied to the most costly aspects of simulator, is required to maintain or improve the level of safety in airline pilot recurrent training and evaluation.

**Project Description:** This project is proceeding in three directions, (1) the collection of Subject Matter Expert opinion, (2) a literature review and (3) new research. First, two workshops were conducted with experts from the FAA, industry and academia. The topic of the first workshop was the aeromodel used to program simulators. The topic of the second workshop was the motion requirements for simulators. These workshops led to proposed changes to the regulations for Level B simulators and gave direction to subsequent research (AC 120-40B)

The second component of this project is an extensive literature review of simulation research. The information has been organized into a large database, which is updated regularly to incorporate new research results.

The third component is the implementation of new research to address the question of the motion requirement for recurrent training. Both the motion workshop and the literature review showed that although motion is presently required for a level B simulator there is no scientific basis for such a requirement despite much research in this area. Given the advances made in simulating visual scenes subsequent to the earlier research, it may now be possible to achieve sufficient perceptual fidelity of

motion with the vision system alone. As counterintuitive as this may seem, one must consider that no amount of technology can erase the fact that the simulator will not take off. Thus, the motion can not be perfect and in fact may create false cues. In contrast, the visual cuing can be close to perfect. Not requiring physical motion would represent a significant cost saving both in acquisition and maintenance of simulators used in recurrent training.

Research is underway at a regional carrier to determine if given a high quality visual system, is a high quality motion platform necessary for recurrent training and evaluation. Data collection is expected to be completed in FY 98.

**Product:** Research report.

**Title:** *Rapidly Reconfigurable Event Sets*

**Performing agency:** Naval Air Warfare Center Training Systems Design and the University of Central Florida

**Background:** Line-Oriented Evaluations (LOEs) are a methodology used in Advanced Qualification Programs (AQPs) to evaluate pilot training performance and establish trainee proficiency. LOEs consist of flight simulation scenarios that are developed by the training organization and approved by the FAA.

In the past, LOEs were developed and appraised individually . That is, each LOE was separately conceived, developed and tested by the training organization and individually reviewed and approved by the FAA. The development and approval of new LOEs was costly and time-consuming. As a result training organizations usually had only a limited number of LOEs available for evaluation, each of which were only approved for a limited time period. In some cases, this may have increased the risk of LOE scenarios being compromised within a training organization, thus reducing the validity and reliability of the scenarios.

One way to reduce the negative aspects of the previous LOE design process is to develop a set of LOE event sets (modules) which upon FAA approval could be assembled by the training organization into a number of unique LOEs, without requiring additional specific approval of each individual LOE by the FAA.

The availability of such a system will not only significantly reduce the time and cost required by training organizations to create valid scenarios for training and evaluation, but will increase the diagnosticity of scenarios by directly linking the event sets with CRM and technical training objectives. Further it will increase the validity of the evaluative process by giving evaluators detailed guidelines and materials for the administration of scenarios. This system will increase the opportunities for air carriers to create valid training and evaluation scenarios for global aviation operations.

This research team is developing the appropriate software to allow the reconfiguration of separate event sets into entire LOEs. The requirements document has been written and approved. Identification of a database structure for the software system has been developed. Creation of databases for aircraft technical data, event sets and aircraft differences has been established. This project is expected to be completed in FY 2000.

**Product:** Software for reconfiguring event sets.

**Project Title:** *Identifying Leadership and Followership Skills in CRM*

**Performing Agency:** Western Michigan University

**Background:** Historically, leadership has been viewed as a linear, one-way (i.e., downward) action with the primary aim of task accomplishment. This perspective has dominated the cockpit of air carriers. However, as roles in the modern cockpit change, from an individual performance to a crew performance focus, the skills required of the crew members change also. It is now thought that leadership is an activity which involves both leaders and followers as they interact to accomplish goals and that followership skills are as important as leadership skills to the safe and efficient performance of flight crews. It has been unclear how to train leadership/followership as the skills underlying these concepts had not been identified.

**Project description:** This project has analyzed cockpit performance in order to determine the behavioral components of leadership/followership skills.

A model of cockpit leadership was developed which guided the subsequent skill analysis. This model of effective and efficient flight operations points out that the crew must use many skills to deal effectively with operational variables largely out of their control. These variables include the regulatory environment, the corporate environment, the market environment, and the physical environment. Skill analyses were conducted based on the assumptions that (1) a minimum amount of leadership/followership is required to achieve a safe flight (2) individuals constantly oscillate between leader and follower roles (3) weak leader or follower skills will be compensated for by other crew member. Preliminary analysis are complete and determined that in order to deal effectively within the framework described in the model crews must possess skills in: envisioning; modeling; receptiveness; influence; adaptability; initiative. The analysis also showed that these skills are common to both leadership and followership.

Data from line audits with major carriers reveal that the components described in the model are related to overall crew effectiveness, thus validating the skill analysis data. The line audit data showed that, in general, the more severe an abnormal situation the less effective the crew tends to be in exhibiting leadership/followership skills. Captains tend to be less likely to articulate a vision for the flight, meet company standards or obtain commitment from other crew members or be adaptable. Both Captains and First Officers were less likely to initiate actions in response to an operational deficiency when the severity of abnormal was high. When handling a complex situation, but not an abnormal, these crews displayed "outstanding" leadership/followership skills. These ratings were not dependent on position or time together as a crew. This suggests that present leadership training may not be meeting its objectives especially when the crews are faced with an abnormal situation.

The data also showed that Captain do in fact, set the tone in the cockpit. When Captains articulated a vision for the flight the First Officer would initiate a response to an operational deficiency without further direction from the Captain. When Captains exhibited good conduct and high standards, First Officers exhibited similar behavior. When Captains were receptive, First Officers were likewise receptive. This finding supports the need for appropriate pre-flight briefing.

Maintaining vigilance during the flight sees to be dependent on each crew member at different phases. During predeparture, takeoff, climb, and cruise, vigilance was related to the Captain's envisioning, modeling and receptiveness. However, during the descent and approach phases, vigilance was related to the First Officer's conduct and standards.

Workload and task distribution was dependent on the Captain's conduct, standards, and receptiveness. Establishing guidelines for automated systems for all phases of flight was related to the Captain's articulation a vision, conduct and standards (modeling), receptiveness and use of interpersonal skills to obtain commitment from others.

Based on the skill analysis model and this line data, requirements for leadership/followership training curriculum have been identified. This includes the development of classroom exercises extending beyond role-playing, the creation of event sets that address critical leadership/followership skills and addressing company philosophy and policy issues. A complete but preliminary training package was developed for a major carrier. This includes instructor manual, student manual, and training videos. This package will be distributed to industry 5/98.

The current work focuses on extending the training curriculum development and on integrating leadership/followership training into the second generation Model AQP which will integrate leadership training throughout the life span of a pilot. This integration will be validated at a major carrier in FY99.

