International Journal of Applied Aviation Studies

Spring 2003

A publication of
FAA Academy
Oklahoma City, Oklahoma

REVIEW PROCESS

The Federal Aviation Administration Academy provides traceability and oversight for each step of the *International Journal of Applied Aviation Studies (IJAAS). IJAAS* is a peer-reviewed publication, enlisting the support of an international panel of consulting editors. Each consulting editor was chosen for his or her expertise in one or more areas of interest in aviation. Using the blind-review process, three or more consulting editors are selected to appraise each article, judging whether or not it meets the requirements of this publication. In addition to an overall appraisal, a Likert scale is used to measure attitudes regarding individual segments of each article. Articles that are accepted are those that were approved by a majority of judges. Articles that do not meet *IJAAS* requirements for publication are released back to their author or authors.

IJAAS is printed by the Government Printing Office. Individuals wishing to obtain a copy of the journal may contact Beverly Laughead by email at bsl@ou.edu, by telephone at (405) 954-1191, or by writing to the following address:

International Journal of Applied Aviation Studies
c/o Beverly Laughead
AMA-502
PO Box 25082
Oklahoma City, OK 73125

International Journal of Applied Aviation Studies
Spring 2003, Volume 3, Number 1

1st Printing June 2003

This issue is dedicated to the memory of
Bailee Westphalen
IJAAS Consulting Editor
Spring 2001 – Spring 2003

POLICY AND DISCLAIMERS

Policy Statement: The Federal Aviation Administration (FAA) Academy strongly supports academic freedom and a researcher's right to publish; therefore, the Federal Aviation Administration Academy as an institution does not endorse the viewpoint or guarantee the technical correctness of any of the articles in this journal.

Disclaimer of Liability: With respect to articles available in this journal, neither the United States Government nor the Federal Aviation Administration Academy nor any of their employees, makes any warranty, express or implied, including the warranties of merchantability and fitness for a particular purpose, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.

Disclaimer of Endorsement: Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the United States Government or the Federal Aviation Administration Academy.The views and opinions of authors expressed herein do not state or reflect those of the United States Government or the Federal Aviation Administration, and shall not be used for advertising or product endorsement purposes.

PRODUCTION MANAGERS

William E. Traylor	FAA Academy, Oklahoma City, OK
Ned S. Reese	FAA Academy, Oklahoma City, OK
Gwen Sawyer	FAA Academy, Oklahoma City, OK
Cynthia Seikel	University of Oklahoma, Norman, OK

EDITOR

Beverly Laughead	University of Oklahoma, Norman, OK

ASSOCIATE EDITORS

Kay Chisholm	FAA Academy, Oklahoma City, OK
Laurie St. John	FAA Academy, Oklahoma City, OK

CONSULTING EDITORS

International

Finian Connolly	Executive & Professional Training Institute, Drogheda, Co Louth, Ireland
Hugh David	Eurocontrol, East Sussex, UK
Sidney Dekker	Linköping Institute of Technology, Linköping, Sweden
Jarle Gimmestad	Braathens Airlines, Fornebu, Norway
Irene Henley	University of Western Sidney, Sidney, Australia
Graham Hunt	Massey University, Auckland, New Zealand
Magnus Jacobssen	Linköping Institute of Technology, Linköping, Sweden
Elizabeth Mathews	International Civil Aviation Organization, Montreal, Canada
Jens Rolfsen	Braathens Airlines, Fornebu, Norway
Stephen Walsh	Interaction Training Associates, Netherlands
Rogier Woltjer	Linköping Institute of Technology, Linköping, Sweden

U.S.A.

Ellen Bass	University of Virginia, Charlottesville, VA
Jeffrey Beaubien	American Institutes for Research, Washington, DC
Wendy Beckman	Parks College, St. Louis University, St. Louis, MO
Deborah Boehm-Davis	George Mason University, Fairfax, VA
Brent Bowen	University of Nebraska at Omaha, Omaha, NE
John Cain	Florida Technology Institute, School of Aeronautics, Melbourne, FL
David Carmichael	Oklahoma City University, Oklahoma City, OK
David Conway	Southeastern Oklahoma State University, Durant, OK
Richard Fanjoy	Purdue University, West Lafayette, IN
Trent Gabert	University of Oklahoma, Norman, OK
A. F. Grandt	Purdue University, West Lafayette, IN
Mavis Green	Embry-Riddle Aeronautical University, Daytona Beach, FL
Judy Holcomb	FAA Academy, Oklahoma City, OK
Todd P. Hubbard	University of Oklahoma, Norman, OK
Michael Hynes	Western Oklahoma State College, Altus, OK

CONSULTING EDITORS (con't.)

Jefferson M. Koonce	University of Central Florida, Orlando, FL
Mary N. Kutz	Oklahoma State University, Stillwater OK
Victor LaSaxon	The Boeing Company, Midwest City, OK
Rebecca Lutte	University of Nebraska at Omaha, Omaha, NE
Beth Lyall	Research Integrations, Inc., Tempe, AZ
Carol Manning	Civil Aerospace Medical Institute, Oklahoma City, OK
Harry Marks	University of Oklahoma, Norman, OK
Steven Marks	Oklahoma State University, Stillwater, OK
Royce Ann Martin	Bowling Green State University, Bowling Green, OH
Celeste Mayer	North Carolina State University, Raleigh, NC
H. C. McClure	Parks College, St. Louis University, St. Louis, MO
Jesse McMullen	FAA Academy, Oklahoma City, OK
Marjo Mitsutomi	University of Redlands, Redlands, CA
Richard H. Mogford	NASA Ames Research Center, Moffett Field, CA
Kathleen Mosier	San Francisco State University, San Francisco, CA
Mike Pearce	American Airlines
Julia Pounds	Civil Aerospace Medical Institute, Oklahoma City, OK
Stephen M. Quilty	Bowling Green State University, Bowling Green, OH
Esa M. Rantanen	University of Illinois, Urbana-Champaign, IL
Robert Ripley	Auburn University, Auburn, AL
Art Saboski	Embry-Riddle Aeronautical University, Prescott, AZ
Nadine Sarter	The Ohio State University, Columbus, OH
Robert Schlegel	University of Oklahoma, Norman, OK
Guy Sconzo	Oklahoma City Schools, Oklahoma City, OK
Kenneth Sperry	The Boeing Company, Midwest City, OK
Ron Ward	FAA Academy, Oklahoma City, OK
Thomas Weitzel	Embry-Riddle Aeronautical University, Daytona Beach, FL
Bailee Westphalen	University of Phoenix, Oklahoma City, OK
Michael Wiggins	Embry-Riddle Aeronautical University, Daytona Beach, FL
Peter Wolfe	Southwest Airlines
John Young	Purdue University, West Lafayette, IN

PHILOSOPHY STATEMENT

Cornelius Lanczos, a mathematician working in the field of applied analysis, expressed the history of mathematics in three phases:

> 1) A given physical situation is translated into the realm of numbers,
> 2) By purely formal operations with these numbers certain mathematical results are obtained, [and]
> 3) These results are translated back into the world of physical reality (1988, p. 1). [1]

Formal papers, in subjects related to aviation, roughly follow the same course. However, there appears to be a weakness in aviation research, that being the omission of the third phase.

It is not good enough that conclusions are drawn, if those conclusions fail to improve the system observed. Clearly, the observed have a say in implementing the conclusions of research, but their failure to implement the conclusions drawn by the researcher may be more indicative of a lack of understanding than a lack of desire. Researchers tend to peer into complex systems as through a soda straw, forming formal opinions on the finite without understanding the complete system. Industry, ever mindful of the complete system, may find research irrelevant, because it makes much to do about nothing.

The editorial staff, to include those listed as consulting editors, is committed to the improvement of all individuals within the aviation community. We seek to enhance existing systems bearing in mind that small improvements must not upset the delicate balance between too little and too much help. We also seek to promote safety, not by lip service, but by demonstration in how we execute our studies and how we report our findings.

We feel that the best way to translate results back to the physical world is to incorporate the viewpoints of people around the globe. Without the influence of a worldwide community, we deny the significance of diversity, and ignore the perspectives of gifted scientists from different countries. It is our hope that each reader will feel the same.

B.S.L.

---

[1] Lanczos, C. (1988). Applied Analysis. Mineola, NY: Dover Publications, Inc.

EDITOR'S NOTES

Since the 1980s, studies have indicated instances of alarm mistrust in several career environments, including aviation. Those interested in human factors may find this article to be of interest. Bliss conducted two experiments investigating the reactions of independent and dependent teams to marginally reliable alarms to determine whether task interdependence would mediate the impact of alarm reliability on primary and alarm task performances. In many complex task environments, including aviation, there is an expectation that team members collectively share in the ultimate "reward" of safely completing the flight task.

Basing the need that changes in the near future that might have implications for both the human errors that will occur and the chances of discovering and recovering from these errors, Bove and Andersen conducted four empirical studies to validate an error management taxonomy to study human errors and their resolution within the area of Air Traffic Control (ATC). They focused their validation on the basis of Reliability, Comprehensiveness, Diagnosticity, and Usability.

With the goal that the aviation training community will be proactive in identifying deficits in crew member knowledge and developing revised training regimens to respond accordingly, Bowers, Jentsch and Salas explored how the content and competencies engendered by traditional CRM programs should be expanded to address the needs of the new environment after September 11, 2001. They offer several additions that might make CRM training programs more effective in responding to those needs.

Based on earlier work that suggested significant differences between lower and upper division aviation students, Kanske, Brewster, and Fanjoy initiated a five-year study to investigate the learning styles of collegiate aviation students and track them by year group (freshman, sophomore, junior, senior and graduate). The authors hope to determine whether changes of individual learning styles occur during the college experience or individuals with "non-predominant learning styles" tend to self-eliminate from aviation programs. Identification of a predominant aviation student learning style also can be used as an aid in course instructional design. This article is a report of data collected at the end of the first year from four hundred twenty (420) students sampled across eight aviation programs.

