Air Transportation Human Factors Research Review

June 2002

Eleana S. Edens, Ph. D.
Advanced Qualification Program Office AFS-230
Federal Aviation Administration

Table of Contents

The Federal Aviation Administration (FAA) Advanced Qualification Program Office (AFS-230) directs the Air Transportation Human Factors research program. This program centers on methods to enhance air carrier safety by providing the aviation community with guidance on pilot training systems, instructor evaluator education, and performance data collection and analysis. The following report summarizes the research performed in this program and provides the current status of each project.

Introduction

Although the basic technical and human factors concepts of crew performance are widely accepted, continuing efforts to improve methods for effective training, valid and reliable assessment of training programs and the appropriate methods to collect and analyze pilot generated safety reports and digital aircraft data remain at the forefront of aviation research needs.

The general research philosophy guiding efforts to reduce crew error through improved pilot training and assessment is that research must consider distinct segments of aviation training systems, including, but beyond the cockpit crew. 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) and Line-Oriented Evaluation (LOE) development, implementation and evaluation. Research must also consider ways for air carriers to employ pilot report safety data and aircraft digital data that will result in positive changes in line operations. Thus, the research centers on (1) crew training and assessment, (2) instructor training, (3) LOFT/LOE development strategies, (4) organizational and systematic influences on pilot performance, including automation usage, and (5) performance data gathered from aircraft and simulators as important components of safety enhancement.

Address questions or comments to:

Eleana S. Edens, Ph.D.
AFS-230
(703) 960-1085
eleanae@aol.com

Project Title
Establishing Relationships Between Flight Data Parameter Values and Advanced Qualification Program (AQP) Qualification Standards Using APMS Methodology.

Performing Agency
CSSI, Inc.; The University of New Mexico

Background
The overall goal of the Automated Performance Measurement System (APMS) program is the development of tools and techniques to improve the efficiency with which air carriers collect and analyze Flight Operational Quality Assurance (FOQA) data. Primary analysis techniques are being developed by NASA-Ames with FAA funding. That particular project is not described in this report, as the work does not focus, at this time, on training data but rather on operational data. Details can be obtained from FAA Flight Standards, AFS-230. The present project is closely related to the overall APMS project in that it will use some APMS-developed tools to analyze flight training simulator data for flight parameters as well as FOQA data.

Project Description
APMS is a research and development effort whose primary goal is to devise practical systems, methods and analysis techniques for air carriers in the use of digital flight data recorded during line operations. The current project is set in the Flight Crew Training Management and Support emphasis area, and seeks to develop tools and techniques to allow air carrier training departments to make use of digital flight data.

The primary task of the Flight Crew Training Management and Support emphasis area in this project is to develop analysis methods of flight data from line operations to assess the aggregate proficiency of the flight crew population, thereby determining which operational tasks and task components should receive more or less emphasis in continuing qualification training. In order to carry out these analyses, detailed formal standards of proficiency for flight crew operations are required. These standards are most clearly established at carriers who have completed the development of an AQP program. AQP development is based on a detailed task analysis of each flight crewmember's job, and results in a detailed set of qualification standards that are directly derived from job tasks. Accordingly, a major objective of the Flight Crew Training Management and Support emphasis area is the identification of flight data parameter profiles that predict levels of proficiency in AQP Qualification Standards.

The initial approach taken to establish these relationships is to gather flight data during simulator sessions and determine the statistical associations between trainee performance grades in the simulator and variations in simulator flight data parameters. Once these relationships are established and measured, they can serve as a starting point for researchers to study the relative level of proficiency with which these same qualification standards are performed during line operations.

Gathering and analyzing flight data from both simulator and aircraft operations required that a number of preparatory tasks be carried out. Tasks included tool development, data collection, data analysis, report/product development, and project management. While work is ongoing in all product areas, promising results have been obtained in terms of predictive models.

Products

- Modification of the MD80 and 737 simulators at Alaska Airlines to support parameter data export (Software products)

- Data collection and conversion software (Software product)

- Data visualization and custom data snapshot archive software (Software product)

- Integration of Flight Training Management System (FTMS) grade data with simulator data collection files (Procedure product)

- Conversion of relevant AQP performance standards into parameter data transforms to provide search patterns for identification of events of interest in simulator and line operations data (Data table product)

- Predictive models based on simulator data and associated grades for evaluation of de-identified parallel line operations events (Software products, logic table and procedure products)

- Data quality tools (Analysis and procedure products)

Note: The above tasks have been accomplished and simulator data is currently being analyzed by members of The University of New Mexico research team. The research described in the next section, conducted by the University of New Mexico, is part of this project.

Project Title
Training and Assessing Air Crew Skills: Methods to Achieve Reliable and Valid Performance Data

Performing Agency
University of New Mexico, Albuquerque

Project 1– Modeling Simulator Flight Parameters to Predict Instructor/Evaluator Grades - Background
With the advent of FOQA, air carriers collect an extensive range of flight parameter data pertaining to the operation and control of the aircraft during normal flight. These data, if appropriately analyzed, could provide carriers with an objective picture of how proficiently various critical maneuvers are being flown on a day-by-day basis. The goal of this project is to model instructor/evaluator's (I/E) grading of critical maneuvers on the basis of simulator flight parameter data and then apply these models to grade aggregated FOQA data.

Project 1 Description
Flight parameter data from MD 80 Simulators has been collected on Rejected Take-offs, V1-cuts, Normal Landings and Normal Take-offs. The focus has been on Normal Landings, where 492 flights have been analyzed. One hundred twenty different transformations of the Landing flight parameter data have been defined. These were analyzed by the Classification Tree algorithm. The results show that the model correctly classified 79% of I/E grades in training mode (the data set used to train or build the optimal tree) and 76% of grades of test mode (applying the model to a new set of data).

Project 1 On-going Research
- Investigating variations of the Classification Tree algorithm that may improve the performance of the landing model.

- Increasing Sample Size of the Landings. Because the model may be influenced by contextual factors (e.g., airport) a larger sample may be needed to isolate these specific effects.

- Isolation of Crew Resource Management (CRM) effects. Using flight animation software to model I/E grades that are based exclusively on physical parameter data and not confounded by CRM effects on grades.

- Modeling of other Maneuvers. Extending modeling efforts to normal take-off, rejected take-off, and V1 cuts.

- Application of Model to FOQA data. Given models are sufficiently predictive of simulator-based grades; extend the models to FOQA data.

- Obtain I/E grading of FOQA based animation data. Have I/Es grade flight visualization representations of simulator flights to be used as training tools for I/Es and debriefing crews.

Project 2 – Assessment of CRM Sub-Skill Grading - Background
The purpose of this project is to empirically determine whether it is possible to assess specific CRM sub-skills.

Project 2 Description/Completed Research

The research team has completed two studies, one based on LOE grades from a major carrier, and the second based on a large sample of line performance grades involving five different air carriers. In the case of the LOE-based study, it was assumed that OBs grades were measuring five different CRM sub-skills. The results from this study showed no support for this model, and subsequent factor analysis supported a one- factor model. In the case of the line data, there were 10 different measures that were assumed to assess two factors (team and task). While the results showed a significant sub-skill effect, only 4.5% of the variance in grades could be attributed to the team versus task factors. Subsequent factor analysis revealed an alternative two-factor model that provided a better fit to the data. The inability of grades to discriminate specific CRM sub-skills could reflect: (1) various methodological problems (e.g., I/E grading practices, the content of the test items, or the manner in which the tests were administered); or (2) false assumptions regarding the structure of CRM (e.g., there may be only one CRM factor).

Project 2 On-Going Research

- Discrimination of CRM Sub-Skills. During I/E calibration, provide a direct test of whether I/Es are able to identify the CRM sub-skill that was responsible for a specific technical failure.

- Assessment of IE's Knowledge of CRM. Administer a paper-and-pencil test of I/E's knowledge of qualification standards underlying CRM.

- Improving I/E Discrimination of Performance Levels. This will involve changes both in grading instrument (e.g., increasing the number of passing levels on the grade scale) and more fine-grained discrimination training of I/Es using *Flight Viz*© animations.

Project 3 – Evaluation of Instructor/Evaluator Training and Calibration Tool (IETC) - Background

The research team has completed the development of the group and individual versions of the Instructor/Evaluator Training and Calibration software, and can now evaluate its effectiveness in the training and assessment of I/Es.

Project 3 Description/On-Going Research

The following questions regarding the effectiveness of IETC-based training will be investigated:

- Does it improve calibration performance?

- What is the duration of any training effects it produces on calibration performance?

- Does the training generalize beyond the specific content of the videos used in training?

- Do the training effects generalize to LOE and Critical Maneuvers grading in the simulator, and if so, what is the duration of the effects?

Project 4 - Evaluation of Skill Maintenance and Reacquisition Training (SMART) - Background

This project is designed to: (1) determine the rate at which different critical skills decay over time; (2) determine what factors (e.g., pilot demographics, training conditions, etc.) influence the

rate of decay; and (3) develop efficient refresher training models that maintain skills for longer intervals

Project 4 Description/Completed Research

The research team studied the decay of critical skills from two air carriers involving a total of three fleets. In a Boeing 747 fleet involving a large sample size, they compared six- and 12-month intervals and found significant decay for all maneuvers that were measured. Looking at a much smaller sample from a Boeing 777 fleet with the same air carrier, they compared nine- and 12-month intervals and found no significant decay for any of the skills. Finally, with a second air carrier, they looked at intervals ranging from one - to 12-months. Here, they found a significant decline comparing one- to three- month intervals with four- to nine-month intervals for first officers, but no decline for captains when comparing one- to four-month intervals with five- to12-month intervals.