**Products:** Leadership/Followership training guidelines, and training systems.

**Project Title:** *Analysis of Cockpit Management System in Multiple Carrier Environments*

**Performing Agency:** George Mason University

**Background:** Based on previous research, it is widely believed that CRM training can result in significant improvements in flightcrew performance and flight safety. Historically, it has been possible to assess the effects of CRM training on pilot *attitudes* regarding the general management of cockpit duties with self-report inventories administered pre and post-CRM training. It is thought that a change in pilot attitude will result in a change in behavior in the cockpit. It has also been possible to subjectively evaluate the effects of CRM training on pilot performance in LOFT and on the Line by noting the differences in CRM performance between crews that have had CRM training and the crews that have not. Further, performance differences between air carriers and between fleets within one air carrier have been investigated. While this approach to assessment has contributed to the advancement of CRM, it is hampered by the practical constraints of self report and global measurement.

**Project description:** This present research project will help determine the most efficient method to train and assess CRM and whether conceivable performance differences pre- and post- CRM training can be empirically substantiated. This research project is being conducted with a regional carrier and plans to transition the methodology to a major carrier in a later phase. This project is designing and implementing a prototype CRM training program based on a task analysis methodology. The research will evaluate the effects of this specific CRM training on the Line, LOFT, and LOE performance of crews. This study is also integrating CRM performance requirements or procedures into the standard operating procedures of the air carrier. CRM procedures are the implementation of specific calls, checks, and/or guidance into one or more of the following: normal checklists, Quick Reference Handbook, Abnormal /Emergency Procedures, Flight Standards Manual, additional Job aids. This can be viewed as translating critical CRM principles into CRM procedures.

In developing the prototype advanced CRM course at the regional carrier, data from line operations, instructor comments and ratings, along with findings from the NTSB commuter safety study were used as the basis for the CRM training design. This data was incorporated into a proceduralized management system specifically tailored to the needs of this regional carrier. This advanced CRM course has been given to pilots in the research fleet. Their performance in a LOE was assessed prior to this training and is currently being assessed after the training course was completed with an appropriate time lapse. The central focus of this research is to determine if a proceduralized system will increase pilot performance, not only that pilots will use the training and system but does using this approach to CRM actually increase pilot performance and thereby increase safety. This research has thus far proceduralized: Team Management; Crew Communication, Decision Making, and Situation Awareness in the operational environment of the regional carrier.

The proceduralized CRM training course is complete, and has been given to flight crews. The Quick Reference Handbook and the normal procedures have been rewritten to reflect proceduralized CRM.

Data analysis of the base line performance data showed some interesting results regarding instructor/evaluator ratings. In the research process it was necessary to train evaluators to collect data during the LOEs. The evaluators are line pilots working in the training center. One of the critical questions centered on whether the evaluator training was successful and would result in reliable and valid CRM assessment. The analysis showed that the LOE assessment process, which was taught in the evaluator training, going from ratings of observable behaviors to technical and CRM ratings, to overall

Pilot in command (PIC), Second in command (SIC) and Crew ratings for each event set was confirmed. Observable behavior ratings significantly contributed to both technical and CRM evaluations of each crew. This confirms that with appropriate training and calibration, evaluators can use specific, observable behaviors as the basis for judgements which should increase the objective component of CRM evaluation in an LOE and ensure more fair, unbiased evaluations of each pilot.

CRM ratings, across event sets, as well as technical ratings significantly contributed to the overall evaluation of PIC and SIC . This confirms that evaluators, when properly trained and calibrated, can consistently integrate CRM evaluation with technical performance to obtain overall evaluation of crew members on each event set. This is evidence that the LOE evaluation was reliably tapping CRM as well as technical performance. This is important because, in the past, there were questions as to whether CRM ratings were just a reflection of technical performance. It has been suggested that evaluators could not or would not rate CRM behavior as a separate performance domain, but part of the entire crew performance.

Additionally, the analysis showed that Captains and First Officers were evaluated differently. In evaluations of the entire LOE, CRM performance was more important than technical performance for the PIC, while technical performance was more important than CRM performance in rating the SIC. These differences in evaluations are congruent with holding the PIC more responsible for CRM as part of the Captain's role. These findings are also congruent with previous findings that Captains state that the most important feature of a good first officer is technical competence and First Officers state the most important feature of a good Captain is CRM competence.

Further analysis showed stronger relationships of Captain evaluation to Crew performance compared to First Officers. This provided additional evidence that Captains are held more responsible for crew performance particularly CRM performance. This empirically confirms that Captains are the focal point for CRM and crew performance and suggests additional training on establishing and maintaining CRM should be performed in upgrade training when a First Officer transitions to Captain.

LOE data was compared with the data from traditional proficiency check of specific maneuvers. The proficiency check evaluations could predict about 50% of the overall evaluations of PIC and SIC on the LOE. This confirms that the technical proficiency emphasized by the traditional check is relevant to LOE performance. However, the remaining 50% of the variance unique to the LOE confirms that the LOE is also tapping a distinct type of performance in addition to technical proficiency. More detailed evaluation of proficiency check assessments showed that only evaluations of the SIC do not predict LOE results. Additional analyses are required to better understand the role of rating SIC performance in proficiency checks.

Analysis of different event sets found significant differences in the performance of the targeted observable behaviors. For specific event sets, the targeted behaviors (e.g. specific briefings) were observed significantly less often. Isolation of performance problems across crews can be used to rationally and efficiently change pilot training. The lack of specific observable behaviors can be tied to specific instructional objectives using an instructional database such as the model AQP database. The database can also be used to find the specific training components for these objectives that must be strengthened. This analysis shows that isolating poor performance across crews can be used to directly change relevant components of training.

The proceduralized CRM concept has been extended to a major carrier where CRM procedures were integrated into a QRC for emergency procedures. The development and implementation of the QRC

represented a fundamental change in how the carrier addressed emergency procedures in the past. The objectives of employing the QRC were to reduce the number of memory items which can lead to operational errors, and to ensure that all critical actions are taken in sequential order, eliminating errors of commission and omission. The study was initiated to validate the usefulness of the QRC and to identify design problems, training issues and overall ability of the QRC to improve system safety. The study assessed 124 crews performing checkrides with the QRC or by the traditional memory immediate action items. The crews were assessed on the following elements:

- Timing of the crews from the beginning of the Emergency event until the completion or the initiation of reference action items.
- The timing of when each action of the checklist was initiated.
- Errors in crew performance caused by omission of checklist items.
- Errors in crew performance caused by omission of checklist items.
- Errors in crew performance caused by addition of items to the checklist.
- Errors in crew performance caused by not following the sequential order of the checklist.
- Overall crew performance
- Crew survey and feedback from on QRC philosophy
- Specific timing events on critical checklist

The data indicate that the QRC checklist significantly reduced crew error in checklist performance. This was measured in errors of omission, commission and order. The QRC checklist received overwhelming approval by the line crews participating in the study. Significant changes were made to the QRC during the study improving the product based on line input. The data also show that the QRC does increase the time for completion of the emergency checklist. However, with the increased accuracy crew performance is actually enhanced. Further, QRC crews received overall higher ratings in CRM particularly in the areas of communication and workload management.