The Dillman, Lee, and Petrin article described Aircraft Discrepancy Analysis Matrix (ADAM) and Airplane Incident Analysis Matrix (AIAM), database tools to determine where there are weaknesses in aviation safety cultures so that appropriate remedies of accidents/incidents can be devised.

Aviation educators who are interested in improving their instructional design may find this article to be quite valuable. In his study, Barnhart examined the cognitive styles, as defined by KAI theory, of those involved in collegiate aviation education, as well as of those aviation flight professionals currently in the field.

Being the first of its kind, Barnhart feels that the information derived from his study can now serve as the basis for the application of some of the other important aspects and implications of KAI theory in the future.

A critical issue in aviation safety is accurate pilot-controller dialogue. In this article, Mitsutomi and O'Brien define the critical components of aviation English as air traffic control (ATC) phraseology, English for Special Purposes (ESP) and English for General Purposes (EGP) and introduce The Aviation English Model to serve as a framework for subsequent discussions on language issues as they relate to the global aviation context.

In this article, Trippett discusses some of rules of thumb, that he introduces to his student pilots, and evaluates their accuracy. Acknowledging that rules of thumb are rules established by pilots for easier management of aviation operations, Trippett states that his article is intended to show how reliable these rules are and how comfortable pilots should be in using them. The mathematical equations may be somewhat overwhelming for some; however, the results of the evaluation are interesting.

Those who have an interest in international aviation rules and laws may enjoy this article. Ripley presents the genesis of international aviation law and discusses the provisions of the 1999 Montreal Convention, which amended earlier international air carrier agreements.

One of the issues that surfaced from the leaders in a 1998 study conducted in Oklahoma, which involved interviews with successful aviation leadership for advice pertaining to education of future aviation leaders, was the importance of teaching values to tomorrow's aviation leaders. Carmichael, Kutz, and Brown examined the literature for basic answers to some fundamental questions pertaining to values in leadership. Using examples of values in a variety of environments, such as Enron, Global Crossing, Arthur Andersen and other contemporary institutions of misguided organizational leadership, in an attempt provide valuable insights into the leadership pressures that create leadership breakdowns and failure, the authors believe that their findings have far-reaching implications for the education of future leaders in the field of aviation.

The former *IJAAS* editor, Dr. Todd Hubbard, conducted an interview regarding concepts for the design of new equipment systems courses for Air Traffic Control personnel. Readers who are interested in course structure and individual lesson structure for classroom and computer-based instruction (CBI) lessons will find valuable information in the interview with Dr. Welp. Since this is something new for *IJAAS*, we welcome your comments regarding articles of this type.

B.S.L.

International Journal of Aviation Research and Development

Volume 3, Number 1 2003

Formal Papers

Training Development Papers

Interview

International Journal of Applied Aviation Studies, 3,1

Formal Papers

Collective Mistrust of Alarms

James P. Bliss
Psychology Department, MGB 244B
Old Dominion University
Norfolk, VA 23513
(757) 683-4222
jbliss@odu.edu

Abstract

Researchers have proposed machine trust theories for individual reactions. Because teamed alarm reaction scenarios are frequent, we conducted two experiments investigating the reactions of independent and dependent teams to marginally reliable alarms. In Experiment 1, 52 college student dyads reacted to an alarm system that generated 30%, 50%, or 70% true alarms, while performing a complex primary task. In Experiment 2, 40 dyads reacted to two collateral alarm systems while performing the primary task. One system maintained a reliability of 80%, and another fluctuated among 40%, 60%, and 80% true alarms. In Experiment 1, dependent teams reacted more appropriately to alarms. However, independent teams reacted more appropriately when there were two collateral alarm systems. Participants in both experiments responded most frequently to alarms of higher reliability, and least appropriately to those of medium reliability. Adding a collateral alarm system negated performance differences on the primary task. Results suggested that machine trust theories should be extended to account for teamed alarm reactions. Furthermore, designers and trainers should promote team interdependence when operators are faced with marginally reliable signals from a single system.

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

Collective Mistrust of Alarms

Recently, the use of automated alarm systems has increased, because sensor-based emergency signal systems have become more sensitive and because system designers have an ethical and legal duty to warn users about system anomalies that threaten safety (Stanton, 1994). Operators are now more frequently notified about minor violations of preset sensor thresholds. Unfortunately, the greater numbers of false alarms from sensitive systems may lower operators' levels of trust (Breznitz, 1983) and degrade their performance levels (Bliss, 1993).

Since the early 1980s researchers have noted instances of alarm mistrust within aviation (Bliss, 1997), industrial tasks (Kortlandt & Kragt, 1980), mining (Mallett, Vaught, & Brnich, 1992), ship handling (Kerstholt, Passenier, Houttuin, & Schuffel, 1996), medicine (Bitan, Meyer, Shinar, & Zmora, 2000) and automobile driving (MacKinnon, Bryan, & Barr, 1993). In most of these instances, researchers found that participants who encountered marginally reliable alarm systems showed degraded task performance. They responded less quickly, frequently, and appropriately to the alarms. Frequently, ongoing task performance suffered as well. For many years, aviation researchers have devoted effort to understand why pilots mistrust alarm systems (Sorkin, 1988). Recent analyses of aviation accident databases have highlighted alarm systems such as the Traffic Collision Avoidance System (TCAS) and the Ground Proximity Warning System as problematic, because of their high false alarm rates (Bliss, Freeland & Millard, 1999).

Researchers have shown interest in operator trust for years. Muir (1989) presented a broad theory of machine trust based on social trust theories from Barber (1983) and Rempel, Holmes, and Zanna (1985). Barber (1983) claimed that social trust relies on one's belief that a partner follows natural and moral laws, that the partner has technical competence, and that the partner acts to benefit others. Rempel et al. (1985) stated that humans show trust because a partner behaves predictably, has a dependable disposition, and follows moral laws consistently. Muir (1989) suggested that these theories were comparable and that they could be aggregated and applied to the human-machine relationship.

Bliss (1993) demonstrated how violating aspects of Muir's theory could influence alarm reaction performance. By presenting a mixture of true and false alarms, Bliss led participants to believe that an alarm system violated Barber's notion of technical competence, across all levels of Rempel et al.'s theory. Bliss (1993) found that if individual participants did not trust an alarm system, they would not respond frequently or appropriately. In addition, alarm response rates conformed to probability matching theory because participants matched their response rates linearly to the expected rate of true alarms (Herrnstein, 1961).

Empirical alarm mistrust work has generally supported Muir's (1989) theory; however, no work has been done to apply her theory to situations where teams must collectively judge and react to alarms. This oversight is crucial for many environments. Members of airplane flight crews often work together to accomplish tasks during takeoff and landing (Wiener & Nagel, 1988). Furthermore, in a recent investigation of the Aviation Safety Reporting System, Bliss, Freeland, and Millard (1999) found that most alarm mistrust incidents occurred during takeoff and landing, when operator workload is the highest. Increased primary task workload has been shown to exacerbate alarm mistrust, further reducing alarm response rates (Bliss & Dunn, 2000).

For years, researchers have acknowledged the frequency with which military operators must collectively monitor signals (Dobbins, 1962). Early "paired watchkeeping" studies required participants to collectively detect moving signals within background noise. Such research generally found teamed detection performance to be superior to individual detection performance (Wiener, 1964; Morgan & Alluisi, 1965). Subsequently, researchers tried to determine the optimal configuration of monitoring teams. Waag and Halcomb (1972) witnessed better detection performance with larger teams and task interdependent structure, where participants worked together to detect signals.

In another early study of interdependent signal monitoring, Morrissette, Hornseth, and Shellar (1975) constructed an experiment where two people were responsible for detecting visual signals. In one condition, the task was configured so that each team member was responsible for a unique portion of the signal detection task. In another condition, the team members each held the same responsibilities, so that signal detection was performed redundantly. Morrissette et al. found that dyads in independent configurations (where both dyad members redundantly viewed the same information) exhibited longer signal detection times. However, because of the impact of social facilitation, Morrissette et al. suggested that independent team configurations be used only when team members are in close proximity to each other.

Paired watchkeeping research demonstrated the importance of task structure for collective signal detection. However, treatment of the topic was cursory and based around practical concerns. Aside from a general consideration of social facilitation, researchers tended to neglect the social factors that could impact collective performance. Furthermore, they considered only signal detection, not evaluation. Signal trust was not at issue, because all signals were presumed true, warranting the same reaction. Yet, the paradigm used by paired watchkeeping researchers offers a promising method for studying how task configuration impacts collective alarm mistrust.

Studying collective alarm mistrust across different task configurations is increasingly important. The complexity of automated task environments has rendered alarm reactions less certain and more variable. As an example, Foushee

(1982) described an incident report filed with the Aviation Safety Reporting System where an alarm activated during flight and members of the crew began to diagnose the problem. However, because of poor communication between the pilot and the flight engineer, the flight crew disregarded what was indeed a crucial and relevant warning signal.

Aspects of Muir's (1989) machine trust theory may allow the prediction of such tragedies. However, her theory was not initially designed to apply to multiple operator situations, so its relevance must be demonstrated. Because her theory is a product of other sociological theories, it may hold promise for collective alarm reactions. The social dynamics of trust underlying Muir's theory may influence dyad members working interdependently.

Campion, Medsker, and Higgs (1993) defined task interdependence as the degree to which team members interact and depend on each other to complete a task. Wageman and Baker (1996) noted that several typologies of task interdependence exist. Thompson (1967) distinguished among pooled, sequential, and reciprocal interdependence. Pooled interdependence requires no coordination among team members, sequential interdependence requires linear task completion and sequencing, and reciprocal interdependence requires transfer of resources among team members. Van de Ven, Delbecq, and Koenig (1976) later introduced "team interdependence" which requires joint diagnosis, problem solving, and collaboration for task completion.