Project 4 On-Going Research

Study 1. Locus of Decrements. The plan is to continue to collect performance data from two major air carriers, and continue to recruit other air carriers who may wish to participate in this study. As the sample size continues to increase, it will allow the team to determine what factors are most closely related to rapid decay rates (e.g., pilot demographics, type of maneuvers, fleet, etc.). If the results continue to show decrements in regularly practiced maneuvers, researchers will determine whether the decay was related to simulator familiarity.

Study 2. Refresher Training. Assuming the results continue to show significant decrements in performance over intervals of less than 12 months, researchers will direct their efforts toward the development and assessment of low fidelity training procedures (e.g., beginning with methods as simple as having pilots image the maneuver before being evaluated). The first stage of this effort would focus on gathering information from trainers and the training literature to determine what types of training procedures are most likely to be effective. They will begin with focused small sample studies to determine which methods appear most promising.

Products

- Training guidelines
- Analysis models
- Technical research reports

Project Title
Pilot Training and Evaluation: Flight Simulator Fidelity - The Effect of Platform Motion

Performing Agency
Volpe Center, Cambridge, MA

Background
This project is part of the FAA initiative which promotes the availability and affordability of effective flight simulators for all U.S. airlines. Simulators provide a safe and effective means for pilot training and evaluation, enabling presentation of scenarios including emergencies requiring both technical and crew resource management skills. Therefore, the FAA is proposing a rule that would mandate the use of simulators for all air carrier flight crew training and qualification, limiting the use of the aircraft itself as a training option even for small regional airlines. However, there is a lack of sound scientific data on the relationship between certain key training device features, such as platform motion cuing, and their effect on the transfer of performance to and from the airplane. The goal of this project is to provide a scientific basis to ensure that FAA requirements are commensurate with safety objectives. Specifically, it addresses the question of the need for simulator motion for commuter airline pilot recurrent training and evaluation in the presence of a state-of-the-art visual system. The resulting data will also help the FAA to evaluate air carrier proposals for the alternative use of full flight simulators, whose availability and affordability may be limited especially for small regional airlines, and other training equipment.

In the past, technical constraints naturally limited the level of fidelity of a simulation. Today, however, technical capabilities have expanded to a point where they may enable a degree of fidelity that may exceed the one required for a particular purpose. This may lead to a situation where the benefit resulting from increased fidelity no longer justifies its cost. The focus thus needs to shift from ever more sophisticated technology to the level of fidelity required to train and evaluate to a specific safety standard. One of the common misconceptions is that the higher the physical fidelity of the simulation the better the training will be. The level of fidelity of the simulator should be determined, however, by the level needed to support the learning or evaluation of the tasks that will be trained or evaluated using the device. Any other approach would unnecessarily preclude the less affluent sectors of the aviation community from the benefits of using simulators.

Project Description
This project consists of four parts: (1) the collection of subject matter expert opinion, (2) a literature review, (3) empirical research, and (4) validation and generalization of the results of the research. For the first part, two workshops were conducted with experts from the FAA, industry, and academia. The first workshop focused on the aeromodel validation standards used in the flight simulators and the second focused on the motion requirements for simulators. These workshops led to proposed changes to the regulations for Level B simulators (the minimum level of simulators required for recurrent training).

The second part of this project is an extensive literature review. The relevant literature has been organized into a large electronic database, which is constantly updated to incorporate new information. Currently, the database contains 450 references, most of them summarized and annotated.

The third part is the execution of research addressing the need for simulator motion in recurrent pilot training. The research empirically examined the effect of FAA qualified Level C six degree-

of-freedom synergistic motion in the presence of a wide-angle high quality visual system on pilot training and evaluation. Transfer of skills acquired in the simulator to the airplane was measured by comparing the effect of training received in the simulator, with and without motion, on performance and behavior in the simulator with motion. This "quasi-transfer" to the simulator with motion as a stand-in for the airplane ensured the safety of the participants while allowing full experimental control. The research was conducted using regional airline pilots in recurrent training. Every effort was made to avoid deficiencies in the research design identified in a review of prior studies, by measuring pilot stimulation *and* response, testing both maneuvers and pilots that are *diagnostic* of a need of motion, avoiding pilot and instructor *bias*, and ensuring sufficient statistical *power* to capture operationally relevant effects. Two test maneuvers were chosen as diagnostic for an effect of motion on pilot training and evaluation, namely, engine failures on take-off with either rejected take-off (RTO) or continued take-off (V_1 cut). These maneuvers minimally disrupted the host airline's training program while satisfying the criteria recommended in the literature as diagnostic for detection of a motion requirement. These criteria included 1) closed loop, to allow for motion to be part of the control feedback loop to the pilot; 2) unpredictable and asymmetric disturbance, to highlight an early alerting function of motion; 3) high gain and high thrust, to magnify any motion effects; 4) high workload, to increase the need for redundant cues such as provided by motion, out-the-window view, instruments and sound; and 5) short duration, to prevent pilots from adjusting to a lack of cues.

A vast amount of data in the form of instructor grades, instructor and crew opinions, and objective measurements was collected from the experiment and analyzed. This research has been completed and the results are available in two different formats (see below). The results indicate that the motion provided by the test simulator, which may or may not be typical of other FAA qualified Level C flight simulators, does not, in an operationally significant way, affect evaluation, training progress, or transfer of training acquired in the simulator with or without motion to the simulator with motion for the maneuvers and pilots tested. Two *caveats* have to be kept in mind, however. First, the current study used the simulator with motion as a stand-in for the airplane. However, because the crews trained with motion did not show an advantage over the crews trained without motion when transferring to the simulator with the same configuration, it is unlikely that they would have had a greater advantage transferring to the airplane. The second caveat is that the simulator used in this study may not have provided sufficient motion to be effective. The measurements indicate that although the roll and longitudinal accelerations produced by the motion system of the simulator followed the aircraft response fairly well considering the limitations inherent to all simulators, the lateral acceleration seemed to be lacking for the maneuvers tested (RTO and V_1 cut).

Work is continuing on the fourth part of the project, namely, validation and generalization of the results obtained in part three. Part three showed no effect of the platform motion of one simulator, which represented a twin turboprop airplane, on commuter airline pilot evaluation and training of engine failures on take-off in the presence of a wide field-of-view visual system. A first effort, in part four, consisted of soliciting data from other FAA qualified level C and D simulators to determine whether or not the motion used in the previous study is representative. A letter signed by the FAA National Simulator Program (NSP) Manager was sent to simulator operators requesting data from a random sample of 30 simulators. Data from nine simulators has been received and analyzed. Results from these data indicate that the acceleration performance of the turboprop simulator used in the study was typical. If more data is received, it will be analyzed and results will be included in the final report.

Efforts are underway in the planning of a new research study designed to validate, generalize, and extend the previous findings to a different simulator, maneuvers, and pilot population. An

interagency agreement with NASA-Ames Research Center has been finalized to share the cost of running the experiment on its FAA qualified CAE Level D Boeing 747-400 research simulator. The motion system is being optimized with the help of CAE personnel. Several candidate maneuvers have been identified and will be tested

Products (published)

- Transcript of the Joint FAA/Industry Symposium on Level B Airplane Simulator Aeromodel Validation Requirements, Washington Dulles Airport Hilton, March 13 - 14, 1996.

- Transcript of the Joint FAA/Industry Symposium on Level B Airplane Simulator Motion Requirements, Washington Dulles Airport Hilton, June 19 - 20, 1996.

- Proposed Changes to AC 120-40B submitted to FAA Simulator Program Management.

- Longridge, T., Ray, P., Boothe, E., Bürki-Cohen, J. (1996). Initiative Towards More Affordable Flight Simulators for U.S. Commuter Airline Training. In *Proceedings of the Royal Aeronautical Society Conference on Training - Lowering the Cost, Maintaining the Fidelity, 15-16 May 1996, London, UK,* 2.1-2.17.

- Bürki-Cohen, J., Soja, N. N., Longridge, T. (1998). Simulator Fidelity Requirements: The Case of Platform Motion. *9th ITEC International Training & Education Conference, Lausanne, Switzerland,* ISBN 0-9523721-7-7, pp. 216-231.

- Bürki-Cohen, J., Soja, N. N., Longridge, T. (1998). Simulator Platform Motion-The Need Revisited. *International Journal of Aviation Psychology,* 8 (3), 293-317.

- Bürki-Cohen, J., Soja, N.N., Go, T.H., Boothe, E.M., DiSario, R., Jo, Y.J.: Simulator Fidelity: The Effect of Platform Motion. Report No. DOT/FAA/RD-00/XX, May 2000.

- Bürki-Cohen, J., Boothe, E.M., Soja N.N., DiSario, R., Go, T., Longridge T.: Simulator Fidelity - The Effect of Platform Motion. In Proceedings of the *International Conference Flight Simulation—The Next Decade, Royal Aeronautical Society*, 10-12 May 2000, London, UK.

- Go, T.H., Bürki-Cohen, J., DiSario, R.M., Jo, Y.J.: Relationship Between Objective Measures of Pilot Performance/Behavior and Instructor Grades. Report No. DOT/FAA/RD-00/XX, June 2000.

- Go, T.H., Bürki-Cohen, J., Soja, N.N.: The Effect of Simulator Motion on Pilot Training and Evaluation. Paper accepted for publication in proceedings of AIAA Modeling and Simulation Technologies Conference, Denver, August 2000, in preparation.

