Part I. Project Identification Information

1. Institution and Address:

The University of Texas at Austin
P.O. Box 7726
Austin, TX 78713-7726

2. FAA Program: Crew Human Factors

3. FAA Award Number: 92-G-017

4. Award period: 10/1/92 to 12/8/98 (note: research ended 12/8/98 although the grant period extended through September 25, 1999)

5. Cumulative Award Amount: $1,731,269

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

Part II. Summary of Completed Project:

A central focus of the work is training in teamwork, leadership, and communication known as Crew Resource Management (CRM) and associated full mission simulator training known as Line Oriented Flight Training (LOFT). Products of the research include methodologies for assessing team performance in simulator training and for evaluating participant reactions to CRM training. A set of behavioral markers was also developed to enable assessment of crew performance in both LOFT and line operations and a methodology (Line Operational Safety Audits: LOSA) developed to measure crew performance during line operations. These markers have become widely used by airlines as part of their training and measurement programs. In response to a need to tailor CRM training to specific airlines, a survey instrument, the Flight Management Attitudes Questionnaire (FMAQ) was developed. A major emphasis in the later years of the project is determining the effects of national, organizational, and professional culture on crew performance and safety in over 20 countries. An unexpected finding of the study was highly significant national differences in attitudes regarding the use of cockpit automation. An ongoing effort involves investigation of the team aspects of error by flight crews and air traffic control.

Part III Technical Information

a. Abstracts of Theses

1). Waller, Mary Joette, PhD. Multitasking in Work Groups: Coordination Processes in Work Groups with Multiple Tasks. The University of Texas at Austin, 1995.

Small groups research has focused on groups performing a single task. The model and two studies presented in this dissertation examine how groups manage multiple tasks under dynamic and deadline conditions. The model suggests that work groups engage in information gathering, task prioritisation, and resource allocation activities in order to perform multiple tasks. In the first study, field data were collected from 15 groups: five air traffic control crews, five airline ground operations crews, and five human physiology research teams. In the second study, field data and performance data were collected from 10 aircraft flight crews. Results support the descriptive model of group multitasking. Conditional likelihood logit analyses concerning the effects of multitasking behavior on group performance reveal significant
differences between high- and low-performing groups in both the frequency and timing of their multitasking activities. Additional analyses suggest nonmonotonic relationships between the frequency of resource allocation activities and overall group performance. Some resource allocation activities are associated with the occurrence of errors in group performance. The dissertation concludes with implications for both group process theory and future research. Specifically, the implications for theories of self-regulation and performance feedback effects in groups are discussed, as well as directions for future research designed to examine the causal directions of associations identified here.

2) Law, Jonathan Randolph, PhD. Rising to the Occasion: Foundations, Processes, and Outcomes of Emergent Leadership, The University of Texas at Austin, 2/96.

506 junior and senior undergraduate students from 214 intact project teams participated in a three month longitudinal investigation of emergent leadership and leadership effectiveness. Although the relationships between the self-assessed traits of dominance, self-monitoring, friendliness, and task orientation with leadership perceptions were all positive, the magnitude of these relationships was relatively modest. However, relationships between peer-assessed traits and the same leadership evaluations were much stronger than the self-assessed trait method. Although evidence suggested that the consistency between leader trait expectations and actual leader characteristics (i.e., implicit leadership theories: ILTs) were significantly related to leadership perceptions, the magnitude of these relationships was relatively small.

Once identified as team leaders by their peers, leaders’ ability, instrumental traits (task orientation, work, mastery, positive instrumentality) and expressive traits (positive expressivity, friendliness, shows emotions) were significantly related to team outcomes. Although a high level of dominance was related to being perceived as a leader, a high level of leader dominance was related to increased team conflict and decreased ability to engage in constructive conflict resolution, which were in turn related to lowered team effectiveness.

Exploratory analyses revealed several meaningful relationships. First, there were small to moderate, significant, positive correlations between self assessed trait characteristics and peer assessments of the same characteristics. Similarly, there were moderate to large, significant, positive correlations between team member and TA/professor assessments of team conflict, conflict resolution and performance. Repeated measures ANOVAs found that, although many peer evaluations of target traits became more negatively valenced over time, leadership perceptions and team performance did not systematically increase or decrease throughout the semester. However, evaluations of team conflict increased while evaluations of successful team conflict resolution decreased over time.

Findings from this study suggested that future researchers may want to consider using peer assessments of traits and team outcomes. The most interpretable and stable leadership perception and leadership effectiveness results were obtained using these methods.

3) Merritt, Ashleigh Carol, PhD. National Culture and Work Attitudes in Commercial Aviation: A Cross-Cultural Investigation, The University of Texas at Austin, 8/96.