As noted above, flight crews sometimes work together to respond to many tasks including responding to alarm signals. However, the degree to which cockpit tasks are interdependent varies. In commercial aviation, resource-intensive activities such as flight planning, takeoff, and landing typically require team interdependence, where the flight crew must exchange information to perform tasks. In cruise flight, however, it is common for members of the flight crew to work independently. Tasks are completed separately and the results are pooled to affect an ultimate outcome. Variations in task interdependence are common in other complex task environments as well (Shea & Guzzo, 1987).

A number of researchers have manipulated task interdependence experimentally. To do so, experimenters typically control the degree to which team members are required to share information to complete a task. Another strategy has been to control the number of interconnections necessary among team members for successful task completion.

In their recent meta-analysis of team efficacy, potency, and performance, Gully, Incalcaterra, Joshi, and Beaubien (2002) noted that several theories help explain the importance of interdependence for team task performance. Sociotechnical theory suggested that teams are composed of social and technological components and that interactions among humans and between human operators and technology both determine team effectiveness (Kolodny & Kiggundu, 1980). Gully et al. (2002) also pointed out the importance of input-process-output models and team goal setting theories. Each of these theories stresses that social interaction processes contribute to the ultimate success or failure of a team in an interdependent situation. Perhaps the most relevant of these processes is social loafing, where team members work less hard when their efforts are combined with others toward a common goal (Bernstein, Penner, Clarke-Stewart, & Roy, 2003). As mentioned previously, paired watchkeeping and machine trust theorists/researchers have stressed the role of social processes (Muir, 1989; Morrissette et al., 1975). The current research represents an attempt to expand these early investigations. Researchers generally agree that the notion of task interdependence is a logical starting point for examinations of team task performance (Kozlowski & Bell, in press). Similar to Morrissette et al. (1975), we manipulated task interdependence through division of labor. At the same time, we used Muir's (1989) theory as a framework to explain collective reactions to unreliable alarms.

Although past researchers have acknowledged the importance of collective signal monitoring, detection, and reaction, it is clear that theories of machine trust must be expanded to account for the added social dynamics present within alarm reaction teams. Muir's (1989) theory assumed that human-machine trust, like human-human trust, is predicated on social and moral expectations. When dyads must detect and react to signals, task complexity rises because individuals are judging the alarm system and each other at the same time. It is important to determine how alarm mistrust and task interdependence interact.

Goals and Hypotheses

The goal of the current research was to investigate the reactions of dependent and independent teams to alarm signals of various reliability levels. Using a dual-task paradigm, we required teams of individuals (randomly assigned to be dependent or independent) to respond or cancel true and false alarms while performing a complex ongoing task. Our goal was to determine whether task interdependence would mediate the impact of alarm reliability on primary and alarm task performances. In the first experiment, dyads performed a primary task while reacting to signals from a low-, medium-, and high-reliability alarm system. The second experiment was similar to the first, except that dyads were presented signals from two collateral alarm systems with independent reliability levels.

We made several predictions for the first experiment. We expected dependent teams to respond to and cancel alarms more appropriately than independent teams. Earley and Northcraft (1987) showed that independent team members reaping collective rewards might demonstrate social loafing. This and the advantages of cohesion from task interdependence (Shea & Guzzo, 1989; Johnson & Johnson, 1989) suggested a performance advantage for dependent teams.

However, theories of task interdependence stress that team interdependence requires increased communication and information sharing (Campion et al., 1993). For that reason, we also expected dependent teams to react more slowly. We also expected communication demands to degrade ongoing task performances for the dependent teams.

Based on the role of technical competence within Muir's (1989) machine trust theory, and on prior research by Bliss (1993), we expected all teams to react faster and more appropriately and to respond more frequently as alarm reliability increased. However, we also expected all teams to react more slowly and less appropriately to alarms of 50% reliability. Considering Muir's theory, operator trust should be lowest when alarms are 50% reliable, because the alarm system is not predictable. Furthermore, prior research by Bliss (1993) suggested that such mistrust will translate into degraded performance.

In the current research, our consideration of teamed alarm reaction responsibilities is clearly relevant to theories of task interdependence, and less so for goal or outcome interdependence (Saavedra, Earley, & Van Dyne, 1993). In many complex task environments, including aviation, there is an expectation that team members collectively share in the ultimate "reward" of safely completing the flight task. It also has been shown that shared goals facilitate cooperative strategies and group performance (Mitchell & Silver, 1990). For these reasons, in the current research all teams shared interdependent outcomes.

Experiment One Method

The experiment was structured according to a 2 x 3 mixed design. Team interdependence was manipulated between two groups. Independent team members required no interaction to react appropriately to the alarms. Dependent team members, however, had to share information to react appropriately to the alarm signals. Alarm system reliability was manipulated within groups. The three levels, 30%, 50%, and 70%, indicated the percentage of true alarms within a particular session. The sequence of reliability levels experienced by each team was random.

Participants

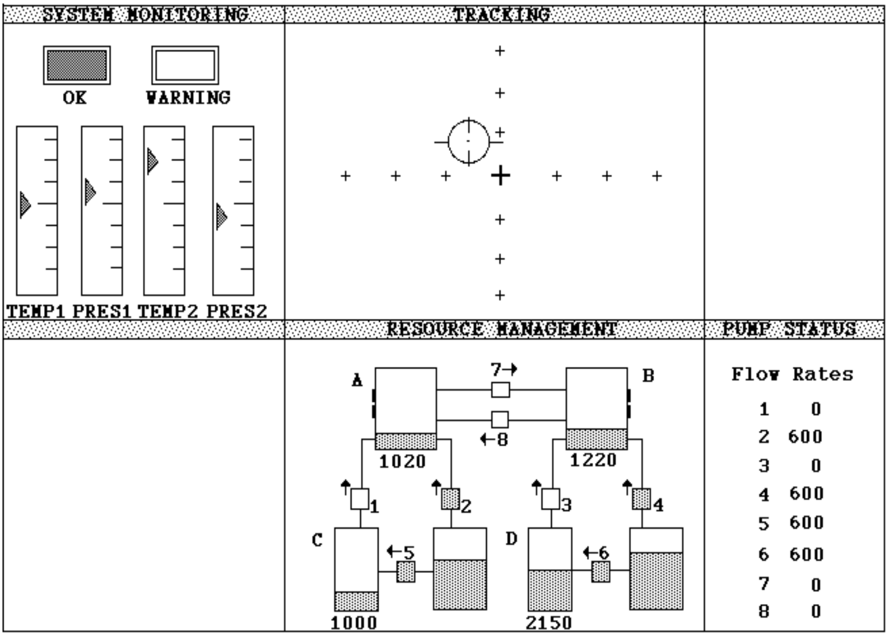
One hundred four students [66 women and 38 men] from psychology courses at The University of Alabama in Huntsville participated in this study.The ages of the participants ranged from 19 to 40. The participants formed 52 dyads. There were 13 same-sex, dependent teams (5 male and 8 female), 15 same-sex, independent teams (2 male and 13 female), 13 different-sex, dependent teams and 11 different-sex independent teams. Participation was voluntary. As a collective outcome, a $20 performance bonus was promised to the team with the best collective performance on the primary and alarm tasks.

Materials

In addition to the Informed Consent Form, participants completed demographics and opinion questionnaires. The Multi-Attribute Task (MAT) battery (Comstock & Arnegard, 1992) was used as the ongoing experimental task. The MAT battery is a microcomputer-based task that measures cognitive and spatial abilities and effectively simulates aircraft piloting demands. The task features dual-axis compensatory tracking, gauge monitoring, and resource management tasks. Because of its continuous nature, the tracking task is particularly suitable for measuring operator attention shifts in multiple-task situations. It required participants to center a continuously moving ball on the center crosshairs (see Figure 1). The program periodically assesses the distance of the ball from the crosshairs.

During the resource allocation task, participants manipulated eight pumps to control fluid transfer among six holding tanks. The goal was to ensure that the fluid levels in tanks A and B stayed between the dark boxes on the sides of the tanks (see Figure 1). The number of times participants activated pumps was measured to determine performance.

While tracking and allocating resources, participants monitored the TEMP1 and TEMP2 gauges at the upper left corner of the screen. The pointers continuously fluctuated. If a pointer traveled further than one mark from the center in either direction, participants were to press the corresponding function key on the keyboard to reset it (F1 for TEMP1, and F3 for TEMP2). The frequency of gauge resetting was measured.



1	0
2	600
3	0
4	600
5	600
6	600
7	0
8	0

Figure 1. The MAT Battery

Participants performed the MAT back-to-back on IBM compatible 486 computers, using 14" color VGA monitors. They used the mice and keyboards to make responses. At the same time, auditory and visual alarms were presented 90 degrees to the side on a Macintosh Quadra 610 personal computer, using a 14" VGA monitor.The auditory stimulus was the fire bell digitized from a Boeing 757/767 simulator, followed by a male voice announcing "temperature, temperature." The visual stimulus was a yellow panel with the word "TEMPERATURE" on it. When an alarm occurred, participants were to determine whether the MAT TEMP1 and TEMP2 gauges were out of tolerance. If both were, the alarm was true. Participants were then to reset the MAT gauge and press the F12 key (marked "R" for "RESPOND") on the Macintosh keyboard, in that order. If none or only one of the TEMP1 and TEMP2 gauges were out of tolerance, the alarm was false. Participants then pressed the F9 key (marked "C" for "CANCEL") on the Macintosh keyboard and resumed the MAT tasks.

Members of dependent teams were required to communicate to determine alarm validity, because TEMP gauge fluctuations were distributed between the two computers, and did not always occur at exactly the same time (though they always occurred within 5 seconds of each other). Independent team members monitored both gauges, because all gauge fluctuations appeared on both computers. Alarm stimuli were presented at 60 dB(A) (ambient sound was 45 dB(A)). Both types of teams could choose which member actually made the alarm reaction.

Alarm reaction measures included speed to react (in seconds), appropriateness of reactions (whether or not participants responded to true alarms and canceled false alarms), and response frequency (the percentage of alarms participants responded to within each experimental session).