- Longridge, Thomas, Burki-Cohen, Judith, Go, Tiauw, Kendra, Andrew (2001) Simulator Fidelity Requirements for Today's Airline Pilot Training. Proceedings of the 11th International Symposium on Aviation Psychology, Columbus, Ohio: The Ohio State University Press

- Burki-Cohen, Judith, Go, Tiauw H., Longridge, Thomas, Kendra, Andrew J. Fidelity considerations to Meet Today's Total Flight Simulator Training and Evaluation Needs. AIAA Modeling and Simulation Technology Conference, Montreal, August 2001.

Products (not published in proceedings)

- Presented preliminary results of investigation of Simulator *Fidelity Requirements: The Case of Platform Motion* to FAA AAR-100 Air Carrier Training Research Review (October 1998).

- Discussed investigation of flight simulator motion requirements with AIAA Simulation and Modeling Committee at NASA Ames: *Simulator Fidelity: The Effect of Platform Motion* (April 1999).

- Presented *Relationship Between Grades and Objective Data* to help flight operational quality assessment (FOQA) program with data evaluation at FAA AAR-100 Air Carrier Training Research Review (July 1999).

- Presented final results and conclusion on *Simulator Fidelity: The Effect of Platform Motion* to FAA sponsors, customers, participating airline, and training center (October 1999).

- Presented *Simulator Validation: The Case of Platform Motion* at Human Factors in Transportation 33rd Annual Workshop at Transportation Research Board 79th Annual Meeting (January 2000).

Project Title
Realistic Radio Communications Simulation (RRS) in Airline Pilot Training and Evaluation

Performing Agency
NASA-Ames and Volpe Center, Cambridge, MA

Background
Radio communications are an integral part of every flight and require not only procedural knowledge, but also increase the need for sound task management, situation assessment, and decision-making skills. In simulator training, however, the incorporation of air traffic control (ATC) and company communications is often left to an already overburdened instructor, who must attend to many other administrative, training, and evaluation tasks. The lack of realistic radio communications simulations (RRS) impoverishes the simulation environment and compromises the transfer of skills between simulator and airplane. Full skill transfer to and from the airplane is a critical issue, if simulator use for training and evaluation is to be mandated.

Project Description
Volpe and NASA-Ames collaborated on a feasibility study for the development and implementation of realistic radio communications (ATC and air carrier) in line operational simulations (LOS). The feasibility study consisted of three parts: (1) literature review related to the requirement for RRS, (2) field survey of instructors and evaluators on current radio communications simulation practices, and (3) review of emerging technologies supporting the use of RRS.

A summary of results follows. The literature provides both practical and theoretical support for the necessity of realistic radio communications to achieve a variety of training objectives; for example, situation awareness, crew resource and task management, as well as pilot/ATC communication and coordination. Realistic radio communications are very important to ensure that training in a simulator will generalize and transfer to the complexity of the real world. A realistic radio communications environment avoids the negative training acquired in an impoverished environment leading to tunnel vision.

The accident and incident literature show that radio communications are a major contributor to pilot workload and are a causal factor in cockpit confusion, task management and monitoring errors, and flight path deviations. Inadequate ATC service has been identified by the Flight Safety Foundation Approach and Landing Accident Reduction Task Force as a factor in a third of all approach and landing accidents and incidents (Khatwa and Helmreich, Flight Safety Digest, November 1999). Crews need to be trained to recognize and challenge overly demanding ATC instructions such as last minute runway changes which may have contributed to the December 1995 American Airlines Boeing 757 accident in Cali, Columbia. In general, crews need to practice effective communication with ATC, particularly during emergencies and when dealing with non-native English speaking ATC (or if they are non-native English speakers themselves).

The importance of realistic radio communications during full mission simulations has been clearly recognized by the FAA Advisory Circulars related to AQP and CRM, which underscore that coordination with air traffic control and company is an integral part of line operations and that frequency monitoring is important for maintaining traffic and weather situation awareness. The CRM process includes not only on-board personnel (flight and cabin crew), but also air traffic controllers and company dispatcher and maintenance personnel.

These opinions were widely shared by the 29 I/Es from 14 airlines participating in the survey of I/E opinion and airline current practices. I/Es consistently rated the presence of air traffic control communications to own aircraft in simulations as important, especially in the terminal environment. Company communications were given above medium importance ratings as well, but slightly lower than air traffic control. The highest importance for company communications was given to communications from cabin personnel and maintenance. I/Es indicated that realistic radio communications increase the effectiveness of training such skills as CRM, and dealing with new ATC procedures or non-routine ATC situations. Radio communications are also important to train and evaluate distraction management and situational awareness skills. For the latter, simulating communications between ATC and other aircraft, the so-called "party line," was considered especially important. One I/E commented that the "party line" was his "biggest concern, so that pilots are listening." Another stated that the "party line enables CRM elements such as workload and distraction to be assessed more effectively."

Several I/Es commented that radio communications are a "critical need for LOS realism." "There is a correlation between realism in communications and training effectiveness; better training results from more realism." A final comment from an I/E was that "[t]his is a subject near and dear to my heart. […] I believe the 'simulator mindset syndrome' must be f[r]ought with realism. How can we expect crews to 'treat the simulator like the aircraft' when the audio environment belies the condition so often?"

Despite the agreement between literature, advisory circulars, and I/E opinions, the review of current practices showed that radio communications are almost exclusively role-played by the already burdened I/E. This increases I/E workload and detracts from the I/Es main function of instructing and observing, "dividing his attention." I/Es reported that they spend about one quarter of their time and effort role-playing radio communications, regardless of whether the simulation is for training or evaluation purposes. This is about the same amount of effort and time as they spend running the simulator.

At the same time, pilot workload is diminished because the complexity of the radio communications provided is "all relative to the creative ability of the instructor," which additionally may lead to uneven training and evaluation across crews. "I/E [radio communication] is less than actual, therefore it reduces pilot workload" and "pilots just listen when the instructor keys the mike." One I/E commented that "company communication] s [are] not normally used, too time consuming." "Party line" communications are "very seldom used due to I/E workload" and because "none of our formal training documentation requires it."

Despite the best intentions of I/Es, then, "when the pilot trainee finally arrives in the 'real world,' he must add another component, which was not learned during training. This new component can really complicate line flying." This last quote of one of the survey participants shows that I/Es need support in providing realistic radio communications to remedy this situation. The question is how. One possibility is increased training of I/Es to provide more realistic radio communications. Given that the awareness of I/Es regarding the importance of providing realistic radio communications appears high, however, it is questionable that additional I/E training alone would be beneficial. A second option would be to use additional instructors to distribute the workload of realistically providing radio communications. In fact, three I/Es from one airline indicated that two instructors may share the communications workload occasionally, but only when instructing a three-person crew. In the literature, there has been mention of other airlines using multiple instructors to simulate ATC, company, and cabin personnel (Lyall, Vint, Niemczyk, Wilson and

Funk, December 1998, AAR-100). It appears, however, that personnel costs have prevented this from becoming a widespread practice.

Another option is to make use of technology to fully or partially automate the provision of ATC, company, and "party line" communications. The most complete such technology identified in the review is the United Airlines in-house development called Interactive Real Time Audio System or IRAS. It automates both communications to own aircraft and the "party line" for all flight phases. In response to a customer requirement that visual terminal environment traffic also be represented by audio, the simulator manufacturer (CAE) developed its Ground and Air Traffic Environment System (GATES). GATES, however, is limited to the terminal environment and communications between ATC and other aircraft. It doesn't provide any communications to own aircraft. Commercial off-the-shelf, more advanced as well as supporting technologies such as the Rapidly Reconfigurable Event-Set Based LOE Scenario Generator (RRLOE, see this report) and speech generation and recognition systems were also examined. Any solution, however, be it the low-tech solution of additional training or I/Es or investing in existing or future technologies, or using a combination, will come at a cost. Before airlines invest in RRS, they will require objective data on the cost/benefit trade-offs achieved by different levels and methods of simulating realistic radio communications for different training and evaluation purposes. Phase 1 of the RRS project is now completed. Follow-on activities will focus on the research needs identified above.

Volpe examined incidents reported in ASRS. It found that over 70% of incidents reported are related to factors involving radio communications. Following the feasibility study, NASA-Ames developed a test plan to conduct a systematic, comparative evaluation using its Boeing-747 simulator. Crew performance under realistic simulation of radio communication vs. traditional training conditions in which the instructor/evaluator plays the role of ATC were systematically controlled and analyzed. The purpose of this comparison was to evaluate the efficacy of ATC simulation in providing airline aircrew with a training and evaluation environment comparable to actual line operations. Specifically, researchers focused on the degree to which workload levels, particularly in crew coordination and communication, could be made comparable to operational levels. Ten two-person Boeing 747 crews (five in each condition for a between groups design) participated in a two-hour LOFT simulation (LAX-SFO) designed to elicit CRM and communication skills. In April 2001, the simulation runs were completed and data analysis was begun. To date, preliminary findings include the following: (1) Substantially higher instructor/ evaluator (I/E) rater reliability for RCS fidelity than for low RCS fidelity. This was true for most CRM ratings as well as for technical proficiency ratings; (2) High RCS fidelity also resulted in increased aircrew-initiated communications to ATC (and other sources) than in the lower RCS fidelity condition; (3) Increased airspace situation awareness was also demonstrated by aircrews with high RCS fidelity when compared to those in the low RCS fidelity condition. The analyses and a report will be completed. Plans are underway for a follow- on validation study to be conducted in cooperation with several major air carriers.

Product

- Bürki-Cohen, J., Kendra, A.J., Kanki, B.G., Lee, A.T.: Realistic Radio Communications in Pilot Simulator Training. Report No. DOT/FAA/RD-00/XX, Final Draft March 2000.