A paper was written describing the study results and distributed to industry, 1/98.

Work was continued with several carriers to identify appropriate questions, data collection strategies and statistical methods to analyze pilot performance data. This data collection system will allow carriers to evaluate their AQP training effectiveness in many dimensions. The overall goal of this part of the research is to determine the appropriate data necessary to extend training intervals for AQP programs. Extensive AQP statistical analyses activities have been performed on data from major and regional carriers.

In order to insure that reliable and valid performance data is being collected by carriers and being reported to the FAA this research team found it necessary to address rater calibration issues. The overall goal of this segment of the research is to assist carriers with incorporating interrater reliability training strategies into existing instructor/evaluator (I/E) indoctrination training. This project involves: Selecting and evaluating existing videotapes or crew performance to use as "gold" standards for the I/E

indoctrination class; developing behaviorally-based grading sheets for the selected video tapes; and collecting and analyzing data from indoctrination classes on levels of agreement and emerge as a result of training.

This rater training has continued at major and regional carriers. The purpose of this training is to convey the importance and knowledge/skills necessary for instructor/evaluators to provide stable, reliable and accurate assessments of pilot performance. Continuing refinements have been made to the inter-rater reliability software, (developed earlier in this research) instruction materials and the process itself. Some of the improvements include better feedback to the instructors, a training guide for conducting inter-rater reliability training, and a new summary report for the IRR session. This summary report closes the loop from IRR training to fleet captains and managers and facilitates the needed changes in flight standards operations.

This research team has developed an entire Instructor recurrent training package for Aer Lingus in an effort to take the research results to the international community. This training was delivered to Aer Lingus 1/98.

**Products:** A cockpit management training system. Research report on performance differences due to proceduralized CRM training. Inter-rater reliability training. Pilot performance analysis systems.

**Project Title:** *Crew Resource Management: Design and Evaluation of Human Factors Training in Aviation*

**Performing Agency:** The University of Texas, Austin

**Background:** This research consists of several projects that center on evaluating the impact of CRM and LOFT and developing methods to optimize training and performance evaluation. This research group has maintained an international, longitudinal, data base of pilot performance since the late 1980's. Prior findings from this work indicate that formal CRM programs combined with LOFT have a significant, measurable positive effect on crew behavior and attitudes and by inference on system safety. However, the data also indicate that all human factors problems in aviation have not been solved by the current CRM and LOFT programs. Specifically the following negative outcomes have been documented through this research program:

- (1) There is great variability in the impact of programs and in observed crew performance, even after implementation of CRM training. Incomplete acceptance of CRM concepts and practices by instructors and evaluators has been identified as one source of problems.
- (2) In the absence of continuing reinforcement behavior and attitudes regress almost to pre-training levels.
- (3) LOFT programs in several organizations show great variability in execution and impact.
- (4) Considerable controversy remains regarding the evaluation of individual and team performance, as required of organizations participating in the Advanced Qualification Program (AQP).
- (5) There is an absence of reliable data on the human factors of incidents occurring in the aviation system.
- (6) Recent investigations of major accidents including air carrier crashes have shown that organizational cultures have a profound impact on operations and on the effectiveness of human factors programs.
- (7) Human Factors training and awareness in Air Traffic Control (ATC) has lagged behind programs in air carrier operations.
- (8) Human Factors aspects of cockpit automation are not being addressed in most CRM/LOFT programs and have been identified as critical in a number of accidents.

**Project Description:** The research program consists of several projects that centers on evaluating the impact of Crew Resource Management (CRM) and Line Oriented Flight Training (LOFT) and developing methods to optimize training and performance evaluation. Research findings from this work indicate that formal CRM programs combined with LOFT have a significant, measurable positive effect on crew behavior and attitudes and by inference on system safety. However, the data also indicate that all human factors problems in aviation have not been solved the current CRM and LOFT programs. Specifically the following issues have been documented through this research program:

This research program is designed to build on current knowledge and to address the issues noted above. This research program has developed and maintains a massive longitudinal data base on the effectiveness of CRM in the Aviation industry. Operational line and LOFT observation data comprise the data base.

Confidential data are entered into the data base and reports are delivered to each participating carrier. This information is used to remedy the weaknesses apparent in the observations or to highlight strengths. It is common practice for carriers to use these reports to base the following year recurrent training. This research group disseminates research findings to scientific meetings and is widely published each year.

This project has developed an observational methodology for assessment of crew performance in line operations. This line audit approach using the Line/LOS Checklist (LLC) employs a team of expert observers to collect systematic data on line operations *without jeopardy*. The research team collect data in each organization giving a common comparison point that allows comparisons of performance across organizations. To date, line audits have been completed in five major airlines. A large, multivariate analysis of the total data is in progress that will form the basis of technical reports and doctoral dissertations. The results show large differences in human factors practices between organizations and between fleets within organizations. A side benefit of the audit process is that organizations have used the findings to target specific issues during recurrent training.

*Error Management.* The project has recently refocused CRM training on error management. Curriculum issues of error management CRM are being developed. This change in emphasis is mirrored by a revision of the LLC to capture the dynamics of error management and the relationship with other aspects of flight management. This should result in new training emphasis for instructors and evaluators. Presently, this training is in place at a major carrier.

*The Role of Cultures.* The research group is collecting data from different national and organizational cultures to determine how culture effects crew performance. The goal is to develop training guidance for carriers employing cross-cultural cockpits or for carriers operating with non-Western cockpits. One of the side benefits of the investigation of culture has been the discovery that automation use varies significantly as a function of culture.

*Automation.* Because flight deck automated systems can execute most control functions performed by pilots, the flight management computer has been likened to an "electronic crewmember". This view of automation has led to speculation that the manner in which crew members interact with the FMC may influence crew behavior and performance. The link between interpersonal flight management attitudes and performance in aviation suggests that attitudes regarding flight deck automation are likely to be related to pilots' use of automated systems and overall performance. This research project has collected data from more than 10,000 pilots from more than 10 cultures addressing attitudes towards automation. Analysis of this data showed significant fleet, organizational and national differences in both general attitudes regarding the conduct of the flight and attitudes regarding the use of flight deck automation. Currently, this research project is investigating these issues with an expanded methodology and will further explore the link between national differences in attitudes toward automation and a multi-dimensional model of national culture.

An initial finding regarding national culture and automation resulted in a new emphasis on automation training and use within the U.S. A new survey was developed in collaboration with pilot groups and airline personnel. Data were collected from most US carriers regarding automation training and use. Use of automation has also become a central element of the line audit process using the LLC.