Procedure

After arriving individually at the laboratory, participants were paired and randomly assigned to the independent or dependent experimental condition. They signed the Informed Consent Form and completed the Background Information Questionnaire. Next, the experimenters carefully presented unique instructions for the experiment to dependent and independent teams. Independent teams were told that they had all of the necessary information on their MAT screens to make reactions to the alarms and did not necessarily have to communicate with the other team member. Dependent teams were told that they did not have all of the necessary information on their MAT screens, and so had to communicate to determine alarm validity.

Next, participants received two 120-second training sessions with the MAT battery.They also were shown the alarms, and were told how to respond to or cancel them. Then participants completed a 200-second practice session with both tasks.

After the practice sessions, the participants began the first of three ten-minute experimental sessions, with mandatory 5-minute breaks between sessions. Ten alarms were presented during each session on a variable interval schedule (average interstimulus interval was 65 seconds). The true alarm rate of alarms during a session was 30%, 50%, or 70%, randomly determined. During the experiment, each team encountered all reliability rates. The reliability of the alarms was told to the participants before they began each session.

The appropriateness of reactions was reflected by a score, which was presented to the participants at all times on the Macintosh screen. Responding to true alarms and canceling false alarms were correct reactions and increased the team score. Responding to false alarms or canceling true alarms were incorrect reactions and decreased the team score. After completing the three experimental sessions, participants completed the Opinion Questionnaire. They were then debriefed and dismissed.

Experiment One Results

After examining the raw data to confirm that they were distributed normally, we calculated a series of 2 X 3 mixed analyses of variance (ANOVAs) to determine support for our hypotheses.

There was no interaction between team dependence and alarm reliability level for alarm response frequency and no main effect of dependence ($\underline{p} > .05$, see Figure 2). However, alarm reliability increased linearly with alarm reliability, $\underline{F}(1,50) = 174.06$, $\underline{p} < .001$.

Although there was no significant interaction between dependence and reliability for alarm reaction appropriateness (see Figure 3), a main effect showed that dependent teams made more appropriate reactions to alarms than independent teams, $\underline{F}(1, 50) = 4.906$, $\underline{p} = .031$. We also noted a quadratic main effect for reliability, with teams making less appropriate reactions to alarms that were 50% reliable, $\underline{F}(1,50) = 10.359$, $\underline{p} = .002$.

Figure 4 shows that although there was no significant interaction or dependence main effect for alarm reaction time, there was a quadratic main effect for reliability, with participants reacting to 50% reliable alarms more slowly than the others, $\underline{F}(1,50) = 4.505$, $\underline{p} = .039$.

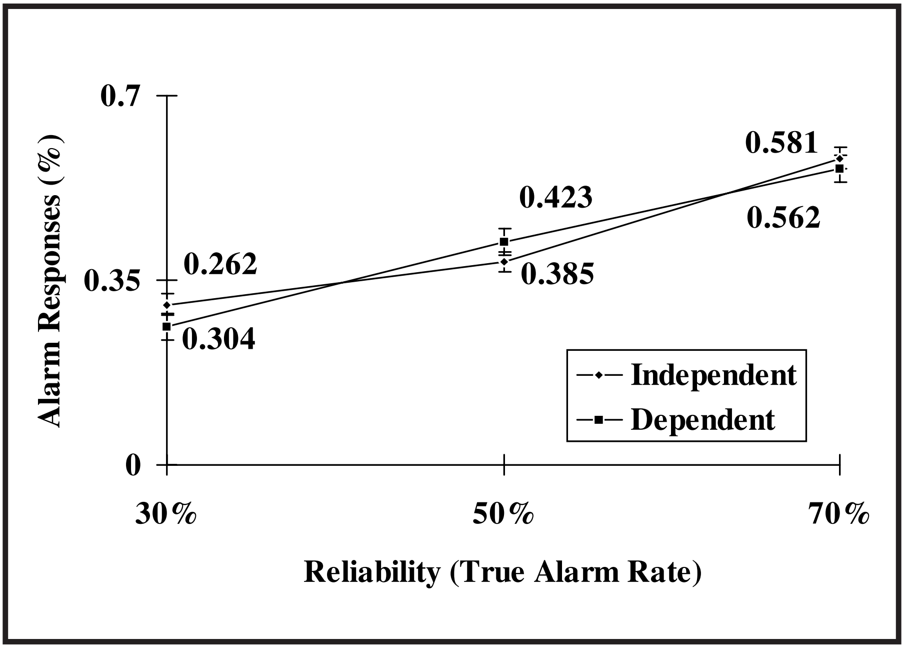


Figure 2. Alarm Response Frequency as a Function of Team Dependence and Alarm Reliability.

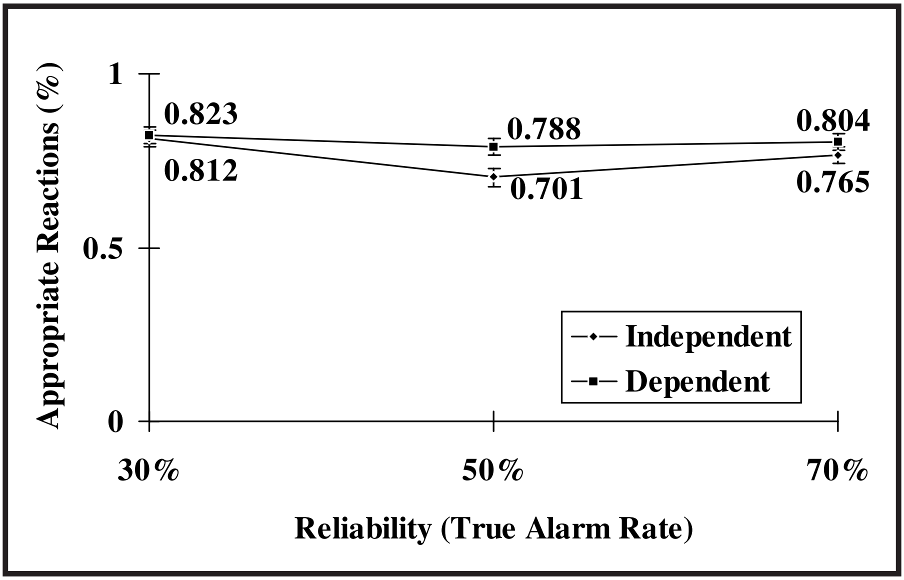


Figure 3. Alarm Reaction Appropriateness as a Function of Team Dependence and Alarm Reliability.

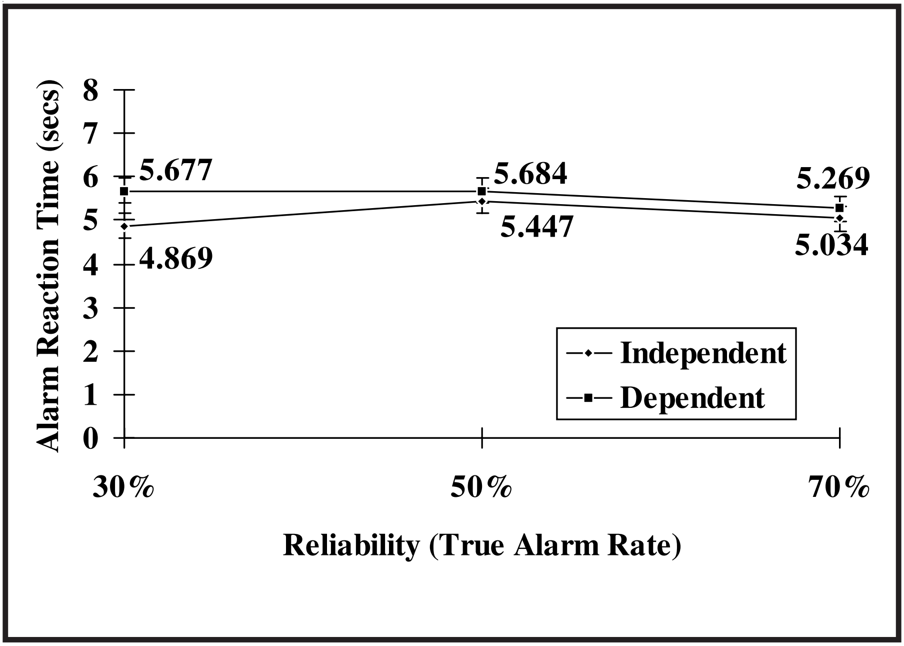


Figure 4. Alarm Reaction Time (seconds) as a Function of Team Dependence and Alarm Reliability.

We also analyzed task performance for the ongoing MAT tasks. There was no interaction of reliability and dependence for frequency of monitor resetting; however, a main effect for dependence showed that independent team members reset their monitors more frequently than dependent team members did, F(1,50) = 23.185, p< 001. Also, teams reset their monitors more frequently when there were more true alarms. Contrasts indicated that the data followed linear (F(1,50) = 138.082, p<.001) and quadratic (F(1,50) = 7.723, p = .008) trends (see Figure 5).

An examination of pump activation frequency revealed no interaction or reliability main effect; however, dependent team members activated pumps more often than independent team members, F(1,50) = 3.959, p = .05 (see Figure 6). We found no significant interaction or main effects for MAT task tracking accuracy, p>.05.