Project Title
Rapidly Reconfigurable Line Operational Evaluation (RRLOE) Generator

Performing Agency
FAA/UCF Partnership for Aviation Team Training Research at the University of Central Florida, Orlando, Florida

Background
LOE is a methodology used in AQP to assess trainee performance in scenarios that are designed to challenge the integration of both technical and crew resource management skills. It is employed in AQP for jeopardy grading purposes - i.e. LOE failures are formally reported as a matter of record to the FAA, and pilots who fail an LOE may not be returned to duty until deficiencies are remediated, and performance on a subsequent LOE is determined to be satisfactory. Ratings of pilot performance on all LOEs must be electronically submitted to the FAA AQP office.

LOEs must be approved by the local FAA office responsible for oversight of a given air carrier. The approval procedure has entailed an FAA review of each such gate-to-gate LOE scenario. That has required each LOE to be separately conceived, developed, and tested by the training organization, and then to be individually reviewed and approved by the FAA. Thus, the development of LOEs has been both costly and time consuming. As a result, training organizations typically have had only a limited number of LOEs available for evaluation, each of which was approved for only a limited time period. Repeated use of a small number of fixed LOE scenarios creates the potential for the LOE events to become known in advance by the trainees in the organization, thus reducing the validity of both individual trainee evaluation and fleet proficiency assurance. The FAA is concerned about this potential deficiency as it could impact the overall effectiveness of the program.

One way to improve the validity of the LOE methodology would be to develop a process that avoids the use of a small set of fixed scenarios over an extended period of training within an organization. If events that form the building blocks of a full LOE could be individually developed and approved by the FAA, then these events could be used as a database from which to assemble complete, unique LOEs. The training organization could rapidly build new LOE scenarios with desired events, without seeking FAA approval of each complete scenario. The specific content could be varied while controlling the general content and overall difficulty. This would lead to fair evaluations for all trainees on fresh, variable, and valid training scenarios. In addition, to the extent that a methodology could be developed for rapidly generating alternative events that target skill set deficiencies, FAA goals for improvements in the remediation strategy for retraining and rechecking pilot LOE failures in AQP could be achieved.

Project Description
The development of the RRLOE methodology is the focus of the current research project.

Realism. A first challenge in developing this type of methodology and associated tools has been to conduct a program of research to identify those aspects of LOE scenarios that are required to make LOEs a valid and realistic assessment situation. The RRLOE research project has approached this problem by integrating the results of modern knowledge elicitation techniques with innovative database development. The result of this process was a theoretical approach (i.e., the "Domino-Method," see Bowers, Jentsch, Baker, Prince, & Salas, 1997) that became the basis for the existing expert system that tracks realism and continuity in the RRLOE program. This

novel approach allows the addition of individual event sets to a library of event sets for LOEs without the need to check the compatibility with other event sets manually. In fact, the resulting system can track more parameters relevant to a flight than a human LOE developer could in the past.

A second issue pursuant to realism is the impact of weather in LOEs. Recent research has demonstrated that weather is a prime consideration in aircrew decision-making (e.g., Jentsch, Irvin, & Bowers, 1997). However, weather has traditionally been only an afterthought in LOE development. Consequently, weather reports given to pilots in LOEs often were developed on the basis of small changes to existing weather paperwork and did not capture the complexities of the operational environment. To remedy this situation, the RRLOE research team conducted extensive research into the variables that affect pilot decision- making and have generated both a mechanism and a sample library of advanced weather scenarios that systematically vary those critical elements. It should be noted that this is the first time that theoretical research on the effects of weather on aviator decision-making has been made available to aviation training organizations in the form of a practical tool.

Difficulty. Another major research thrust of the RRLOE project relates to the estimation and measurement of LOE difficulty. Every LOE, notwithstanding whether it is generated by a human or a computer, must fall within an acceptable range of difficulty. A human LOE developer achieves this task largely through an experiential process that takes the combination of event sets, environmental conditions, and assessment expectations into account. Therefore, the challenge for the RRLOE project was to identify and describe the process used by humans and translate it into a mathematical model that the machine could execute. It was also important to validate the consequent model. The RRLOE research project has accomplished this through several studies at air carriers and has presented the results at international meetings (cf. Jentsch, Abbott, & Bowers, 1999). The research team now feels confident that the proposed methodology leads to a range of equivalent and fair scenarios. Again, this is the first time that a method for LOE difficulty assessment has been developed, validated, and described for the operational community.

Human factors aspects of the software. Finally, the results of the research and development effort described above needed to be translated into a user-friendly tool set, so as to facilitate transfer of the research results to the operational arena. In developing this tool set, the research team has had to study and apply human factors principles related to airline culture, operational environment, and computer sophistication across a wide range of operators participating in AQP. The resulting software tool set includes both tools for operators that are just beginning the AQP process and those which are far advanced within it. Finally, following the skill-based training approach described above, the project had to integrate existing task and skill analyses with the goals and tools of RRLOE. This required the consequent application of HF guidelines in Computer-Human Interface (CHI) design to their program. Further, through research at various airlines, the RRLOE research team was able to establish a standardized way to describe event sets that can be used by all participants in the AQP process.

The following research needs were identified for follow-on development of the RRLOE software:

Automation. The proliferation of high-technology systems in aviation has important implications impending on instruction. Specifically, future training for high-technology aircraft will not only require the identification of new knowledge and skills, but also the study of how the interaction with automated systems impacts the training and maintenance of traditional pilot skills. Furthermore, there is a need to create training opportunities and materials that take the technical capabilities of simulators and training devices into account. Both the research

methodologies and the software tools created under the RRLOE effort can now be employed to significantly advance the development of simulation scenarios for automation training. Indeed, it would appear that the only means of assuring that the decision-making skills required for the selection of an appropriate level of automation for a given set of flight conditions is through the systematic development of scenarios that specifically target such skills.

Since the distribution of the original software, four software updates have been produced. These include expanded worldwide functionality, functionality for long-range flight, and improved functionality for use with automated aircraft. Currently, 37 organizations (i.e. airlines, training centers, aircraft manufactures, pilot unions, and certification agencies) in the U.S. and overseas have received the RRLOE software and are receiving updates and software extensions. Requests from non-U.S. organizations were received from Transport Canada, Air Canada, Air Canada Regional, Air Nova, Air Ontario, Air BC, Ansett Australia, Quantas, Lan Chile, Ladeco (Chile) and Airbus Industrie.

Measurement. As described above, valid measurement of aircrew skills is the cornerstone of the AQP Program. The development of tools to accommodate consideration of measurement requirements during the LOE scenario development process itself could significantly enhance the precision of scenario based pilot assessment in the final product. While the groundwork for such an approach has been established by virtue of the RRLOE work accomplished to date, further RE&D to explicitly embed tools validated for that purpose is planned.

Products/References

- Research Reports Describing Difficulty Research
- Methodology for Generating, Cataloguing, and Assembling LOS Event Sets
- Event Set Library
- RRLOE Software and Documentation
- RRLOE Training Course

- Bowers, C., Jentsch, F., Baker, D., Prince, C., & Salas, E. (1997). Rapidly Reconfigurable event-set based line operational evaluation scenarios. Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting, Albuquerque, NM (pp. 912-915). Santa Monica, CA: Human Factors and Ergonomics Society.

- Jentsch, F., Abbott, D., & Bowers, C. (1999). Do three easy tasks make one difficult one? Studying the perceived difficulty of simulation scenarios. Proceedings of the Tenth International Symposium on Aviation Psychology. Columbus: The Ohio State University.

- Jentsch, F., Irvin, J., & Bowers, C. (1997). Differences in situation assessment between experts and prospective first officers. Proceedings of the Ninth International Symposium on Aviation Psychology (pp. 1228-1232). Columbus: The Ohio State University.

Project Title
Training for Unexpected Events

Performing Agency
FAA/UCF Partnership for Aviation Team Training Research at the University of Central Florida, Orlando, Florida

Background/Project Description
Unexpected events occur that surprise pilots and require that they respond to the event in novel and creative ways. Given the low likelihood of many failures that lead to the unexpected event and the high cost of training, it is realized that training for all possible situations is not only impractical but also likely to be impossible. Therefore, research is needed on how traditional training approaches can be augmented to improve pilot ability to respond to an unexpected event.

This research project is presently being initiated. A literature review is underway to consider the concepts important to recognition of unexpected events. A concept that will be reviewed is "Probability of an event: The pilot's perception of the probability that an unexpected event will occur is not only determined by the actual probability of occurrence, but it is also based on several other criteria". Apart from the actual probability that the event will occur, these factors are hypothesized to be: (a) the number of times the event has occurred in the past; (b) the amount of attention the event received following its occurrence; (c) the time since the last occurrence; and (d) the amount of preparation that the crew had for the event. Another factor affecting the assessment of the event's probability is the availability heuristic which says that decision makers estimate the frequency of an event based on how easily examples of a similar event come to mind.

Surprise: One of the key cognitive processes involved in the handling of unexpected events is that of surprise. There are several key elements that lead to surprise. One of these elements is the lack of anticipation of an event; another is the perception that an event is occurring purely by chance and is uncontrollable.

Cognitive Rigidity: Cognitive Rigidity may result from mission rehearsal that is too realistic. Research has shown that there can be too much, and too immersive mission planning, creating the illusion that the pilots already have experienced an event and know exactly what the correct response is.

Process Loss in Teams: Research has shown that the number of novel or alternative courses of action that a group or team of individuals can generate is lower than if the members work separately to conceive novel solutions.

When the literature review is complete, a research plan will be developed to investigate training methodologies for the unexpected event.

Products
- Training guidelines

Project Title
Human Factors of NOTAMS (Notice to Airmen)

Performing Agency
FAA/UCF partnership for Aviation Team Training Research at the University of Central Florida, Orlando.