One of the critical elements of the project is observation of how crews actually use automation in line operations as part of the line audit approach. Data show, for example, that concepts taught by carriers in philosophy or automation courses do not uniformly find their way to the line. One of the project goals is

to identify better strategies to integrate CRM training and automation training and facilitate transfer to line operations.

*Validating the Importance of Behavioral Markers in world wide aviation.* This research is correlating US line audit data, that uses the behavioral markers developed under this project, with a world wide accident data base. This is being done to develop specific recommendations to reduce CFIT and approach and landing accidents in collaboration with Flight Safety Foundation. An additional goal of this project is to validate further the role of the behaviors markers defined in the Line/LOS CHECKLIST in safety. The results are also being analyzed in the framework of the NTSB's 1994 Safety Study to determine common casual elements worldwide and in audit data. A report with recommended strategies to reduce accidents will be issued in FY98. This is being done in order to determine the influence of certain cockpit behaviors and accidents worldwide rather than just in the United States.

*Optimizing Air Safety Action Partnership Data.* This project, in collaboration with several major airlines, is developing a self-report human factors checklist to be used by crews to report incidents to their parent organization. This effort is in support of the FAA's Aviation Safety Action Program designed to elicit reports from crew members under non-jeopardy conditions. the project's further goal is to develop a statistical methodology to integrate data from surveys, line audits, formal evaluations and incident reporting systems to provide the most comprehensive and valid picture of an organization's operations. In collaboration with the FAA, these de-identified data can also provide a valid picture of elements of the aviation system.

*Cabin Safety Issues.* An extension of the research to examine issues related to safety and CRM training for cabin crewmembers has also been undertaken. A new survey instrument was designed and data have been collected from more than 2,000 flight attendants at a major airline. Results pointed to a number of safety-related issues and problems in leadership between the cockpit and the cabin.

A large number of direct observations along with instructor and line check airmen comments have been analyzed regarding incidents that were successfully and safely resolved. This data can be used to develop training scenarios and is available to air carrier training departments.

*CRM for Regional Carriers.* A segment of the work done recently by this program was to analyze the particular CRM requirements of the regional carriers. This research suggests that regional CRM programs should focus on specific elements of the commuter environment. The programs must address low experience level and high turnover of both pilots and flight attendants. Also, pilots may be thrust into automated equipment without extended training in non-automated equipment which may affect trust in the automated systems. Also, these programs must address the environmental variables: high frequency, low altitude, turbulence and small airports.

*Performance in LOFT.* Another component of the program focused on the effectiveness of LOFT. The goal is to determine the key components for high quality LOFT. Scenarios were analyzed for appropriate workload, creative problem solving required, problem realism, and simulator mechanical condition. Briefings were analyzed for how did the instructor integrate the crew into briefings, how was a review of CRM concepts handled, and instructor enthusiasm. Scenario execution was also analyzed for simulator operation, flight attendant roles, adhering to script and radio calls. Debriefings were analyzed for crews involvement with self-appraisal, the comprehensiveness of the debrief, and the integration of technical and CRM. Findings showed that, overall, the briefing and debriefing quality were the most important aspect of LOFT quality followed by the quality of the scenario and least important was the scenario

execution. These findings were made available to the Airline Transport Association subcommittee on LOFT development.

*Air Traffic Control Research.* This research program also investigated team and group determinants of performance in the Air Traffic Control system. This project was conducted with the Southwest Region to investigate the relationships between the behavior of organizations, work teams and individuals and three possible operational outcomes: loss of separation, acceptable performance (absence of remarkable event) or outstanding performance in the presence of increased challenge. Data for each type of outcome have been collected and specific behavioral predictors of each outcome determined. The results were reported in a doctoral dissertation, a technical report and a forthcoming paper. Results have also been presented at industry meetings.

**Products:** Research methodologies and data collection tools. Technical reports and scientific papers (available on the project's homepage, [www.psy.utexas.edu/psy/helmreich/nasaut.htm](http://www.psy.utexas.edu/psy/helmreich/nasaut.htm).) Presentations of project results have also been given at industry and scientific meetings. Robert Helmreich serves on ICAO's Human Factors Team, to deliver training in human factors to airlines around the world. During the project's tenure, seven doctoral dissertations have been completed. Most of these Ph.D.s are working in aviation positions.

**Project Title:** *An Investigation of Training Issues Concerning the Advanced Qualification Program (AQP)*

**Performing Agency:** Battelle Memorial Institute, Columbus, Ohio

**Background:** In an attempt to encourage the use of innovative training programs the FAA as proposed the concept of Advanced Qualification (AQP) training. Central to AQP is proficiency based training. Under AQP, carriers who have applied for inclusion in the program can develop their own proficiency objectives, which must address the range of conditions and contingencies that might be faced by pilots working within the carrier's operational domain. These proficiency objectives define the set of skills and tasks a pilot must be able to perform to be proficient on a given aircraft type within the carriers operational domain. The goal is to ensure that the training program meets each carrier's specific requirements and does so in the most efficient way possible. Utilization of a strong analytical frame work for developing a carrier's program helps to ensure that training is systematically oriented towards those objectives of greatest relevance to the individual carrier and also supports meaningful crew and program evaluation. One of the objectives of AQP is to provide seamless integration of CRM and technical skills within the curriculum to ensure that CRM skills are practiced together wit all other flight skills and procedures as required by each flight situation. One goal of AQP is that CRM skills should be utilized as normal and inherent part of aircraft operation little different form operating the automation or performing a proper checklist.

While AQP permits significant departures from the traditional FAR requirements for training and checking airmen, the price of that regulatory flexibility is a detailed front end analysis, the methodology for which is described in Advisory Circular (AC) 120-54. Accomplishing the analyses necessary to create AQP qualification standards has proven to be particularly challenging for participating air carriers. Training developers have had difficulty selecting an appropriate level of analysis detail, efficiently executing their analyses and determining how to incorporate cognitive and crew resource management considerations in a manner that will generate meaningful proficiency objectives, standards, and conditions. These difficulties stem, in part, from the fact that as a new program, AQP lacks concrete examples for reference purposes. The methodology was developed by the FAA on an a priori basis, with the explicit intention that it be refined subsequently as experience with the new programs accumulate. There was a need for a methodology that, if faithfully followed, would produce an effective AQP.

**Project Description:** The goals of this research project are to assess whether the Instructional Systems Design (ISD) process can be effectively used to develop a prototype AQP, including specification of the process that derives learning objectives from the task analysis; (b) determine whether the ISD methodology needs to be modified so as to best support AQP development; and © verify the modified ISD methodology that has been standardized through the development of a database comprised of templates and instructional guidance for developing an AQP.