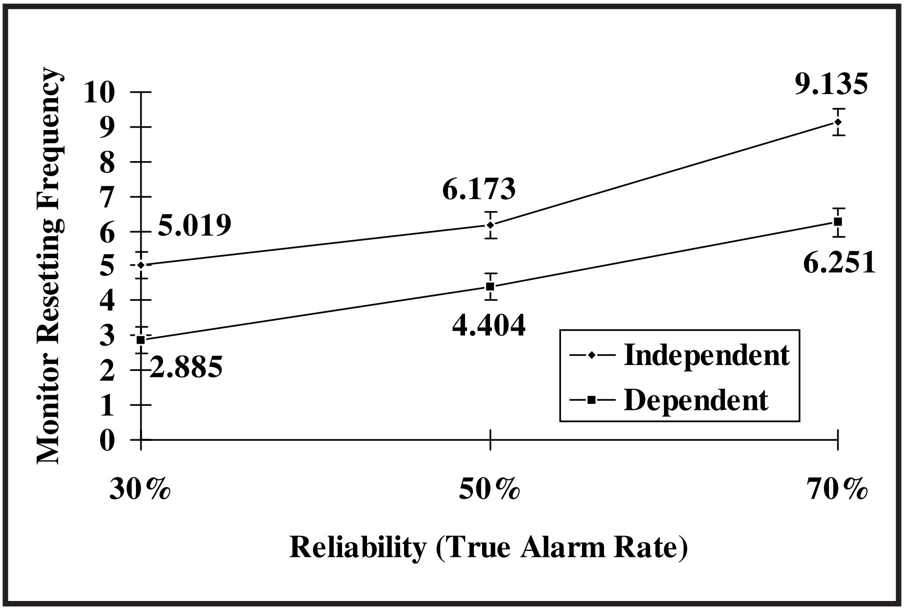


Figure 5. Frequency of Monitor Resetting as a Function of Team Dependence and Alarm Reliability.

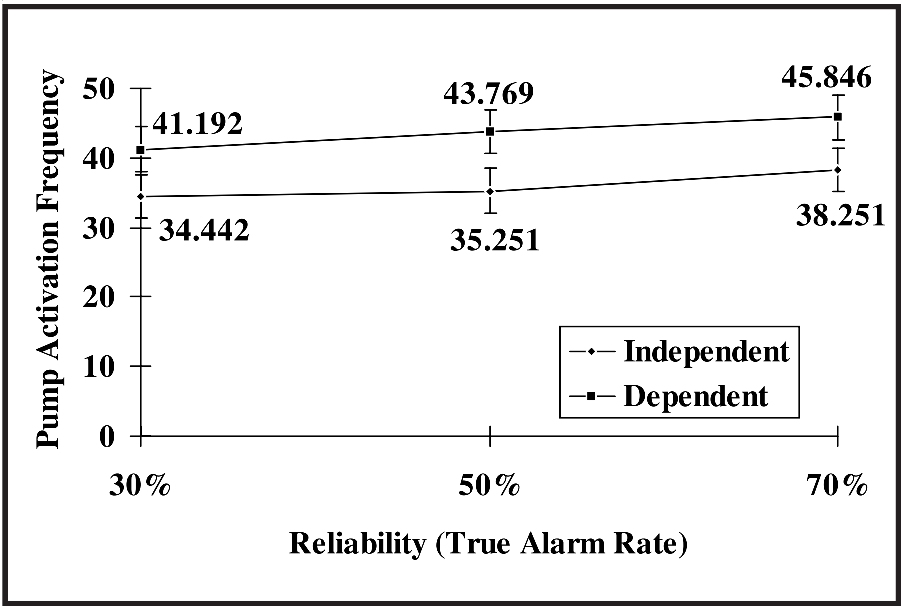


Figure 6. Pump Activation Frequency as a Function of Team Dependence and Alarm Reliability.

Experiment One Discussion

In general, our results confirmed our hypotheses. The fact that dependent teams reacted more appropriately to alarms suggested that they realized the benefits of task interdependence. This extended the findings of Morrissette et al. (1975), showing that the advantages of interdependence exist when dyads are required to make complex reactions to stimuli (not just detect them). Sharing information before reacting to alarms was important for teams, and ensured that reactions of dependent team members were more appropriate than those of independent team members. The lack of an effect for reaction speed suggested that dependent teams formed strategies for task response quickly. Also, we noted that even independent team members communicated with each other, though it was not required.

Social loafing research may help to explain why dependent teams reacted more appropriately than independent teams. As noted by Williams and Karau (1991), social loafing tends not to occur in situations where one team member believes they can contribute information or resources not available to the other team member. In task dependent divisions of labor, such as was present here, this is clearly the case. What is not clear from this study, however, is whether the differences between task dependent and task independent divisions of labor will remain when the alarm reaction task is more complex.

An interesting finding concerned the dip in alarm reaction performance at the middle level. Probability matching theory (Herrnstein, 1961) suggested that teams should exhibit mediocre performance at the 50% level. However, although the response frequency data reflected this expectation, participants in both team conditions exhibited the worst reaction appropriateness to alarms that were 50% reliable. Prior researchers (e.g., Bliss, Jeans, & Prioux, 1996) have found similar results while testing individuals. Muir's (1989) theory of machine trust may provide an explanation for this phenomenon. Muir showed that operator trust is a factor of predictability, dependability, and faith. An alarm system that is 50% reliable is not predictable, because there is an equal chance of true and false alarms being generated. In contrast, alarm systems that are 30% or 70% reliable are equally predictable, though not equally reliable. The implications of this may be dire for certain medical or process industry alarm systems that approximate a 50% true alarm rate (see Tsien & Fackler, 1997).

As mentioned previously, teams generally matched their response rates to the true alarm rates in a linear fashion. However, the match was not exact, as teams responded less frequently than the reliability levels might have suggested. It is possible that the workload inherent in the primary task may have lessened overall response rates (Bliss & Dunn, 2000). To test this possibility, it is important to determine if response rates would decrease further given additional alarm system complexity.

Performances on the primary tasks suggested that dependent team members may have benefited from the distributed workload. While their tracking performances were similar to the independent teams, they seemed able to distribute monitor resetting responsibilities equitably while retaining more cognitive resources to activate the pumps.

Experiment Two

In the first experiment, teams interacted with one alarm system while performing a complex primary task. However, often teams of operators must interact with multiple alarm systems. In transport aviation, for example, flight crews may confront collateral alarm signals indicating altitude deviation, excessive speed, impending traffic, proximity to the ground, and the disconnection of the autopilot, to name a few. Furthermore, these alarms commonly reflect varying reliability levels. For example, the Traffic Collision Avoidance System has been criticized for its high false alarm rate (Shapiro, 1994), as has the Ground Proximity Warning System (Billings, 1997). However, data suggested that the altitude deviation, overspeed, and autopilot disconnect alarm systems are typically more reliable (Bliss et al., 1999).

The presence of multiple alarm systems, each with its own reliability level, also is common outside of aviation. Medical personnel must react to reliable alarm systems like the arterial catheter blood pressure signaling system (Tsien et al., 1997) and less reliable alarm systems such as the pulse rate oximeter (Wiklund, Hok, Stahl, & Jordeby-Jonsson, 1994).

Because collateral signals occur often in aviation, Doll and Folds (1986) recommended that researchers study multiple concurrent signals. McDonald, Gilson, Mouloua, and Deaton demonstrated the influence of collateral alarm signals on operator reactivity. They found that features of collateral alarm signals might be associated (properly or improperly) by operators in predictable ways, based on Gestalt principles of proximity, similarity, or continuity (McDonald et al., 1995).

The available research concerning collateral alarm reactivity has not included the effect of such alarms on teams. However, as mentioned above, reactivity to alarm signals frequently is a shared responsibility. As demonstrated in the first experiment, dependent team members may react more appropriately to alarm signals from an individual alarm system. What is not known is whether these tendencies will persist given signals from multiple alarm systems. This question is of theoretical interest as well as practical importance. Morrissette et al. (1975) suggested that additional task complexity could enhance social facilitation. However, they presented no data to support this claim. Similarly, Zajonc (1965) has suggested that social facilitation may be mediated by arousal, so that increased task complexity may indeed increase social facilitation. An important question is whether the effect would lead to differential performances for independent and interdependent teams.

Another reason for increasing the complexity of the experimental task is to address concerns voiced by Wageman (1999). She suggested that individual differences may moderate the effects of task dependence in situations where the task is not challenging. To account for the possibility that individual differences may have influenced the results from Experiment One, it is important to increase the complexity of the tasks performed.

In this experiment, dyads responded to two separate alarm systems, one of which (temperature alarms) had an 80% true alarm rate. The reliability of the second (pressure alarm) system fluctuated among 40%, 60% and 80% true alarms.

As in the first experiment, we expected dependent teams to react more appropriately to alarms than independent teams (Shea et al., 1989; Johnson et al., 1989). We also expected that added alarm complexity would augment the reaction time difference between dependent and independent teams, rendering it statistically significant. Muir's (1989) theory emphasized the importance of technical competence of a machine for trust. Because we presented two unreliable alarm systems, and because of the added primary and secondary task workload associated with collateral alarm systems, we expected operator trust and performance to drop compared to the first experiment.

Based on Muir's (1989) theory of trust, and prior research by Bliss (1993), we also expected all teams to perform worst for alarms of intermediate reliability or worst predictability (in this case, 60%).

Experiment Two Method

The second experiment was structured as a 2 X 3 mixed design. Team member dependence was manipulated in the same manner as in the first experiment. In this experiment there were two active alarm systems: pressure and temperature. The pressure alarm system reliability (40%, 60%, or 80% within a session) was manipulated within groups. Teams experienced all reliability levels across three sequential task sessions. The reliability of the temperature alarm system remained constant at 80% during all sessions. Participants experienced both pressure and temperature alarms during each session. Dependent measures for the ongoing and alarm tasks were the same as in the first experiment.

Participants

Eighty undergraduate and graduate students (40 dyads) from Old Dominion University participated in this experiment for course credit. The ages of the participants ranged from 18 to 43. There were 10 same-sex, dependent teams (1 male, 9 female), 11 same-sex, independent teams (2 male, 9 female), 10 different-sex, dependent teams and 9 different-sex, independent teams. A $20 performance bonus was promised to the team with the highest combined score on the primary and alarm tasks.

Materials

The Multi-Attribute Task (MAT) battery (Comstock et al., 1992) again was used as the primary task. All MAT tasks operated identically to the first experiment, with one exception: during gauge monitoring, participants monitored all four gauges at the upper left corner of the screen. Two of the gauges reflected temperature alarm information, and the other two reflected pressure alarm information. Reset buttons were F1 for TEMP1, F2 for PRES1, F3 for TEMP2, and F4 for PRES2 (see Figure 1).