This study asked if Hofstede’s dimensions of national culture, based on data from IBM employees in the late 60’s and 70’s, could be replicated with data from a vastly different profession, in the 1990’s. More than 8,000 male commercial airline pilots, from 22 airlines, in 15 countries (Argentina, Australia, Brazil, British Hong Kong, Cyprus, Ireland, Japan, Korea, Malaysia, Morocco, New Zealand, the Philippines, Switzerland, Taiwan, and the USA) participated in a survey to ask the question: To what extent, and in what areas, are pilots’ work related attitudes and values universal, i.e., part of a professional pilot norm, and to what extent are they influenced by national culture.

The replication study was successful in that Hofstede’s formula for Power Distance (PD) produced a correlation of .79 between the country-level scores of the pilots and the IBM personnel of 25 years
Communication and teamwork were universally acknowledged as important, although the expression of that teamwork did vary cross-culturally. Attitudes toward stress reflected a universal pilot norm that the true professional is invulnerable to environmental stressors. The strongest cross-cultural differences were observed in the areas of command (PD), and flexibility with rules and routines (UA). Effect sizes as large as .4 and .5 for 12-country comparisons were observed. The pilots from Anglo countries (USA, Australia, New Zealand, Ireland, British pilots from Hong Kong) had very similar views, while amongst the non-Anglo countries, the more hierarchical command style could be differentiated by the relative priorities given to rank (Brazil), rules (Taiwan), and relationships (Philippines). The study’s unambiguous conclusion was that national culture is a powerful influence on professional’s performance, and that training and international regulations need to reflect an awareness of those differences.

4) Human Error: Jones, Sharon Gayle, Ph.D. The Role of Group Dynamics in Error Tolerant Systems, The University of Texas at Austin, 9/96.

Data regarding operations in ten Air Traffic Control (ATC) facilities in the Southwest Region were collected from controllers and expert observers. Outcomes resulting from the events which were studied were of three types: Mishaps (less than required spacing between aircraft), Normal Operations, and instances of Exemplary performance. Ethnographic, case study, and survey data were examined to determine possible associations between human error and non-task-related, contextual factors or team behaviors.

Of four conceptual categories derived during the research, three indicated the presence and applicability of team behaviors: Task Management, Information Exchange, and Interpersonal Relations. The fourth provided an indication of Situational Challenge. Conceptual scales demonstrated consistency internally and variability across the subject population. Subjects’ conceptual category scores were combined during Logistical Regression analyses into statistical models that successfully differentiated between outcomes of differing desirability (Mishaps, Normal Operations, and Exemplary Performance).

A model including all four conceptual scales correctly categorized 66% of Mishaps. This was a significant improvement over the 50% accuracy that could be obtained through random prediction. Additional models comprised of other scale combinations indicated distinctive patterns for events that culminated in Normal, Exemplary, and Mishap outcomes.

The study revealed the multi-variate, interdependent nature of operational and environmental correlates of organizational effectiveness. Factors associated with system performance included administrative (resource procurement and development), technical (equipment adequacy), social (cultural, behavioral norms), group process (team building and maintenance), and personal (fatigue) ones. Data indicated the greater influences of team over task activities on operational outcomes. They also demonstrated the restricted number and variety of desirable, influential behaviors relative to undesirable ones. Discussion addressed the enhancement of system effectiveness through incorporation of positive, group-oriented, behavioral concepts into organizational investigation, assessment, and training.


The present study examined 1,718 commercial airline pilots’ evaluations of the training they received for use of aircraft automation, automated systems on their current aircraft, and their attitudes toward the management of these automated systems. Examination of training ratings showed that, overall, roughly one-quarter of pilots felt that initial training did not adequately prepare them for operating their
aircraft. Substantial differences in ratings of training efficacy were found across airlines, aircraft types, pilots’ experience level, and exposure to discretionary, unstructured opportunities for practice during training. Examination of automated equipment evaluations revealed that ratings of automation usability are related to ratings of training efficacy. Analyses also demonstrated differences across aircraft types on automation usability, quality of troubleshooting and problem solving, and awareness of aircraft energy state; some of these differences seem to be related to aircraft manufacturer and some to automation generation. Finally, three scales were derived measuring automation preference, respondents’ discretion in use of automation, and recognition of the increased communication needs on the automated flight deck. Correlational analyses showed relationships between the scales and measures of experience, perceptions of company policies regarding automation use, and a measure of respondents’ need to avoid uncertain, ambiguous situations. Generally, more experienced pilots showed slightly higher recognition of the increased need for communication on the automated flight deck and preferred automation slightly less than did younger pilots. Overall, these results allow identification of some potential threats to safety that reside in the crew-automation interface. They also suggest that crew-automation interaction can be conceptualized from the systems viewpoint – i.e., that crew-automation interaction is determined by multiple factors, including training quality, the automated equipment itself, and the organization’s policies and procedures regarding automation use.