The approach to integrating CRM into technical training adopted by this research utilizes ideas from several sources: (1) ISD methodology used for AQP curriculum development (2) The event set approach developed in prior AAR-100 research and (3) The situation assessment model being developed by this research team for the Model AQP. Each of these sources provide a useful and unique perspective. The ISD methodology, as currently implemented, focuses of the technical skills, knowledge and procedures required to accomplish specified tasks and subtasks. The methodology is especially effective for activities that occur at predictable times and in a standard order. The event set approach, in contrast, focuses on a selected sequence of situations which attempts to mimic real-world situations with all of the

attendant complexity. The objective is to evaluate crew performance in situations that require pilots to utilize both technical and CRM skills. This approach is currently used by many carriers to support line operational simulation. This research project's situation assessment model attempts to provide a cognitive perspective by focusing on those factors that influence a flight crew's assessment of a situation and subsequent management of available resources. In the past there has been little continuity between the task-oriented front end analysis provided by the ISD methodology and the situation-oriented event set approach. This lack of continuity is exemplified by differences in types of CRM skills addressed by each approach. The ISD methodology is best suited to handle phase-specific skills, that is, those activities which are always performed by the crew for a given task or subtask. Traditional ISD's behavioral orientation supports its emphasis on the specific task and subtasks that must be performed to complete a job.

This task orientation does not support those aspects of the pilot's job which fall outside of the sequential tasks and subtasks found in a task list. Instead, the unique dynamics of the aviation environment necessitate a change in focus to the situation as a whole, including conditions under which a task or subtask must be performed (e.g. weather, aircraft system, failure and the requirement to utilize phase - independent flight management skills either on a need basis or continuously to ensure that the flight is properly managed. Appropriate utilization of phase-independent skills depends upon crew judgement: Accurate assessment of the requirements of the situation together with effective utilization of those skills and information sources most likely to be useful in that situation. This judgement depends upon an understanding of the situation as a whole, not simply the task in isolation.

This situational focus is the strength of the event set approach. An effective training program will enable flight crews to experience these situations so as to allow them the opportunity to practice the phase-independent skills required to cope with these situations.

Both the ISD and event-set methodologies bring important and unique perspective to an AQP. Merging them into a coherent approach will support the development of a complete training program. One means by which this integration can occur is to place the focus of flight training on situations rather than tasks throughout the program rather than waiting until LOFT. A situational orientation throughout training helps to ensure that phase-specific and phase-independent technical and CRM skills are practiced in an integrated fashion. In addition, the situational orientation gives flight crews the opportunity to practice those skills involved with assessing situations and utilizing available.

This transition from the task focus of the ISD methodology to situation orientation takes place in the Model AQP by means of the concept of an event. An event includes a specific task (i.e. a maneuver or set of procedures) together with the conditions (weather, malfunctioning aircraft system, etc.) under which the task is to be performed. To handle an event successfully requires that the crew quickly and accurately assess the situation, plan how to manage the event and utilize the technical and CRM skills appropriate for that event. In addition, the set of events included in curriculum can be selected to ensure that important technical and CRM issues are addressed. Each event has a specific topic or theme that is the point of that event. The flight training curriculum can be designed by strategically selecting and positioning events in accordance with these themes.

One of the strengths of the event concept is its applicability to both ground and flight training. Continuity throughout all parts of an AQP is a critical goal for the Model AQP project. Events can be used as the buildingblocks for both the ground and flight training curricula. For ground school, one of the goals of the Model AQP is to utilize scenario-based training, where students would be required to not only acquire

new information but also learn how to apply that information to solving problems. Events are a natural tool for designing a scenario-based ground school curriculum. Similarly, events can serve as the individual units for FTD and simulator training. Finally, they will continue to serve as the building blocks for LOS scenario development.

The first generation Model AQP which has been targeted to regional carrier training, has been completed and is in use by approximately 20 carriers. Several workshops have been conducted to train air carrier personnel in the use of this model.

During FY 98 development will begin on an Advanced Model AQP incorporating many features requested by the carriers and allowing each carrier to pursue its own level of development.

**Products:** Model AQP and research report delineating process, methodology and lessons learned. Advanced Model AQP.

**Project Title:** *LOFT Debriefing Study*

**Performing Agency:** NASA Ames

**Background:** Training effectiveness of LOFT simulations is heavily influenced by the debriefing that occurs afterwards. The LOFT simulation is a very busy, intense, and sometimes stressful experience. Thoughtful discussion after the experience is necessary in order for the crew to sort out and interpret what happened and to consolidate the lessons learned into longterm memory in a form that can be used later in actual line operations. The debriefing is a window on the entire CRM process. LOFT debriefings can demonstrate how well crews are able to analyze their performance along CRM dimensions. It is thought that in order to implement CRM effectively in day-to-day line operations crews must have the skill and the habit of analyzing their own performance in terms of CRM.

**Project Description:** The purpose of this project is to determine which techniques are actually being used by LOFT instructors, how effective the techniques seem to be, the extent to which those techniques are consistent with FAA Advisory circular guidelines, how practical the guidelines are for real-world training and what obstacles instructors encounter in trying to teach to these guidelines.

Data have been collected from five major US air carriers. At each carrier five to eleven instructors from different fleets, made audiotapes of the debriefings they conducted. The audiotapes are transcribed and coded, generating a large database. Findings are as follows:

- (1) Typically, the debriefing is scheduled after the practice of proficiency maneuvers, rather than immediately after the LOFT. Interposition of the proficiency maneuver practice constitutes and "interference experiment" that probably impedes the ability of the crew to remember what happened in the LOFT.
- (2) The similarities among carriers are greater than differences in how they conduct LOFT and LOFT debriefings. There are large differences among individual instructors within carriers in facilitation effectiveness.
- (3) Most instructors conscientiously attempt to elicit crew participation, but some unwittingly sabotage their own efforts with behaviors that discourage participation.
- (4) Most instructors fail to make clear at the beginning of the debriefing the nature of the participation that they expect of the crew and do not explain why is important that the crew take an active role.
- (5) Although instructors have been told to facilitate crew self-debriefing, they have been given only very general advice on how to do this. Training departments should provide much more explicit hands-on training in facilitation and should mentor new instructors as they start to facilitate.
- (6) The content of the debriefing is strongly driven by the instructor's observations and questions, rather than the crew's self analysis.
- (7) The crews are responsive to the instructors, but are more reactive than proactive.

(8) Most crews would probably have difficulty in conducting a deeply analytical debriefing on their own, however they appear to be quite conversant with and accepting of CRM and could with practice develop debriefing skills and apply them routinely on the line.

This data and analysis has resulted in major carriers developing a list of do's and don'ts for facilitators and is being used as the core of annual recurrent training for instructors.

The findings from this research will be incorporated into a CBT training system for instructors at a major and regional carrier in FY99.

**Products:** CRM Debriefing training guidelines were distributed in FY 97. A handbook for instructors was developed and distributed in FY 97. The handbook was also incorporated into the instructors manual for Leadership/Followership curriculum development at Alaska airlines, distributed 5/98.