Responding to and canceling alarms was done in the same manner as in Experiment One. True temperature alarms occurred when the MAT gauges TEMP1 and TEMP2 were out of tolerance, and true pressure alarms occurred when the PRES1 and PRES2 gauges were out of tolerance. To ensure auditory uniqueness of the temperature and pressure alarms, we employed a procedure followed by Bliss and Kilpatrick (2000): the temperature alarm consisted of a steady-state burst followed by four discrete pulses and a male voice annunciation ("temperature, temperature"). For pressure alarms, the sequence was reversed, so that the four pulses preceded the steady-state burst and the male voice annunciation ("pressure, pressure"). Alarm annunciation lasted two seconds.

As before, dependent team members had to communicate to determine alarm validity, because TEMP and PRES gauge fluctuations were distributed between the two MAT computers. Independent team members saw all gauge fluctuations. Alarm stimuli were again presented at 60 dB(A) (ambient sound level was 45 dB(A)).

Procedure

The procedure used in Experiment Two was similar to Experiment One. Participants completed an Informed Consent Form, Background Questionnaire, three practice sessions on the MAT, and an alarm familiarization session.

After the practice sessions, participants began the first of three experimental sessions, separated by 5-minute breaks. Ten alarms were presented during each session (five temperature, and five pressure). The reliability of pressure alarms during each session was 40%, 60%, or 80% (randomly determined). The reliability of the temperature alarms was 80% during each session. The reliability of both alarm systems was told to the participants before they began each session. The appropriateness of reactions was reflected by a team score, present at all times on the Macintosh screen. After three experimental sessions, participants were debriefed and dismissed.

Experiment Two Results

We calculated separate ANOVAs for temperature and pressure alarm reactions. First, we calculated one-way ANOVAs to determine whether there was an effect of team dependence on temperature alarm reactions. Results indicated no significant differences for reaction time, reaction appropriateness or response frequency ($p>.05$).

After examining the data for temperature alarm reactions, we then calculated 2 X 3 mixed ANOVAs to examine pressure alarm reactivity. There was no interaction between team dependence and pressure alarm reliability for alarm response frequency, and no main effect of dependence (see Figure 7). However, there was a linear main effect of reliability, $F(1,38) = 129.600$, $p<.001$.

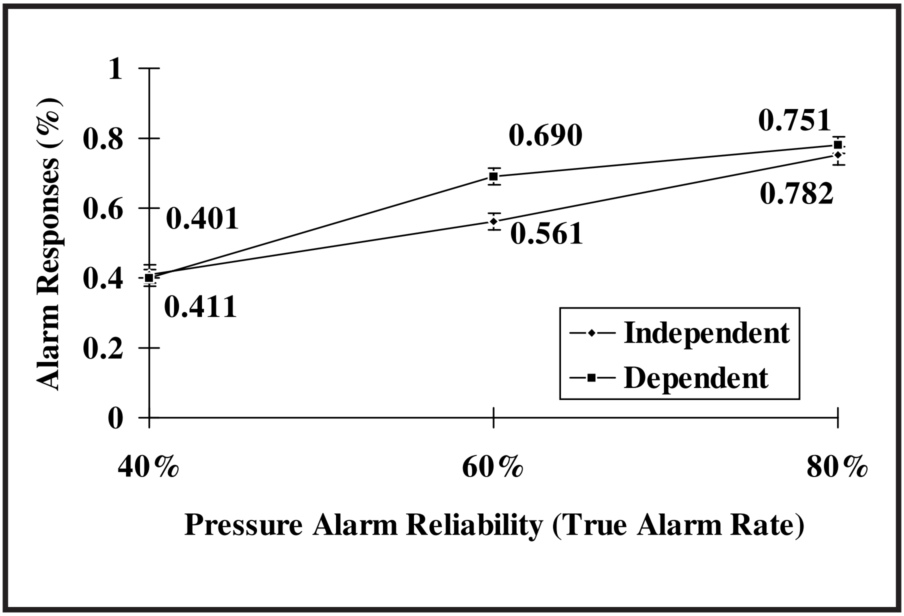


Figure 7. Alarm Response Frequency as a Function of Team Dependence and Pressure Alarm Reliability.

There was a significant interaction between dependence and reliability for alarm reaction appropriateness (see Figure 8), $F(2,76) = 10.193$, $p<.001$. We also found that independent teams made more appropriate reactions to alarms than dependent teams, $F(1, 38) = 4.000$, $p=.05$. A quadratic main effect for reliability was also significant, with participants showing less appropriate reactions to alarms that were 60% reliable, $F(1,38)=19,563$, $p<.001$.

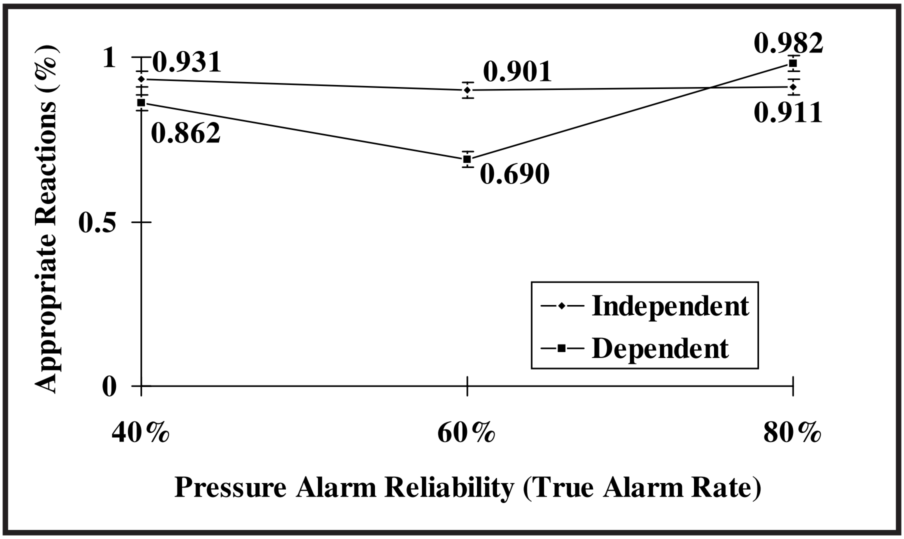


Figure 8. Alarm Reaction Appropriateness as a Function of Team Dependence and Pressure Alarm Reliability.

Although there was no significant interaction or dependence main effect for alarm reaction time, there was a linear main effect for reliability. Participants reacted more quickly to alarms that were more reliable, F (1,38) =8.181, p=.007 (see Figure 9).

We also analyzed task performance for the ongoing MAT tasks. We found no significant interaction or main effects for any of the primary task measures.

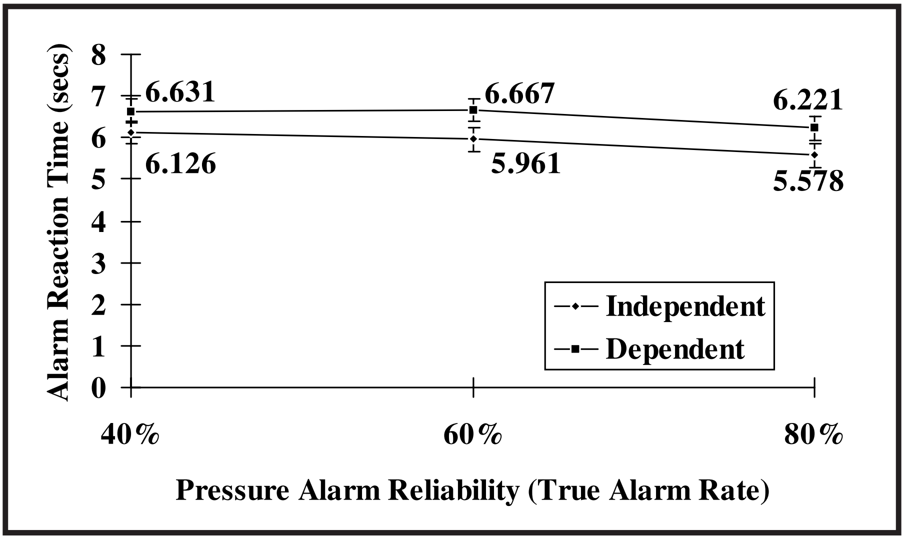


Figure 9. Alarm Reaction Time (seconds) as a Function of Team Dependence and Pressure Alarm Reliability.

Experiment Two Discussion

It is clear from both experiments that response frequency rose linearly with reliability and reaction appropriateness was significantly lower at the intermediate reliability level. These similarities were consistent with past research, and showed support for Muir's theory of machine trust (1989).

There were some interesting differences in the Experiment Two data. The fact that reaction appropriateness and alarm response frequency generally were higher in the second experiment suggested that the added task complexity may have increased social facilitation. However, we also noted that independent teams reacted more appropriately than dependent teams when the task was more complex. This was contrary to our expectations. Both Morrissette et al. (1975) and Zajonc (1965) suggested that increased task complexity could increase social facilitation. Yet, it is possible that the demands of an additional alarm system coupled with information sharing degraded performance for teams that were required to share information.

Another difference concerned primary task performance. Whereas data from the first experiment showed that alarm reliability affected primary task performances, the current research did not reveal such differences. One possible explanation is that the presence of two alarm systems increased the participants' alarm task workload and masked primary task differences (Bliss & Dunn, 2000).

Another interesting observation was the lack of performance differences for the temperature alarm reaction data. Prior research by Bliss (1993) suggested that reaction strategies may be formed prior to the start of a session, particularly for alarm systems that are predictable (Muir, 1989). If this is true, teams may have focused more on the pressure alarm system, continuously adjusting to the shifting reliability of those alarms. Adjustment to the fluctuating pressure alarm reliability may have been easier for the independent teams, as they were not required to coordinate before reacting.