Project Description
NOTAMs are used to convey critical air traffic control, air space, and airway facilities information to pilots. NOTAMS are disseminated electronically and in paper format and are available to pilots during their preflight preparations. The NOTAM system has been in use nationally (FAA NOTAMs) and internationally (ICAO NOTAMs) for an extended period of time. Some have alleged that the current NOTAM system is outdated and not very user friendly. Specifically, there are a number of human factors issues in the current NOTAM system that need to be evaluated.

This project is currently being initiated. It will analyze NOTAMs and write human factors recommendations for improving the usability of the documents.

Product
Technical report

Project Title
Automation Skills and Training

Performing Agency
FAA/UCF Partnership for Aviation Team Training Research at the University of Central Florida, Orlando, Florida

Background
Recent advances in aircraft instrumentation and equipment have led to a new generation of advanced automated aircraft. The development of electronic Flight Management Systems (FMSs), in particular, has given flight crews very powerful tools to plan, execute, and control specific flight plans by automating navigation and performance calculations. Indeed, the latest generation of FMS with four-dimensional (4-D) guidance (i.e., along the three spatial axes, and in time) can conceivably conduct an entire automated flight from engine spool-up at takeoff to deceleration after landing.

While these technical advances have undoubtedly increased the capabilities of modern aircraft, they have also led to a change in the pilot's job. Previously, the pilot was largely "in-the-loop," that is, an active controller who received direct and immediate feedback from the aircraft about the effects of his/her control inputs. With today's modern FMS systems, however, the pilot's role has changed from an active controller to a more passive supervisor/monitor/administrator. System inputs made early in the flight, such as route entries, may have adverse consequences much later in the flight. Also, as with other computerized systems, the capabilities of the system are not always entirely transparent to the operator. In fact, recent studies have shown that pilots frequently are uncertain about a system's status and its future performance (Sarter and Woods, 1998): Questions such as "What is this thing doing now?" and "What is it going to do next?" are, unfortunately, frequently heard in today's modern flight decks.

Sometimes, such uncertainty can have deadly consequences: In the decade since the introduction of the second generation of automated aircraft (Airbus A320, Boeing 757/767), several fatal accidents have occurred in which flight crew uncertainty/error in the operation of advanced autoflight systems has been cited as a causal or contributing factor (e.g., Boeing 757 at Cali, Colombia; Airbus A320 at Strasbourg) (Billings, 1997).

Project Description
Given the complexity of modern automated aircraft systems and the potential threats that inappropriate flight crew understanding of the systems pose, training for flight crews has become even more important in these aircraft than in traditional aircraft. This has been recognized by the FAA and has been the motivating force behind this research project. A structured analysis of air carrier needs with respect to automation training (Abbott et al., 1996) confirmed earlier FAA and industry reviews which showed that the identification of the skills needed to operate automated aircraft successfully is one of the key issues that need to be studied. The research also responds to the need for better airline-level tools by developing methods for scenario generation and performance measurement that specifically target cognitive skills for automated aircraft in the air carrier environment.

This research is dedicated to the appropriate assessment of automation skills. The eventual goal of this automation research is to create a 'toolbox" for use by training departments in assessing whether trainees truly understand the automated system. The toolbox is comprised of several different measurement products, each designed to tap a critical aspect of automation ability.

Obviously, an important aspect of auto-flight skill is knowledge about the system and how it works. Thus, one part of the toolbox will be a set of assessment tools to measure this knowledge. However, it is also important that operators organize their knowledge of auto-flight systems such that it supports accurate inferences. Training guidelines will be developed to target this issue. The development of new tools to measure these knowledge structures has been completed and is being validated at four regional carriers.

In addition to knowing the facts about automated systems, it is important that the operator be able to execute a complex sequence of activities to bring about desired outcomes. Operators usually acquire this "procedural knowledge" through practicing on low-fidelity trainers. However, it is difficult to capture performance on these trainers so that effective feedback can be provided to facilitate learning. In response to this problem, development of a software tool that will capture trainee inputs is under way. By collecting these data, comparisons of the trainee's performance to pre-determined sequences will provide rapid feedback about procedural knowledge deficiencies. Guidance on how to transfer this technology to other low fidelity trainers is being developed.

After acquiring basic knowledge of automated systems and their operation, trainees must be able to use these skills in realistic flight contexts. Consequently, a final aspect of the autoflight skill measurement toolbox includes modifications to the RRLOE that will allow the generation of useful training scenarios. The scenarios will be used to determine whether the trainee can use the automation effectively in specific situations that have been shown to be troublesome to novices.

Overall, the toolbox offers substantial advantages over current assessment methods. It allows the assessment of knowledge through several levels of acquisition, it considers both learning and application, and it promotes assessment of autoflight skills in the broader operational context. Furthermore, it offers a range of data that can be used to diagnose and remediate shortcomings in automation training programs.

Products

Presentations at Professional Meetings

- Hitt II, J.M., & Jentsch, F. (1999). Differences in general and commercial aviation automation. Paper presentation at the 18th Digital Avionics Systems Conference, Oct. 23-29, 1999, St. Louis, MO.

- Hitt II, J.M., Jentsch, F., Bowers, C., & Oser (under review). An information-processing model approach to identifying training issues for advanced aviation automation. Paper proposal for the Human Performance, Situation Awareness, and Automation: User-centered Design for the New Millenium, October 15-19, 2000, Savannah, GA.

- Jentsch, F., Hitt II, J.M., Agliata, D., Milham, L., Evans, B., & Bowers, C. (under review). Theoretical limits of card sorting as a mental models assessment tool. Paper proposal for the Human Performance, Situation Awareness, and Automation: User-centered Design for the New Millenium, October 15-19, 2000, Savannah, GA.

- Jentsch, F., Hitt II, J.M., Bowers, C., & Salas, E. (Accepted, 2000). Do pilots know how they manually control the vertical flight path? Studying conceptual knowledge of vertical flight path guidance. Paper accepted for presentation at the APA Annual Meeting, 2000.

- Milham, L., Jentsch, F., & Bowers, C. (Accepted, 2000). Understanding: Comparisons between assessment methods of knowledge structures. Paper accepted for presentation at the

APA Annual Meeting, 2000.

Technical Reports and Project Memoranda

- Hitt II, J.M. (1999). Reference list/bibliography for automation-related publications. Internal Team Performance Laboratory Technical Report compiled for the FAA Automation Training Team. Can be found at http://www.researchintegrations.com/ Note: password protected. Orlando: University of Central Florida

- Jentsch, F. (2000). Aviation Safety Reporting System (ASRS) - Air Carrier Report Search: Narratives for the 357 reports mentioning "VNAV". Technical Memorandum for the FAA Air Carrier Automation Skills and Measures Team. Orlando: University of Central Florida.

- Jentsch, F., Hitt II, J.M., Bowers, C., & Oser, R. (in preparation). Identifying training areas for advanced automated aircraft: Application of an information-processing model. Interim Technical and Scientific Report to the Naval Air Warfare Center Training Systems Division and the Federal Aviation Administration. Orlando: University of Central Florida.

- Jentsch, F., Hitt II, J.M., Bowers, C., Oser, R., & Salas, E. (1999). Do pilots know how they manually control the vertical flight path? Studying conceptual knowledge structures of vertical flight path guidance. Research White Paper for the FAA Automation Research Team. Orlando: University of Central Florida.

Airline-level Products

- VNAV white paper and assessment vignettes
- Automation knowledge assessment questionnaire
- Automation question catalogue
- Draft of a technical memorandum for the creation of automation vignettes and event sets
- Prototype software for data collection in PC-based flight management systems trainers

Project Title
Analysis and Proceduralization of Pilot Skills for Automated Flight Decks

Performing Agency
George Mason University

Background
The potential relationship between crew resource management (CRM) and flight deck automation management has been evident for some time. A number of carriers have defined specific automation skills as well as CRM skills, but a common framework that can be used to coordinate research in these two related areas has not been established. This research project is divided into two separate but related research efforts. Part 1 extends the use of Advanced Qualification Program (AQP) databases from previous research to include automation topics. Part 2 is a detailed analysis of cognitive skills and procedures related to specific crew performance problems in automated flight decks. These two areas overlap where the information in AQP databases is used to assess crew automation performance. Two research efforts will be pursued concurrently.

Project Description
This research is designed to provide the framework to integrate and extend what was learned in previous work about proceduralized CRM to the detailed procedures and cognitive skills relevant to flight deck automation. Establishing this framework provides potential common ground for research on CRM and team skills to manage automation. More importantly for the air carriers, this framework should also provide a path for applying the relevant research results to the development of better assessment of crew automation use and new operational procedures or crew training for automation use.

Analyses of pilot cognition and teamwork in the context of automated flight deck systems that may have a substantial effect on crew performance are under way. The research team has established a data collection and analysis foundation at a major carrier, and the team is using that base to develop and evaluate ways of assessing automation performance. A unique combination of cognitive modeling and performance assessment to help carriers more precisely evaluate and change automation procedures and training strategies is being used. In addition, a cognitive model of automation use for several focal automation problems is being developed. This model will be used to develop assessment items and procedures for LOFT/LOE evaluations of automation use that will be evaluated for scientific reliability and validity.

Part 1: Database
To date, some Performance Proficiency Database (PPDB) information has been received. Path analyses have shown that there is significant structural validity for the evaluation process. The observable behaviors and task skills predict the Technical and CRM ratings, and these in turn predict the Pilot in Command (PIC) and Second in Command (SIC) ratings. Compared to previous data analyses for the proceduralized CRM project, the strength of these relationships is slightly stronger.