**Project Title:** *A methodology for developing Gold Standards for Rater Training Video Tapes*

**Performing Agency:** NASA Ames and University of Central Florida

**Background:** A central issue in the Advanced Qualification Program (AQP) is how to train check airmen to provide reliable, consistent, and valid assessments of Crew Resource Management (CRM) skills. It has been suggested that to enhance reliability of CRM assessment, measures should focus on clearly defined observable behaviors. These behaviors should be explicitly identified so that no ambiguity exists for check airmen conducting CRM assessments. Raters should be calibrated so that check airmen are consistent in their evaluations. Check airmen should be trained by rating and discussing video tapes of aircrews. The purpose of the present project was to develop and validate a methodology for constructing gold standards video tapes for rater training.

**Project Description:** The purpose of this present project was to develop and validate a methodology for constructing a "gold standard" for rater training video tapes. Gold standards are defined as criteria that are assigned by experts to a training video tape. These tapes can be used of training rater reliability and accuracy.

To meet this objective, a five step methodology was developed.

- Step 1 Select Video Tapes of LOS Scenario
- Step 2 Develop Gold Standards Data Collection Forms
- Step 3 Conduct Gold Standards Development Meeting
- Step 4 Analyze Gold Standards Data
- Step 5 Establish Gold Standards

The methodology was tested at an air carrier that was interested in developing gold standard video tapes for check airmen rater calibration training. CRM, technical and event set gold standards were developed for a video tape of a crew performing an LOE scenario.

**Products:** A technical report was published and distributed to industry in 3/98.

**Project Title:** *Integrating CRM into Crew Procedures: Checklist Standardization*

**Performing Agency:** NASA AMES

**Background:** Non-compliance with checklists has become a potentially serious problem for air carriers and the FAA. In today's aviation system, procedural changes occur frequently and are instituted for a large variety of reasons; company mergers, changes in air traffic procedures, changes in airport environments, introduction of new technologies and new aircraft, etc. While one obvious goal is to maintain compliance and standardization of crew member performance, there must not be a sacrifice in the match between procedures and the operational realities in which the procedures are to be used. Not only must procedures reflect a reasonable match to technical task demands, they must reflect consistency with company and system requirements. Within an air carrier for example, philosophy, policies and procedures should be made consistent both within teams (within aircraft type) and across teams (across fleets within the company) to the extent that standardization is feasible. Since there are legitimate aircraft and mission differences across fleets, a means of evaluating the appropriate level of standardization must be developed and reviewed when changes are anticipated.

**Project Description:** The goal of this research project is to integrate CRM principles into crew roles and procedures in order to: 1) improve the match between current procedures and operational reality, and 2) develop a process for adapting procedures to anticipated changes.

The basic approach is to analyze procedures from a team perspective and to conduct a team task analysis in order to assess whether current procedures match the task flow including team processes. The integration of CRM principles into the task analysis provides the basis for reviewing and re-designing procedures. The development of standard methods and measures for evaluating procedure usage is important for assessing both current and proposed procedures. Measures used during research process will be refined for validation purposes as well as eventual use in training.

The objectives of this research project are to develop a systematic approach to checklist standardization across all fleets, to develop guidelines for implementing and evaluating procedure changes that optimize, 1) within company standardization, 2) compatibility with aircraft and operational differences and 3) that the procedures are consistent with CRM principles. This research is complete has established company-wide philosophy and policy statements: Operations, Automation, Training, and Checklists. Also, methods for data collection have been identified and developed. Pilot feedback on proposed checklist changes has been completed.

**Products:** Guidelines for standardization of procedures.

*Note:* This project led to the following project: Operating Documents Human Factors Project

**Project Title:** *Operating Documents Human Factors Project*

**Performing Agency:** NASA Ames; George Mason University

**Background:** The original project described earlier was intended to develop a systematic approach to checklist standardization and to develop guidelines that would optimize within-company standardization for eight fleets, to insure compatibility with aircraft and operational differences and be consistent with CRM principles. This project considered the form of the checklist with respect to content, structure, logic, format, terminology, roles and usage. However, this approach was inadequate as checklists are part of the whole operating document system. It became apparent that operating documents must show internal consistency across fleets (aircraft types, and route structures) across departments, pilot's handbooks, company policies, etc.) Further, these documents must be externally consistent with regulations and manufacturers. Document systems must be consistent across philosophies, policies, procedures and practices. This concept is not in place in the air carrier industry.

**Project Description:** This research team developed surveys and held workshops to bring together air carriers (including regionals and cargo), manufacturers, and the FAA to work on a better, more human centered approach to these document systems. Two workshops were held in FY 97. The most recent workshop was attended by 78 participants, indicating industry's interest in this topic.

Based on the information provided on the surveys and in the workshops, guidelines are being written for the development of operating documents that will include: the Organization of Documents, Standardization of Documents, Usability of Documents, Developing and Maintaining Documents, and Transition to Electronic Media.

Feedback from the air carriers indicate the need for additional workshops. The structure and content has not been determined nor has it been determined whether this process should be turned over to industry to lead. Feedback from internal FAA sponsors indicate that the integration of AQP documentation should be considered in future work by the research team. This focus is presently being considered.

**Products:** Reports from the two workshops have been written and distributed to industry, 1/98.

**Project Title:** Team Situational Awareness

**Performing Agency:** Naval Air Warfare Center, Orlando

**Background:** There is little question that situational awareness is critical for safe flight. Much research has centered on attempts to identify and quantify individual situational awareness. However, since cockpit crew performance depends heavily on team processes, it is necessary to determine exactly what aspects of individuals and the environment either enhance or degrade team situational awareness.

**Project description:** This research project has identified a model of team situational awareness, developed effective instructional strategies to improve situational awareness training, developed situational awareness assessment tools, and developed guidelines for situational awareness scenario development.

The initial framework for team situational awareness has been completed. This framework will serve as a basis for the team situational awareness model. Preliminary training guidelines based on this framework, follow:

- (1) Tailor training to the experience level of the crew in terms of time in aircraft, time in position, and familiarity with locality. For example: Consider a newly hired first officer or flight engineer that is transitioning to or learning a new aircraft and flying out of a new area. It is important to ensure that the crew member is familiar with cockpit roles, the geographic area, and the aircraft equipment before expecting him/her to have the awareness to handle dynamic problem situations in that environment.
- (2) Ensure job/task analysis provides information about the cognitive, perceptual and behavioral demands required for situational awareness. For example: A job analysis of a pilot who flies a highly automated or "glass cockpit" aircraft should solicit information particularly relevant to that environment, emphasizing factors such as information seeking from different sources, collation and interpretation of information from complicated displays, scan through the FMS, mode errors with the FMS, manual data entry task demands, and communications associated with the advanced technology.
- (3) Include training which specifically focuses at the team level of situational awareness in addition to training which focuses on individual skill development. For example; Individual skill development in situational awareness may include: providing information that increases the individual's knowledge base(so that he/she knows what information to seek and when) helping in the development of planning skills, so that he/she is better prepared for completing tasks within the flight; and providing tips that help free up time from basic tasks in the cockpit, so that more attention may be paid to the situation. Team skill development in situational awareness should include team process skill training such as communication and leadership behaviors to enhance and maintain crew awareness.
- (4) Based on the complexity of the situation, provide training which emphasizes adequate information, demonstration, and active practice and feedback, all are needed to produce a behavioral change. For example: To enhance the overall training experience provided in a realistic LOFT scenario emphasizing situational awareness, the crew members should receive advanced preparation about the training objectives, including specific information on the skills to be trained and demonstrations or examples of effective and or ineffective behaviors relevant to those skills. Providing general information about situational awareness concepts would be insufficient to expect the crew to obtain the full impact of LOFT. Crew members who are a low experience level need to have specific information on the kind and timing

of information transfer that will enhance crew situational awareness. Alerting them to the importance of communication, if they do not know what and when to communicate, would not prepare them to practice good communication skill in the LOFT session.