Data from 3,266 flight observations aboard regularly scheduled domestic and international flights from five U.S. airlines in both standard and automated aircraft are presented. Trained raters assessed flight crew performance on a four point scale (1 = poor, 2 = minimum expectations, 3 = standard, 4 = outstanding) in the areas of leadership, crew interaction, and automation management. Raters de-identified all observations, and crews were under no jeopardy of retribution for any negative behavior they exhibited during the flight. Seven core measures of crew performance, and four core measures of automation management were extracted from regression analyses and a review of commercial airline accident and incident investigations. Examination of the data indicated that less than 14% of the flights observed were rated as below standard, 71% were rated as standard, and 15% were rated as outstanding. Substantial differences in crew performance were found as a function of airline. Crew performance was also found to vary as a function of the quality of flight briefings, the complexity of the operating environment, the degree of crew familiarity, and the length of the flight. Crews in automated and standard aircraft were not found to perform differently on core measures shared by both aircraft types. Large performance differences between airlines were found for core automation management measures, such as establishing guidelines for automation management and using automation appropriately. Additionally, higher scores on the automation management measures were associated with good flight briefings. Further analysis showed that there were pilot flying effects in complex operating environments: crews performed better when the first officer was the pilot flying and worse when the captain was the pilot flying. Principal Component Analysis was used to derive three higher order scales of crew performance: Command, Crew Interaction, and Automation Management. Cluster analysis indicated that there were three natural groupings of crews based on their performance. Crews in automated aircraft had High, Standard, and Low groupings, and crews in standard aircraft had Above Standard, Below Standard, and Low groupings. Further analysis indicated that the introduction of automation has posed new issues and modes of possible errors in crew performance, and that automation can have either very good or very bad effects on crew performance in modern aircraft. Recommendations for training and policy interventions are discussed.

b. Publication Citations


c. Data on Scientific Collaborators

Captain John Bell – Consultant, former Chief of Pilots, San Francisco Base, Pan American Airlines

Captain Roy E. Butler – Research Associate, former Director of Training, Pan American Airlines

Captain E. Peter Connelly – Research Associate, former Airline pilot
d. Information on Inventions: None

e. Technical Description of Project and Results

In addition to research products, the project is producing a new generation of able researchers. Three of the six Ph.D’s produced by the project have been recognized for writing the best dissertation in the Department of Psychology. Two of the Ph.D’s have won the University of Texas Best Dissertation Award, most recently Dr. William Hines in 1998 for his analysis of line audit data from major U.S. airlines (Hines was also nominated for a national dissertation award for this work). Ph.D. graduates of this graduate program have gone on to make significant, safety-related contributions at the FAA, NASA, the NTSB, and major airlines.

The current grant research has resulted in a number of significant findings related to crew performance and system safety. These studies are described in detail in publications and are available on the project’s Website, which was developed under this grant (http://www.psy.utexas.edu/psy/helmreich/nasaut.htm). The products of the research are both methodological, in the form of tools that can be applied by researchers and operators, and empirical findings that aid in system assessment and the validation of training programs such as CRM. The research has also provided a foundation for the research and development efforts of other organizations including the Air Force in its CRM doctrine, the Air Transport Association, ICAO, AQP, and individual researchers. Members of the project participate in industry projects, such as ICAO and IATA Human Factors seminars to make research findings available to the worldwide aviation community.

In addition to producing research products for the industry, the research team has been a continuing contributor to the development of regulatory guidance for the FAA, for example, in the Advisory Circulars for CRM and AQP.

Specific research products include the Flight Management Attitudes Questionnaire (FMAQ), used by airlines in more than twenty countries as well as major and regional airlines in the US to assess attitudes about teamwork and crew coordination. The FMAQ was extended to measure crew perceptions of safety practices and commitment in airlines and is used as a tool to assess the safety culture. This survey is used to tailor CRM programs to the specific needs of pilots in different airlines. A second major effort was the development of an observational methodology, based on observable behaviors, to allow expert observers to measure crew performance in line operations. This methodology, called the Line Operational Safety Audit, has been employed by major airlines in the US in collaboration with the research team. The results provide insights into safety issues and help define training needs. More than 3,000 line flights were observed during the project. As part of this effort, new definitions of crew-based error have been developed and employed in the creation of a model of how crews assess and manage threat in the environment and how they avoid and manage error during flight operations. In a related investigation, errors by air traffic controllers in one Region were analyzed and training needs developed to facilitate controller performance.