In addition to using the PPDB to establish baseline reliability and validity, it has been used to study specific automation performance indicators in the PPDB data; however, that attempt has not been fruitful. Qualitative data analysis included assembling and coding automation-related performance comments into a taxonomy of errors. Consensus on coding was difficult, and the results did not seem to illuminate basic automation performance issues.

Quantitative data analysis has focused on the carrier PPDB LOE performance ratings for automation-relevant compared to non-automation event sets. Technical performance was found to be significantly different in the automation-related event sets. However, the CRM, PIC, and SIC ratings as well as specific observed behaviors and task skills were not significantly different in automation vs. non-automation event sets. Therefore, evidence for the scientific quality of current assessments of automation performance is not compelling. The sensitivity, reliability, and validity of these assessments may be increased by a more detailed analysis of pilot interactions with automation.

Part 2: Cognitive Modeling of Automation Use
The most common problems in automated flight reported in research and by trainers/evaluators are concerned with vertical navigation. Specifically, there seems to be a lack of understanding of the relationships among vertical speed mode (V/S), flight level change mode (FLCH), and the VNAV mode. Thus, the research team has developed a cognitive model focusing on the issues surrounding changes in vertical path during the climb and descent phases of flight.

Completed Research

In collaboration with major carriers, event sets were designed that emphasize vertical navigation. These event sets were used to model the interaction of pilots with cockpit automation in climb and descent. A desktop Boeing 747-400 flight simulator was used to gather verbal protocol and eye-tracking data for the cognitive modeling from pilot subjects. The data from this study have been used to collect basic information about how pilots track data in an automated cockpit. Specifically, the data suggest that pilots coming from a major airline do not use a common procedure to execute a descent using automated systems. . The cognitive model was coupled with the simulator to test alternative strategies for vertical navigation. The use of the model-simulator package uncovered a strategy that works well in the operational aviation environment. This package should integrate procedural and conceptual elements and guide better approaches to assessment of crew automation performance, development of automation procedures, or crew automation training.

Integrating the Database and Modeling Efforts
These efforts will lead to measurement and analysis of pilot performance data associated with flight deck automation in a high-fidelity simulator located at a major airline. This analysis will provide a baseline for current flight deck automation performance with current carrier training and provide information on critical measures or indexes of crew automation performance. The critical behaviors and processes identified by the model will be used to develop more precise and detailed measures of crew automation performance for the airline. The reliability and validity of these revised measures will be evaluated. This baseline can also be used by the carrier to compare the performance of pilots with revised or updated flight deck automation training against pilots without this new training. Similarly, the development of new procedures for crew automation use can be based on the problems identified by the cognitive model. These new procedures would be developed in a similar manner with the proceduralized CRM intervention developed and successfully evaluated under a previous FAA grant. Future work may evaluate the effects of either training interventions or automation procedures based on this model.

Products

- Training guidelines
- Procedures development guidelines

Project Title
Extending Proceduralized CRM to a Major Carrier

Performing Agency
George Mason University

Background/Project Description/On-Going Research
The evaluation study at a major airline was a field experiment designed with two distinct goals. The first goal was to examine whether the Advanced CRM (ACRM) results would be generalized to major airline pilots, equipment, and evaluation scenarios. This required a comparison of a set of crews training to use ACRM procedures integrated as Standard Operating Procedure (SOP) in their fleet FOM (integrated ACRM group) with a set of crews using the ordinary airline training and procedures (control group). The second goal was to examine whether the ACRM procedures could be effectively condensed into a separate, stand-alone checklist, rather than being integrated with all the details for each procedure in the FOM. To examine this, a stand-alone version of the ACRM procedures was constructed and used for a third set of crews (stand-alone ACRM group) in this study.

For both CRM and technical performance, the overall pattern is that the integrated ACRM group is superior to the normal group, which in turn is superior to the stand-alone ACRM group. These results confirmed the main expectation of this study - that the integrated ACRM procedures would facilitate crew performance. However, these results disconfirm the hope that a generic, stand-alone version of the ACRM procedures could have a positive effect. In fact, the stand-alone ACRM group is worse than either the normal or the integrated ACRM group. However, the data leave open the question of whether integrated ACRM procedures, combined with limited training, can also facilitate technical performance. Hopefully, this issue can be clarified by further analysis of the case-by-case data when they become available.

Product
- Technical report

Project Title
The Effects of Perceived Efficacy, Perceived Cohesion and Selected Non-Cognitive Individual Differences on Aviation Crew's Performance in a Dynamic Flight Simulator Task

Performing Agency
George Mason University

Background
Previous research suggests that group cohesion and efficacy are strong predictors of subsequent group processes and performance across a wide range of work-related groups. However, the commercial aviation domain's unique structural characteristics (i.e., short-term crew pairings, strong emphasis on standard operating procedures, and technology-constrained work environment) may inhibit their development. As a result, group cohesion and efficacy may be unrelated to pilot/crew processes and performance. To date, this issue has yet to be tested empirically.

Project Description
This project tested a model of the antecedents (e.g., personality characteristics, self-efficacy, and teamwork skills) outcomes, and moderators (e.g., task-related stress) of perceived cohesion and efficacy using a sample of commercial airline pilots. Retrospective questionnaires were linked with pilot performance ratings that were collected during a Line Oriented Evaluation. The participants included 118 pilots from the Boeing 737-300 (n = 63) and Boeing 757/767 (n = 55) fleets at an international air carrier. Preliminary results suggest that pilots who were high in conscientiousness, self-efficacy, and teamwork skills perceived their crews to be higher in cohesion and efficacy. However, despite the fact that the study had a sufficient level of statistical power, perceived crew cohesion and efficacy were unrelated to crew processes and performance. Moreover, these near-zero effects were not moderated by task-related stress.

Product
- Technical report

Project Title
Human Factors, Human Error and System Safety

Performing Agency
The University of Texas, Austin

Background
Current research extends investigation of the human factors of flight operations previously produced by this research group. The original research focused on Crew Resource Management (CRM) training as a method of optimizing crew performance and avoiding and managing operational threats and human error. CRM training focuses on issues that have been implicated as causal in the majority of air crashes and incidents. These investigations have demonstrated that CRM training is effective in changing flight crew attitudes and behavior, and by implication, system safety.

Despite demonstrated effectiveness, CRM has come under attack for its failure to eliminate all human error and to prevent accidents involving human error. Some of the attacks correctly reflect the fact that some organizations have provided minimal and outdated CRM training. Other criticisms demonstrate a lack of understanding of the nature of human error and the role of training interventions such as CRM, LOFT, and AQP in system safety.

Humans are cognitively limited and further subject to the deleterious effects of stress, fatigue, lapses and cultural biases. The effects of human limitations are exacerbated by complex operating environments that include difficult ATC interactions, increases in cockpit automation, and operation into difficult regions such as Latin America and Africa. Given the physical and psychological limitations of humans, the elimination of error is impossible. What can be achieved is the establishment of defenses against the occurrence and consequences of error by managing the threats that occur in daily flying. This is done by cockpit crew deployment of strategies to manage threat and error, and mitigate the consequences of error. In this context, CRM represents a training architecture for developing a set of tools to be used as countermeasures and error management tools. Focusing on error management and threat recognition also has the advantage of basing training on superordinate goals that are globally accepted. However, much remains to be learned about the extent and nature of human error and threat in the current aviation system.

A focus on system issues (latent system threats) and using CRM as a training intervention for threat and error management contrasts with two increasingly common responses to human error accidents and incidents. The first is to increase the level of automation to remove the human from the loop. As behavioral data and other research have shown clearly, automation is not a panacea for human limitations. Rather it adds a new set of errors and problems. The second approach is to proceduralize behavior defining Standard Operating Procedures (SOP) for pilots to prevent human lapses. This also is a strategy with limitations since data indicate that an over-emphasis on procedures can actually increase the incidence of errors and undesirable behaviors, as crewmembers may ignore highly complex procedural requirements. In other words, proceduralization may have the unintended consequence of increasing violations.

Factors such as national, professional and organizational cultures play a significant role in safety and in the acceptance of training programs such as human factors/CRM. Cultural factors further influence the use of flightdeck automation, which raises issues regarding the certification of new systems. It follows that attempts to increase the safety of flight operations must address broader

system issues as well as training at the individual and crew level. These efforts require accurate and extensive data on the way both individuals and organizations function within the aviation system to guide safety initiatives.

On-Going Research

The research program extends prior research in this area. It has eight components that address critical flight operations. These areas are:

- Development of a new conceptual model of threat and human error in flight operations. The model will be of practical value in understanding crew performance on the line, in developing training programs, and in the analysis of incidents and accidents. The model is driven by data collected from pilots flying regularly scheduled airline flights.

- Extending current data on performance in line operations to include data on environmental threats to safety, error, and responses to error. This approach (called LOSA – Line Operations Safety Audit) employs a team of expert observers from the jump seats of scheduled flights to collect systematic data on line operations ***without jeopardy.*** The research team collects common data in each organization, giving evaluation points that allow comparisons of performance across organizations. Organizations have used the findings to target specific issues to address in flight operations and during training. The specific data collected are guided by the conceptual model. LOSA has been endorsed by the International Civil Aviation Organization (ICAO) as a preferred means of assessing system safety.

- Developing a curriculum for organizations to train line audit observers. The curriculum should help organizations develop consistent internal audit methodologies.

- Assessing the use of automation in line operations. As part of the line audit methodology, a particular emphasis will be placed on the use of automation, including the number and nature of errors made in automation usage and documentation of the automation anomalies and threats that occur in normal line flying.