(5) Design training scenarios to be realistic and to elicit the desired cognitive, perceptual and behavioral task demands encompassing the situational awareness construct. For example: Consider that including an unfamiliar event, such as icing, for crews who have only flown in Florida would require them to recognize the relevant cues, comprehend their meaning and project a course of action. A very different type of problem, such as a gradual change in an instrument due to a developing condition, may require perception comprehension and projection as well, but the cognitive and behavioral demands related to this problem are very different from those imposed by icing conditions. Including both types of problems within a scenario would allow the trainee to exercise these different cognitive and behavioral skills, and would help him/her to learn to generalize the application of these skills across various problem situations.

(6) Conduct training which focuses on improving team or process skills related to situational awareness(e.g.planning, communication, leadership) because this may aide the development of team situational awareness. For example: Training focusing on communication skills related to situational awareness can include: training that encourages the flow of the information that needs to be transmitted in the planning phase of the flight (e.g., what questions the Captain can ask to solicit information, what questions the First Officer should ask, what information should be covered with Flight Attendants.) It can also include practice in the communications that are most effective (e.g., what information to provide to another crew member or when to communicate the information to the other).

(7) Include training which emphasizes developing accurate and appropriate knowledge structures for expectations related to situational awareness. For example: Training a crew about how traffic flows around a specific airport, prior to arriving there for the first time, helps the pilots to understand what can be expected, thereby increasing the readiness with which he/she can recognize that an abnormal situation is occurring.

(8) Use realistic and dynamic scenarios to provide training on: information seeking and processing; detecting anomalous information; prioritizing attention to cues, especially when there is information that may be related to different problems; and taking appropriate actions based on cues. For example: Design realistic scenarios with events that require crew members to: look for and use information; handle multiple cues; determine when information is signaling a problem or possible problem; and act on the information. Intermittent fluctuations of an instrument during a scenario require crew members to recognize the fluctuations and to diagnose the problem. This may include cross checking instruments, discussing among themselves, and seeking additional information. Putting this event in the scenario at the time of moderately heavy workload or when another troubleshooting event is occurring gives crew members practice in dividing their attention and determining priorities.

(9) Design and develop performance measurement tools for the behaviors, the knowledge and perceptions required for situational awareness. For example: A realistic scenario provides information about situational awareness behaviors and an observation scale can be developed for those behaviors. Tapping into the crew member's knowledge and perceptions may require asking questions, either during the scenario or after it is completed to determine if he/she saw the problem and understood its possible impact. For pilots with low experience levels, questions can be either a part of the scenario (e.g., ATC can question the crew members) or the scenario can be stopped briefly for questions to be asked.

(10) Measurement tools should measure individual situational awareness skills as well as team interactions related situational awareness. For example; An observation scale can be designed that measures what individual team members contributed to the team situational awareness and what actions they performed independently that demonstrated their situational awareness. The same scale should include opportunities to document team interactions (e.g., planning, communications, and leadership) that helped the crew members gain and maintain team situational awareness.

(11) Situational awareness measurements of behaviors should include both outcome and process measurements. For example: Outcome is the action taken by the crew in response to an event (e.g., the crew diverted to Airport A) and process is how they achieved the outcome (e.g., the Captain gathered information from others about the decision, crew members discussed the merits of each airport, the crew members noticed that the runway at Airport B was not adequate for their needs). Both are important to understanding performance.

(12) Include instructions and information about training objectives to instructors about each training method or tool. For example: Write out the specific objectives to be accomplished at the top or on a cover sheet of scenario instructions. Include clear objective behaviors as part of the scenario goals. The instructor is then informed about how to use the measurement tool to re-enforce the training objectives.

(13) Link the measurement criteria to the scenario events and to the debrief plans, for consistency in training. For example: If a generic measurement instrument is used, prepare the raters to recognize and document specific examples observed in the scenarios that relate to the general behaviors on the rating form. It is likely that the raters will need through training so they can map the scenario events and behaviors observed to the rating form. Thus in the debrief, crew members will know what specific actions they took and how those relate to general requirements for situational awareness.

(14) Multiple measurements over time are necessary to yield a good measurement of situational awareness; single measurements at one time are insufficient. For example: In a single LOFT scenario, build a number of opportunities for crew members to demonstrate situational awareness (e.g., NOTA information, ATC calls to other aircraft, PIREPs, subtle changes in instruments, weather, unexpected circumstances) and make sure that observers know when, where, and how to observe and document situational awareness actions. The same could be accomplished by using a number of small scenarios (each with opportunities to observe situational awareness) and then evaluating the aviator's situational awareness across the scenarios.

(15) Use specific feedback tools along with realistic and dynamic scenarios to provide feedback that; is specific to the trainees on their situational awareness skill performance; reinforces the development and maintenance of relevant knowledge; and stresses adapting knowledge to meet situation demands. For example: Give observers feedback forms with the specific situational awareness elements that were built into the scenario. Observers should use these forms to document behavior they see in the scenario. This is then used to give specific feedback on what the crew members did, how they did it, and the context of the behaviors. Explanation of the rationale for the feedback helps the crew members develop and add to their existing knowledge.

(16) Train instructors on making ratings and giving feedback. For example: This training can be done by providing information to the raters on situational awareness, showing them examples of the behaviors associated with situational awareness (so raters can be knowledgeable about the behaviors), giving them

practice in rating those behaviors, and giving raters practice in providing feedback on specific situational awareness behaviors.

The above guidelines have been expanded, published and distributed in 2/98. The guidelines are too lengthy to be contained in this report. Although these guidelines have received positive comments from the air carrier training community there remains a fundamental problem with situational awareness training and assessment. Several carriers believe that new instructors, and line pilots, in general, are not able to dissect pilot performance and to identify and articulate when and what exactly led to poor situational awareness. Thus crews and instructors are not adequately debriefing flight scenarios, either in training or on the line. In response to this critical requirement this research team will design a CBT based system to provide the fundamental principals in situational awareness cue recognition. This system will be designed at one carrier for pilot training and at another carrier for instructor/evaluator training. It is expected to be complete in FY 2000.