Wageman (1999) proposed that individual differences could moderate the effects of task dependence when tasks are not challenging. Because Experiment One was conducted at a small university, teams consisted of individuals who were often already acquainted. However, Experiment Two occurred at a larger university, and most participants had not met each other prior to participating. This relative unfamiliarity may have influenced performance for dependent teams, making their interaction more forced or unnatural. On the other hand, the experimental tasks and rewards were intrinsically challenging, and participants in both experiments knew their individual and team performances were subject to monitoring by the experimenter and their teammates. Therefore, it is likely that the amount of social loafing was similar in both experiments (Williams et al., 1991).

General Discussion

This research addressed the paucity of empirical work concerning collaborative alarm reactions. Although laboratory research may not fully capture the complexity or represent the consequences of an actual alarm situation, the experimental control achievable within a true experimental framework allowed the precise manipulation of alarm system reliability and lent rigor to the statistical comparisons.

It is clear from Experiment One that dividing signal monitoring detection and evaluation responsibilities between teammates may at times help teams make more appropriate reaction decisions. Since outcome dependence existed for all teams in Experiment One, such findings supported the idea that independent efforts may suffer when outcomes are dependent (Earley et al., 1987). However, the results from Experiment Two revealed an interesting fact: independent efforts actually may be superior when task complexity increases. Added task responsibilities may make team coordination more difficult, and the results of that coordination less certain.

As Bliss and Dunn demonstrated in 2000, increased workload may interact with alarm mistrust. Whether the workload occurs because of more frequent alarm stimuli (as in Bliss & Dunn, 2000), or because of multiple active alarm systems (as shown in this research), the result appears similar: performance is degraded on the primary task. This is of considerable importance, because in many complex environments the presence of alarm signals often accompanies increases in workload. In fact, in some environments it has been noted that the necessity to react to alarms (true or false) can itself constitute a considerable drain on operators, and may ultimately increase primary task workload (Kortlandt & Kragt, 1980; Sorkin, 1988).

A recurring theme across alarm mistrust research was that performance showed definite degradation as alarm reliability approaches 50%. This finding highlighted the importance of predictability within Muir's (1989) machine trust theory. Low- and high-reliability alarm systems may not be equally useful, but they are both predictable. For that reason, decisions about performance are made more appropriately.

It is true that a 50% or 60% alarm reliability level may not represent the reliability of alarm systems in some task environments (such as aviation or nuclear power). However, alarm systems in other environments do generate false signals at such rates. For example, Tsien et al. (1997) reported that true positive rates for alarm systems in an intensive care unit ranged from 0% to 54%. Also, Kortlandt and Kragt (1980) described industrial process alarm systems that achieved true positive rates of only 13%! Clearly, a consideration of low and intermediate reliability rates is important.

In past research, participants have shown a tendency to respond to alarm systems at a rate that matches the advertised reliability rate. Bliss (1993) attributed this to the phenomenon of "probability matching" discussed by Herrnstein in 1961, where animals and humans showed a tendency to match operant responses to the rate of reinforcement available from a certain source. Since 1993, Bliss and his colleagues have demonstrated probability matching in a number of alarm mistrust experiments (Bliss, 1997; Bliss et al., 1996; Bliss & Dunn, 2000).

The current research showed that teams also tend to respond to alarms at a level approximating their stated reliability. However, there was a notable difference between the first and second experiments. In Experiment One, participants' response rates were considerably lower than the stated reliability rates of 30%, 50%, and 70%. This may have occurred because of the workload associated with the primary-secondary task combination. However, in the second experiment, where the primary task workload was higher (because of the need to monitor more gauges) participants more closely matched their response rates to the reliability rates of 40%, 60%, and 80%. This was a paradox, because Bliss and Dunn (2000) found that increased primary task workload reduced alarm responsiveness. Perhaps the relatively high reliability rate of the temperature alarms (80%) increased participants' responsiveness to the pressure alarms. If the presence of high-reliability alarms does provide such an influence, this may constitute a method for increasing response rates to alarms in general. Future research should address this idea in more detail.

From prior research, intermediate alarm reliability levels are problematic. One fairly obvious recommendation is to improve the algorithms that determine alarm signal activation. However, this is not always an easy task because of the numerous factors that influence alarm activation and the general complexity of alarm systems. Another solution may be to design training programs to enhance operator responsiveness, as recommended by Bliss and Gilson (1998). Bliss et al. (1998) suggested the creation and adoption of an adaptive training program to help naïve and experienced operators react appropriately to marginally reliable alarm systems. Such a program, combined with alarm system improvements, may indeed hold promise for alarm manufacturers and users. However, research is needed to determine the optimal design and implementation for such training.

Conclusions

The stated goal of the research presented here was to determine whether task interdependence would mediate the impact of alarm reliability on primary and alarm task performances. The answer to this question appears dependent on the criteria of interest. If alarm reaction appropriateness is most important, the influence of interdependence may be moderated by task complexity: in situations with collateral alarm systems, independent team member efforts seemed to yield more appropriate alarm reactions. Added task complexity also appeared to enhance social facilitation among team members, improving response frequency and reaction appropriateness in general. Reaction speed did not appear to be sensitive to differences in team interdependence or task complexity. An important and stable finding was degraded performance at medium reliability levels, regardless of team task structure or task complexity.

The research presented made an important contribution to the state of knowledge regarding teamed alarm reactivity. However, many questions remain concerning the function of teams in such circumstances. It is important that researchers study the impact of variables such as leadership style, team cohesiveness, and communication content to better understand why collective reaction appropriateness changes with added task complexity. It is also important to test strategies for counteracting collective alarm mistrust.

Acknowledgements

This research could not have been completed without the contributions of a number of individuals. Sincere thanks are extended to Laticia Bowens, Rebecca Krefting, Ashley Byler, Anastasia Gibson, Gary Capobianco, Susan Sidone, and Holly Mason.

References

Barber, B. (1993). *Logic and the limits of trust*. New Brunswick, NJ: Rutgers University Press.

Bernstein, D.A., Penner, L.S., Clarke-Stewart, A., & Roy, E.J. (2003). *Psychology* (6th ed.). Boston, MA: Houghton Mifflin.

Billings, C.E. (1997). *Aviation automation: The search for a human-centered approach*. New Jersey: Lawrence Erlbaum Associates.

Bitan, Y., Meyer, J., Shinar, D., & Zmora, E. (2000). Staff actions and alarms in a neonatal intensive care unit. *Proceedings of the IEA 2000/HFES 2000 Congress.* July 29-August 4: San Diego, CA.

Bliss, J.P. (1993). The cry-wolf phenomenon and its effect on operator responses. Unpublished doctoral dissertation, University of Central Florida, Orlando.

Bliss, J.P. (1997). Alarm reaction patterns by pilots as a function of reaction modality. *The International Journal of Aviation Psychology, 7* (1), 1-14.

Bliss, J.P., & Dunn, M.C. (2000). Behavioural implications of alarm mistrust as a function of task workload. *Ergonomics, 43* (9), 1283-1300.

Bliss, J.P., Freeland, M., & Millard, J. (1999). Alarm related incidents in aviation: A survey of the Aviation Safety Reporting System database. *Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomics Society*. September 27-October 1: Houston, TX.

Bliss, J.P., & Gilson, R.D. (1998). Emergency signal failure: Implications and recommendations. *Ergonomics, 41* (1), 57-72.

Bliss, J.P., Jeans, S.M., & Prioux, H.J. (1996). Dual-task performance as a function of individual alarm validity and alarm system reliability information. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*, 1237-1241. Santa Monica, CA: Human Factors and Ergonomics Society. October 2-8: Philadelphia, PA.

Bliss, J.P., & Kilpatrick, F. (2000). The influence of verbal content on alarm mistrust. *Proceedings of the 2000 Human Factors and Ergonomics Society Annual Meeting*. July 30-August 4: San Diego, CA.

Breznitz, S. (1983). *Cry wolf: The psychology of false alarms*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Campion, M.A., Medsker, G.A., & Higgs, C.A. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel Psychology, 46*, 823-850.

Comstock, J.R., & Arnegard, R.J. (1992). *The multi-attribute task battery for human operator workload and strategic behavior research*. (Technical Memorandum No. 104174). Hampton, VA: National Aeronautics and Space Administration.

Dobbins, D.A. (1962). *Monitor performance task – Status report*. (USAPRO Technical Research Report No. 1123). Washington, D.C.: USAPRO.

Doll, T.J., & Folds, D.J. (1986). Auditory signals in military aircraft: Ergonomics principles versus practice. *Applied Ergonomics, 17* (4), 257-264.

Earley, P.C., & Northcraft, G.B. (1989). Goal setting, resource interdependence, and conflict. In M.A. Rahim (Ed.), *Managing conflict: An interdisciplinary approach* (pp. 161-170). New York: Praeger.

Foushee, H.C. (1982). The role of communications, socio-psychological, and personality factors in the maintenance of crew coordination. *Aviation, Space, and Environmental Medicine, November*, 1062-1066.

Gully, S.M., Incalcaterra, K.A., Joshi, A., & Beaubien, J.M. (2002). A meta-analysis of team-efficiency, potency, and performance: Interdependence and level of analysis as moderators of observed relationships. *Journal of Applied Psychology, 87* (5), 819-832.

Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*, 267-272.

Johnson, D.W., & Johnson, R.T. (1989). *Cooperation and competition: Theory and research*. Edina, MN: Interaction Book Company.

Kerstholt, J.H, Passenier, P.O., Houttuin, K., & Schuffel, H. (1996). The effect of a priori probability and complexity on decision making in a supervisory control task. *Human Factors, 38* (1), 65-78.

Kiggundu, M.N. (1983). Task interdependence and job design: Test of a theory. *Organizational Behavior and Human Performance, 31*, 145-172.

Kolodny, H., & Kiggundu, M.N. (1980). Towards the development of a sociotechnical systems model in woodland mechanical harvesting. *Human Relations, 33*, 623-645.