- Refining incident reporting and analysis methods. This component will develop and refine data collection methodology for incident reporting by flight crews with the goal of generating data that can be used more effectively for determination of trends and latent systemic threats. With the recent approval of the ASAP program, a unique opportunity for studying incidents has presented itself. The conceptual threat and error management model will aid in understanding the dynamics of causes of cockpit incidents – and of their effective and ineffective management. A complete ASAP reporting system had been developed for a major carrier. This system is based on the threat and error model. All pilots at this carrier have been trained on this concept and should be able to easily identify the threats that led to the errors that resulted in an ASAP report. Using this system, the carrier can link its FOQA data with its ASAP data. This system is available to other carriers wishing to design their ASAP on the treat and error model.

- Determining relationships among safety-related data sources. Different methods (incident reports, line audits, surveys, etc.) produce different data on organizational and system performance. In this phase of the project, relationships among data from these varied sources will be explored with the goal of improving the assessment of system performance and safety.

- Extending data on flight attendant CRM. The project has developed a survey instrument to

determine human factors attitudes and training needs for flight attendant CRM. The research will extend the database and make the findings available to organizations initiating CRM for this group. To date, survey data from several organizations have been collected showing significant differences from pilot responses.

- Developing tools for organizational CRM and safety surveys. Most U.S. major and many regional airlines have used the Flight Management Attitudes Questionnaire (FMAQ) to assess CRM curriculum needs and safety concerns. It has also been used by airlines in more than 20 foreign countries. This task will produce a semi-automated FMAQ self-administration package that airlines can use for data entry and reporting. The package contains norms from both US and international airlines.

Products

- Methodologies, data collection and analytic tools, and training products for use by airlines and other agencies.

- Technical reports and scientific papers are also presented on the project's home page, which is widely used nationally and internationally –www.psy.utexas.edu/psy/helmreich/nasaut.htm

- Dr. Robert Helmreich, the project's principle investigator, also serves on the ICAO human factors team providing worldwide instruction in human factors issues. This role helps disseminate research findings.

- The project also trains doctoral students who go on to work in aviation positions after receiving their degrees.

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 has proposed the concept of Advanced Qualification Program (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 carrier's 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 framework 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 with all other flight skills and procedures as required by each flight situation. One goal of AQP is that CRM skills should be utilized as a normal and inherent part of aircraft operation, little different from operating automated equipment or performing a proper checklist.

While AQP permits significant departures from the traditional Federal Aviation Regulation 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 FAA 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 CRM 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 a priority basis, with the explicit intention that it be refined subsequently as experience with the new program accumulates. There was a need for a methodology that, if faithfully followed, would produce an effective AQP.

Project Description/On-Going Research
The goals of this research project are to design and develop a task analysis and a performance audit database. Air carriers entering the AQP process will then use this model. 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 provides a useful and unique perspective. The ISD methodology, as currently implemented, focuses on 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 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 that are always performed by the crew for a given task or subtask. Traditional ISD 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 that fall outside 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 judgment: 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 judgment 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 perspectives to an AQP. Merging them into a coherent approach will support development of a complete training program. One means by which this integration can occur is to place the focus of flight training on situations instead of 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 information.

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 building blocks 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 Flight Training Device (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, is complete. Several workshops have been conducted to train air carrier personnel in the use of this model.

Development of the Advanced Model AQP (the task analysis component) has been completed and it is being distributed to air carriers. This second-generation model incorporates many features requested by the carriers, allowing each carrier to pursue its own level of development and will include a performance proficiency database. (To be delivered in FY 2002).

Products

- Model AQP and research report delineating process, methodology and lessons learned.
- Advanced Model AQP

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

Performing Agency
NASA-Ames Research Center

Background
Interruptions, distractions, and preoccupation with one task to the neglect of another task are among the most common causes of pilot error incidents and have contributed to many accidents. Rather than attenuating this problem, glass cockpits 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 is a collaborative project, co-sponsored with the NASA Aviation Safety Program. The project goals are: characterize the interruptions that most frequently occur, the types of situations conducive to distraction, and factors that impede recovery from distraction; identify specific techniques crews can use to control interruptions, recover from distraction, avoid habit capture and prevent tunneling of attention; identify ways to systematically design interruptions and distractions into LOFT/LOE scenarios to realistically challenge crews' task management skills; and explore ways to modify cockpit Standard Operating Procedures (SOP) to reduce this form of crew error.

To date, this project has analyzed ASRS incident data and NTSB reports of accidents attributed primarily to crew error. In roughly half of the accident reports, lapses of attention by the crew occurred. The review of ASRS data revealed that the most commonly neglected tasks involved lapses in monitoring (2/3 of reports), or lapses in memory, especially failure to remember to complete deferred actions (1/4 of reports). The competing tasks that distracted or preoccupied the crews were: communication (50%), "head-down" tasks (16%), abnormals (14%), searching for traffic (8%), and miscellaneous (12%).

Two questionnaire studies were developed and administered to senior captains, instructors, and check pilots. The first questionnaire elicited information on what techniques experienced pilots currently use to reduce their vulnerability to lapses in attention and memory. The second questionnaire probed the extent to which experienced pilots agree on the efficacy of these techniques and the practicality of using them in line operations. Preliminary data analysis suggests that some techniques are currently available, however, experienced pilots differ substantially in their opinions of the efficacy and practicality of these techniques.
Analysis of cognitive issues suggests that these lapses of attention and memory are not primarily the result of overload, but of difficulty in switching attention back in forth between concurrent tasks reliably in a timely manner. Laboratory studies are being conducted to determine the nature of this difficulty.

From this data, specific training guidelines will be developed to aid crews in controlling interruptions, recovering from distraction, avoiding habit capture and preventing tunneling of attention. Also, this research will provide methods to design interruptions and distractions into simulator training to realistically challenge crews' task management skills and ways to modify cockpit SOP to reduce this form of crew error.

A summary of progress to date was published in ASRS Directline. This article was quickly

reprinted in its entirety by several major flight safety magazines, including: Airline Pilot (ALPA), USAirways Safety on Line, Flight Safety (Canadian Airlines), Independence (UPS Pilots Association), and USAF Flying Safety. Several airlines and the U.S. Air Force Reserve have requested presentations on this project.

Products

- Training guidelines and scenario development methodology.

Project Title
Opportunities for and Vulnerabilities to Error in Everyday Flight Operations

Performing Agency
NASA-Ames Research Center

Background/Project Description/On-Going Research
Recent evidence (Dismukes, Young & Sumwalt, ASRS Directline, December 1998) suggests that interruptions and distraction in the cockpit are more prevalent and potentially more threatening to pilots than previously considered. Events such as momentary interruptions (e.g., from another crew member or form ATC), or the need to conduct concurrent activities (e.g., program a new departure on the FMC while taxiing) in response to operational demands often lead to lapses in attention that, in turn, contribute to errors. Manifestation of such errors is abundant in accident and incident reports. It is a fact that pilot performance errors occur on a daily basis, both on the ground and in the air, in the course of every day flight operations. Yet, interruptions and distractions are widely accepted as the "norm" in air carrier operations, and are significantly underestimated because their nature and effects are not well understood.

A NASA jump seat observer flew 50 legs on board the Boeing 737s of two US air carriers. A qualitative but detailed picture of every day, routine pre-flight and taxi-out phases was developed, with emphasis on the nature and timing of interruptions and distractions observed during these flights. This analysis was then contrasted to the procedures, flows, and checklists as described in the appropriate flight manuals and trained in the schoolhouse. Based on this task analysis, it was possible to illustrate the number and variety of distracting events that impose on pilots and threaten to detract them from their standard, habitual cockpit duties during the course of a typical flight. Unlike the sterile, linear presentation of operations in the training environment, the real-life situation is dynamic and fraught with distracting events of unpredictable source, nature and timing.

With a detailed description of line operations charted, aviation incident databases were examined for possible outcomes of interruptions and distraction. Evidence for errors was ample in reports of mistakes that happen without pilot' realization and despite their best, conscious efforts to follow Standard Operating Procedures. Pilots, as all skilled human operators, are particularly vulnerable when forced to interleave novel activities with habitual, well-practiced sequences of actions, and when engaged in making decisions about adding, shedding, and/or deferring actions. Memory, particularly that for deferred actions, side-tracking and preoccupation with interrupting events, automaticity, assumptions and expectations based on habit, are among the major cognitive themes that are derived from these reports. Interruptions and distractions, such as those observed in the course of the jump seat study, are exactly those events that cause pilots to interleave and add or shed activities, and so contribute greatly in creating opportunities for cognitive vulnerabilities which lead to errors.

Expecting to find ways to eliminate interruptions and distraction, or mitigate cognitive vulnerabilities is unrealistic. Observations from line operations, coupled with the kind of task analyses that come out of this research project, however, significantly contribute to a better understanding of the causes for pilot error and aid in design of robust procedures that offer multiple layers of defense against error. One air carrier has already used these findings to guide its effort to do the latter. These findings also suggest ways to enhance existing training programs so that they are more consistent with the realities of both routine line operations and the characteristics of their human element. Discussion of human vulnerabilities to operational

demands can be particularly useful in critically analyzing incident reports and moving away from dismissing errors as the result of complacency and/or fatigue. This research will be extended to include all phases of flight.

Product

- Technical report

Project Title
Operating Documents Human Factors Project

Performing Agency
NASA-Ames Research Center and University of New Mexico

Background
The original project, Integrating CRM into Crew Procedures, developed a systematic approach to integrating CRM procedures with normal and non-normal checklists. This emphasis on checklists is only part of the solution, because checklists are just one element of the whole operating document system. It became apparent that operating documents must show internal consistency across fleets (aircraft types, and route structures) and 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 approach has not been used systematically in the design, development, and implementation of operating documents.