**Products:** Situational Awareness guidelines. Situational Awareness training systems.

**Project Title:** Training Pilot Not Flying (PNF) Back-up Skills

**Performing Agency:** Naval Air Warfare Center, Orlando

**Background:** Regional airline Captains indicate that one of the greatest difficulties they experience is working with inexperienced or new First Officers who lack the skill to effectively provide back-up, when serving as Pilot Not Flying (PNF), to the Captain when it is needed. Since the complex environment of regional flying necessitates the coordination of the two crew members it is critical that this issue be addressed.

**Project Description:** This research project identified the knowledge, skills and attitudes necessary for PNFs to properly perform back-up duties. In conducting this research, (1) structured interviews with pilots from the regional airlines were conducted to inquire about effective and ineffective back-up behavior, (2) the 100 most recent submissions to ASRS database were analyzed using the descriptors of "regional airlines" and (3) analysis of the recent NTSB report which reviewed flight-crew involved major accidents of U.S. air carriers that occurred from 1978 through 1990 and which identified poor monitoring/challenging behavior on the part of the non-flying crewmember as a causal factor in seventy-six percent of these accidents. Combining information from these sources, three broad problem areas requiring additional training were identified. (1) first officer or PNF assertiveness in ambiguous situations, (2) monitoring and challenging by the first officer or PNF, and (3) proficiency in technical and communication procedures.

**Products:** Training guidelines have been developed, published and distributed 12/97.

**Project Title:** Decision Making and Judgement on the Flight Deck

**Performing Agency:** NASA Ames; the Naval Systems Warfare Center, Orlando

**Background:** NTSB reports in a four year period 1983-1987 crew judgement and decision making was implicated in 47% of fatal accidents. NTSB and ASRS cites many more reports where poor decision making was evident in many potentially serious incidents. Decision making is a component of most CRM courses and is encouraged by the FAA Advisory circular 120-51, 1993. However, scientific research has not been available to support this training. Decision making in the dynamic, time constrained cockpit environment does not mirror decision making in a static laboratory setting. Thus, findings from generic decision making research do not necessarily generalize to aviation settings.

**Project Description:** Decision making is an essential component of a Captain's expertise. The Captain is responsible for making the hard decision, for example, choosing where to divert after a system malfunction, when fuel is short and weather is deteriorating; determining how to cope with passenger's medical emergency; evaluating whether to take-off with a placarded system given past experience with the projected weather and traffic at the destination. The captain's judgement is most critical when conditions are ambiguous and no clear guidance is provided in manuals, checklists, or company policy.

This research project has been in progress for several years and has developed a model of factors that make flight decision efforts difficult, along with a taxonomy of decision problems. This model incorporated the nature of the problem, the time available to solve the problem and the risks associated with the options. This work was based on the analyses of crews facing abnormal and emergency events in full motion simulators and further supported by NTSB accident analyses and ASRS reports. For example, analyses revealed differences in the importance of risk and time pressure between Captains and First Officers. For Captains, risk was the most important variable in the decision strategy, while First Officers viewed time pressure as most important. As might be expected, experience does affect decision strategies. Experienced crew members are more sensitive to available decision time, better at handling uncertainty and have better prioritization of information needs. Based on these results it is suggested that to encourage effective decision strategies, crews are trained to : (1) understand and verify the problem, (2) assess time and risk, (3) rush to judgement, (4) consider constraints and consequences of options so as not to oversimplify the problem, (5) use "worse case" instead of " best case" reasoning, (6) plan for contingencies, and (7) manage the situation by setting priorities, assigning tasks and using all available resources.

Presently, this research team is developing decision making training scenarios for major carriers to validate during recurrent training. This research project focuses on designing and validating assessment methods and scales for the decision making scenarios. This research team is developing training guidelines to manage risk in ambiguous situations. The error taxonomy data base is being incorporated into the Rapid Reconfigurable LOE project described earlier. This research team is working with four major carriers to develop "challenging" training guidelines.

**Products:** Decision making training and assessment guidelines for air carriers to incorporate into their CRM training or into AQP will be published in 1998. The "challenging" training guidelines will be published in 1998.

**Project Title:** Interruptions, Distractions and Lapses of Attention in the Cockpit

**Performing Agency:** NASA Ames

**Background:** Interruptions and distractions are one of the most common causes of pilot error incidents and have contributed to many accidents. One of the most striking of these accidents was the Eastern L1011 accident in the Everglades. Glass cockpits, rather than attenuating this problem, have if anything amplified it. Closely intertwined with the issue of interruptions and distractions are problems with habit capture, tunneling of attention and failing to remember to perform deferred actions.

**Project Description:** This study will analyze ASRS incident data and NTSB reports, perform field observations of LOEs and LOFTs and will conduct experimental research to probe the cognitive mechanism pilots use to interweave concurrent tasks and manage attention in dynamic response to competing demands of the overall job. This study will characterize types of interruptions, types of situations conducive to distractions and factors that impede recovery from distractions. From this data specific training guidelines will be developed to aid crews in controlling interruptions, recover from distraction, avoid habit capture and prevent tunneling of attention. Also, this study will provide methods to design interruptions and distractions into simulator training to realistically challenge crew's task management skills and ways to modify cockpit SOP to reduce this form of crew error. A review of the literature is underway as well as the analysis of the ASRS and NTSB data.

**Products:** Training guidelines and scenario development methodology.

**Project Title:** Establishing Relationships Between Flight Data Parameter Values and AQP Qualification Standards using APMS methodology.

**Performing Agency:** JIL Information Systems

**Background:** The Automated Performance Measurement System (APMS) project has as its overall goal the development of tools and techniques to improve the efficiency with which air carriers collect and analyze Flight Operational Quality Assurance (FOQA) data. That overall project is not described in this report, details can be obtained from FAA Flight Standards, AFS-230. This present project is closely related to the overall APMS project in that this present project will be using APMS developed tools to analyze simulator data for flight parameters.

The goal of this project is to develop methods and techniques within the APMS research program that will allow air carriers to determine the extent to which daily flight operations are consistent with training qualification standards established in an AQP program. Thus, the overall APMS project will develop the tools necessary to analyze FOQA data and this project will identify flight parameters and qualification standards links. The goals of this project are: (1) To develop a "string" of flight data parameters and conditions that together would identify each AQP Qualification Standard when used in a FOQA database query. (2) To develop a system to acquire and record simulator flight data at the same level of fidelity and resolution as an FOQA equipped aircraft and (3) To develop a baseline flight data parameter profile of each AQP Qualification Standard for use in evaluating aggregate performance of these tasks in line operations based on FOQA data.

The software to collect the simulator data has been developed and is being used by a major carrier. The research team is presently collecting data on the flight parameters.

**Products:** Software tools and methodology to link FOQA data and AQP qualification standards.