Project Description
The 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 the development and implementation of document systems. There was strong industry interest and participation in two NASA/FAA Operating Documents Workshops. Based on the information provided in the surveys and in the workshops, 120 guidelines have been identified and modified for carrier personnel involved in the development of operating documents. Guidelines are organized into the major steps in the development and implementation process: Organization of Documents, Design of Documents, Developing and Maintaining Documents, and Transition to Electronic Media. A manual entitled "Developing Operating Documents" is in its final revision awaiting publication in hard copy as well as in an interactive on-line version.

Feedback from participating carriers indicated the need for ongoing operating documents workshops. To address that need, the research team sponsored the NASA/FAA Operating Documents Workshop III. This research effort has also identified the need for a database to help manage the large number of documents required to support flight operations. A Documents Database (DDB) is a database of the information topics, requirements, sources, users, review, distribution, and related data essential to the efficient management of an operating documents system. To further specify the content of a DDB, a subgroup of the NASA/FAA Operating Document Group has developed a prototype user interface. This user interface provides a concrete example and a usability-testing platform for how such a database would be structured and used. A DDB should ultimately be tailored to meet the carrier's individual needs while adhering to standards that would allow it to be linked to other operational and training databases. Remaining research challenges in this area include the identification of information topics as well as essential database fields to manage flight document systems.

Work on the prototype DDB user interface has led to a collaborative effort between the NASA, the FAA Operating Documents Group, and the Air Transport Association's Flight Operations Working Group (FOWG). These two groups have addressed similar operating document issues from different perspectives. The FOWG has been setting the standards for digital data exchange between manufacturers and carriers. FOWG members were able to provide review comments on

the DDB while learning about the user interface approach to requirements identification and refinement. In addition, the International Civil Aviation Organization (ICAO) intends to use the manual "Developing Operating Documents" to harmonize flight document development at the international level.

Products

- Operating Documents Workshop III
- Manual "Developing Operating Documents"
- Additional coordination and harmonization with ICAO and the ATA FOWG has been proposed and is currently under review at the FAA.

Project Title
Risk Assessment and Risk Management in Aviation Decision-Making

Performing Agency
NASA-Ames Research Center

Background/Project Description/On-Going Research
Analysis of air carrier accidents found that the majority of flight-crew related errors concern faulty decision-making. A common pattern was the crew's continuation with their original plan when conditions changed, suggesting that another course of action might be more prudent. Many accidents reflect what appeared to be inadequate risk assessment by the crew that led to an inappropriate decision. Additional factors that may have been working against safe decisions include organizational and social pressures, such as on-time departures and arrivals, fuel economy, and peer judgments about one's professional competence. Risk assessment feeds into a crews' decision-making in two ways: (1) during their assessment of the perceived threat that may require intervention, and (2) in their evaluation of potential responses to the perceived threat. Despite its important role in aviation decision-making, there is surprisingly little research on pilot's risk assessment. The research team has conducted two studies to examine the following issues: (1) what types of risks are of the greatest concern to pilots and affect their perceived difficulty in decision-making, and (2) how do they manage these risks?

In the first study, 119 pilots were asked about their every day experience with five types of risk: (1) economic (dollar cost to the company); (2) physical (safety); (3) productivity (disruption of the schedule; (4) professional (possible loss of job); and (5) social (other's impression of you). Analysis indicates that pilots are most concerned with threats to flight safety, which accounted for 79% of their responses. However, economic and productivity risks were the ones the pilots reported they encounter most often, while social and safety risks were rated as least often encountered.

Both captains and first officers mentioned that conflicts between safety-related and other types of goals, especially customer satisfaction, were most difficult to resolve. Lack of any good options also increased perceived decision difficulty. For captains, decision difficulty was associated with high levels of ambiguity and uncertainty in the decision situation. First officers, on the other hand, attributed greater decision difficulty to the possibility that a poor decision may have negative professional consequences.

A second study was conducted to examine how pilots respond to goal conflict, ambiguity, and uncertainty in dynamic flight situations. Two evolving decision scenarios were presented to 30 pilots who were asked to "think aloud" as they decided what to do. In both scenarios, continuing with the original plan (e.g., take-off) posed a threat to flight safety, but if successful would bring economic gains. Changes to the original plan, in turn, would lead to gains in terms of safety, but also economic or other losses. Across both scenarios, pilots chose the riskier (in terms of safety) option 50% of the time. Analysis of their "think-aloud" protocols indicated that pilots did not weigh the consequences of options, but rather tried to control the safety risk associated with their choices. Those who took the "riskier" options invented solutions that allowed them to minimize the safety risk while achieving their productivity goals. Those who took the conservative option basically avoided the safety risk.

These two studies represent initial efforts to understand the risks that pilots face and to address why they make decisions as they do. The inherently subjective nature of risk assessment is

evident in the differences between crew positions in the risk survey and in the differing interpretations of cues by pilots in the "think aloud" study. Subsequent research will seek to improve decision-making by developing effective risk management and decision-making strategies and support tools.

Product

- Technical report

Project Title
Airline Pilot Training Survey

Performing Agency
American Institutes for Research, Washington, DC

Background
The FAA is presently in the process of rewriting existing regulations that pertain to pilot training. The proposed revisions will change the requirements for traditional FAR Part 121 training, and they will codify the requirements of training conducted under SFAR 58, the Advanced Qualification Program (AQP).

A valuable source of information regarding the strengths and weaknesses of existing training, both Part 121 and AQP, are the pilots who have been the recipients of that training. The FAA believes that the quality of its decision making with regard to future changes in those regulations would benefit significantly from a thorough, unbiased assessment of pilot opinion regarding these training programs. Accordingly, the FAA issued a grant to the American Institutes for Research (AIR) to conduct a survey of airline pilots' perceptions of the effectiveness of their training. To ensure the success of this effort, AIR worked closely with the Air Line Pilots Association (ALPA), Independent Association of Continental Pilots (IACP), the Allied Pilots Association (APA), and the Air Transport Association (ATA) to develop the survey's content. The survey is truly a collaborative effort among the federal government, unions, and the air carrier industry. Each group is represented on a technical advisory board that provides oversight to the project. The survey was sent to a representative sample of pilots during summer 2000. The results of the survey will be published in a report prepared by AIR in consultation with the members of the project's technical advisory board. It is anticipated that this report will be of great interest to industry and the air carrier pilot community as a whole.

Project Description
A three-phased approach was proposed to meet the project's primary objective. During Phase I, a detailed research plan was developed. The research plan was derived from a review of important airline pilot training documents, the wider body of training effectiveness research, and consultation with industry and union experts. The research plan specified the pilot population to be studied, key subgroups that comprise the population, strategies for sampling the population, and important variables related to pilot experiences in and perceptions of their training. A survey instrument for collecting data on these variables was developed and administered during Phase II. Survey questions were developed based on the results of several focus groups conducted with airline pilots who have experience in AQP, Part 121, and Single-Visit training (i.e., an interim step in AQP development). The survey instrument was pilot-tested prior to its implementation in the main study. Once the survey was finalized, a large representative sample of airline pilots was identified and administered the survey. The airline pilot sample was drawn at random from a larger population of airline pilots using a stratified random sampling procedure. During Phase III, data collected from this sample was analyzed to identify pilot experiences in and perceptions of Part 121, AQP, and Single-Visit training.

Product
- The procedures employed in the investigation, the analyses conducted, and the resulting findings will be presented in a detailed FAA technical report. The survey is completed and presently under review by the technical board. It will be distributed in summer 2002.

Project Title
Developing a Classification System and Searchable Database for Analyzing Human Factors Issues in ASAP Reports.

Performing Agency
American Institutes for Research

Background
Several carriers have recently implemented Aviation Safety Action Programs (ASAP) that collect de-identified incident reports from line pilots regarding threats to safety. Most ASAP incident reporting forms include fill-in-the-blank or check-in-the-box items that describe the pilot who is submitting the report (e.g., seat position, flying time) and the flight conditions (e.g., weather, phase of flight) that immediately preceded the event. Space is also provided for a short narrative that describes the event, the causal factors that precipitated it, and suggestions for preventing recurrence. Unfortunately, many of the human factors issues described in ASAP reports are located exclusively within these narratives. As a result, researchers have found it difficult to identify statistical trends in the data.

Project Description
This project has three main goals. The first is to develop a taxonomy for classifying human factors issues in de-identified ASAP reports. The second is to embed this taxonomy within a searchable computer database. The third is to extend all or part of this research to the classification and analysis of de-identified AQP performance ratings and FOQA output. The end product will be three separate databases that can help carriers proactively identify threats to safety. The information generated from these databases can then be used by the carrier community in a variety of ways, such as identifying training objectives and developing LOS scenarios.

Researchers are currently reviewing existing aviation human factors taxonomies, accident/incident reporting systems, and data reporting tools. During the remainder of 2002, they intend to use the lessons learned from this literature review to help air carrier partners develop an ASAP human factors taxonomy. This taxonomy will go through several rounds of internal review, testing and revision. An independent panel of human factors experts will then review it. After incorporating their feedback, the research team will embed the taxonomy within a searchable database, develop a user interface, and create several prototype data analysis reports. Finally, they will develop a user manual and training materials to assist users. During 2003, the research team will extend all or part of the taxonomy to the classification and analysis of de-identified AQP performance ratings and FOQA output.

Product
- Beaubien, J.M., & Baker, D.P. (2002). *A Review of Selected Aviation Human Factors Taxonomies, Accident/Incident Reporting Systems, and Data Reporting Tools.* Washington, DC: The American Institutes for Research