Kortlandt, D., & Kragt, H. (1980). Process alerting systems as a monitoring tool for the operation. In *3rd International Symposium on Loss Prevention and Safety Promotion in the Process Industries, 3*, 804-814. September 15-

19. Basel, Switzerland: Swiss Society of Chemical Industries.

Kozlowski, S.W.J., & Bell, B.S. (in press). Work groups and teams in organizations. In W.C. Borman, D.R. Ilgen, R.J. Klimoski, & I.B. Weiner (Eds.), *Handbook of psychology: Industrial and organizational psychology*. New York: Wiley.

MacKinnon, D.P., Bryan, A.D., & Barr,A. (1993). *Four studies on the effects of multiple warnings: The false alarm effect and overwarning*. (Technical Report, Contract # AA8547). Arizona State University: Project ABLE.

Mallett, L.G., Vaught, C. & Brnich, M.J. (1992). Sociotechnical communication in an underground mine fire: A study of warning messages during an emergency evacuation. *Safety Science, 16*, 709-728.

Mathieu, J.E., & Hamel, K. (1989). A causal model of the antecedents of organizational commitment among professionals and nonprofessionals. *Journal of Vocational Behavior, 34*, 299-317.

McDonald, D.P., Gilson, R.D., Mouloua, M., & Deaton, J.E. (1995). The effect of collateral alarms on primary response behavior. In the *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, 1020-1024.

McGrath, J.E. (1964). *Social psychology: A brief introduction*. New York: Holt, Rinehart & Winston.

Mitchell, T.R., & Silver, W.S. (1990). Individual and group goals when workers are interdependent: Effects on task strategies and performance. *Journal of Applied Psychology, 75*, 185-193.

Morgan, B.B., & Alluisi, E.A. (1965). On the inferred independence of paired watchkeepers. *Psychonomic Science, 2*, 161-162.

Morris, J.H., & Steers, R.M. (1980). Structural influences on organizational commitment. *Journal of Vocational Behavior, 17*, 50-57.

Morrissette, J.O., Hornseth, J.P., & Shellar, K. (1975). Team organization and monitoring performance. *Human Factors, 17* (3), 296-300.

Muir, B.M. (1989). Operators' trust in and percentage of time spent using the automatic controllers in a supervisory process control task. Doctoral thesis. University of Toronto.

Rempel, J.K., Holmes, J.G., & Zanna, M.P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49*, 95-112.

Saavedra, R., Earley, P.C., & Van Dyne, L. (1993). Complex interdependence in task-performing groups. *Journal of Applied Psychology, 78* (1), 61-72.

Salas, E., Fowlkes, J.E., Stout, R.J., Milanovich, D.M., & Prince, C. (1999). Does CRM training improve teamwork skills in the cockpit? Two evaluation studies. *Human Factors, 41* (2), 326-343.

Salas, E., Rhodenizer, L., & Bowers, C. (2000). The design and delivery of crew resource management training: Exploiting available resources. *Human Factors, 42* (3), 490-511.

Shapiro, N. (Executive Producer) (1994). NBC Dateline. New York: National Broadcasting Corporation.

Shea, G.P., & Guzzo, R.A. (1989). Groups as human resources. In K.M. Rowland and G.R. Ferris (Eds.), *Research in Human Resources and Personnel Management (vol. 5)*, pp 323-356. Greenwich, CT: JAI Press.

Sorkin, R.D. (1988). Why are people turning off our alarms? *Journal of the Acoustical Society of America, 84(3)*, 11078.

Stanton, N. (1994). A human factors approach. In N. Stanton (ed.), *Human factors in alarm design*. London: Taylor & Francis.

Thompson, J.D. (1967). *Organizations in action*. New York: McGraw-Hill.

Tsien, C.L., & Fackler, J.C. (1997). Poor prognosis for existing monitors in the intensive care unit. *Critical Care Medicine, 25*(4), 614-619.

Van de Ven, A.H., Delbecq, A.L., & Koenig, R. (1976). Determinants of coordination modes within organizations. *American Sociological Review, 41*, 322-338.

Waag, W.L., & Halcomb, C.G. (1972). Team size and decision rule in the performance of simulated monitoring teams. *Human Factors, 14*(4), 309-314.

Wageman, R. (1999). Task design, outcome interdependence, and individual differences: Their joint effects on effort in task-performing teams (commentary on Huguet et al., 1999). *Group Dynamics, 3*(2), 132-137.

Wageman, R., & Baker, G.P. (1996). Incentives and cooperation: The joint effects of task and reward interdependence on group performance. *Journal of Organizational Behavior, 18*, 139-158.

Welch, (1999). Auditory alarms in intensive care. In N.A. Stanton & J. Edworthy, *Human Factors in Auditory Warnings* (pp. 319-342). Aldershot: Ashgate.

Wicklund, L., Hok, B., Stahl, K., & Jordeby-Jonsson, A. (1994). Postanesthesia monitoring revisited: Frequency of true and false alarms from different monitoring devices. *Journal of Clinical Anesthesia, 6*, 182-188.

Wiener, E.L. (1964). Performance of multi-man monitoring teams. *Human Factors, 6*, 279-284.

Wiener, E.L. & Nagel, D.C. (1988). *Human factors in aviation*. San Diego, CA: Academic Press.

Williams, K.D., & Karau, S.J. (1991). Social loafing and social compensation: The effects of expectations of co-worker performance. *Journal of Personality and Social Psychology, 61*, 570-581.

Zajonc, R.B. (1965). Social facilitation. *Science, 149*, 269-274.

International Journal of Applied Aviation Studies, 3,1

Validation of an Error Management Taxonomy in ATC

Thomas Bove and Henning B. Andersen
Risoe National Laboratory
Frederiksborgvej 399
P. O. Box 49 – DK-4000
Roskilde, Denmark
Email: tbv@bang-olufsen.dk

Abstract

While a fairly large number of taxonomies are available describing mechanisms behind human error in technological work contexts, a surprisingly small number of studies have been made about the mechanisms underlying human error detection and recovery. The objective of this paper is to describe and validate an error management taxonomy to be used in the analysis of error events in the area of Air Traffic Control (ATC). The validation is focused on four aspects: Reliability, comprehensiveness, diagnosticity, *and* usability. *Results obtained from four empirical studies – including incident reports, a simulator experiment and a questionnaire study – provided support for the notion that the framework could be of use in future error management studies*
.

Introduction

Studies have shown that human errors have contributed to about 90% or more of ATC incidents (Kinney, Spahn, & Amato, 1977). To minimize the risk of events that may compromise the safety of the air traffic it is therefore important to develop error resistance strategies. The latter can be divided into two main categories, namely *error prevention* and *error correction* (Frese, 1991). Tradi-

---

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

tionally, human error resistance strategies have mostly focused on error prevention. This focus is understandable since many studies of incidents and accidents in safety critical domains indicate that the underlying problem is often to be found in a combination of shortcomings of human performance in man-machine systems and the fact that most of such systems are unforgiving to errors (Rasmussen, 1984). An obvious solution to avoid such is to try to prevent human errors (e.g. through failsafe protection devices, automation and error reducing procedures).

There are, however, several reasons why safety strategies narrowly based on error prevention may not be successful. First, human errors will inevitably occur and it is impossible to anticipate all types of errors that may occur in a specific task context. Second, by focusing exclusively on avoiding various kinds of errors there is a risk of imposing excessive limitations on the performance and, therefore, compromising both effective and adaptive behavior. Indeed, it has been argued that the efficiency of error avoidance strategies has been exhausted in ultra-safe areas such as aviation and air traffic control, and that the end result of increased error suppression may in fact be counterproductive seen from a safety perspective (Amalberti, 2001). Since human errors are inherent to real life and people have powerful capabilities to control errors, it is important, to a larger extent, to try to manage the manageable and to support people's chances of detecting and recovering from errors. Consequently, error management should be considered an important supplementary safety goal.

As indicated by the previous paragraphs, improvement in system safety in the area of ATC requires gaining systematic and detailed knowledge about the underlying mechanisms of not only error production but also error recovery. An important step in that direction is to explore whether existing error taxonomies, expanded with a classification scheme of how errors are managed, can be applied to studies of human errors in ATC. Such a taxonomy can be a useful human-factors tool in diagnosing underlying mechanisms behind air traffic controller (ATCO) errors and their resolution. The results can be useful in relation to analyzing the effects of various ATC safety initiatives, be they changes in system design, operating procedures, or training of personnel.

The goal of this paper is therefore to validate a taxonomy to study human errors and their resolution within the area of ATC. More specifically the focus is to validate the taxonomy on the basis of four product criteria that have been suggested by Wiegmann and Shappell (2002) to be crucial for the utility of a taxonomy: *Reliability*: Can consistent results be reached across raters and time? *Comprehensiveness*: Does the taxonomy cover all of the relevant variables that it purports to cover? *Diagnosticity*: Does the taxonomy provide insight? *Usability*: Can the taxonomy be applied in practical settings?

Four empirical studies have been conducted for the validation. First, a pilot study was carried out on the basis of Swedish ATC incident reports to get an initial indication of whether the core of the framework could successfully be applied to the analysis of real, complex ATC scenarios. The second study examined simulator scenarios from an ATC training curriculum. Here it was possible to analyze error events that in many ways are comparable to those that occur on a normal everyday basis in ATC. The third study was a questionnaire study where different human factors experts provided their opinion about the framework. This feedback provided the means for determining the usability of the framework. The final study focused on analyzing a series of cases generated on the basis of the critical incident technique. Here it was possible to test out a full-scale version of the framework.

The Framework

A literature review has been carried out to determine which categories should be included in the error management taxonomy (for details concerning the literature review please refer to Bove, 2002). The focus of the review was on taxonomies associated with human error as well as on human performance issues occurring both before and after errors. With respect to the phase before the occurrence of an error the focus was on threat management which concerns how operational factors that have the potential of leading to errors and jeopardizing safety are controlled. In relation to the phase after the occurrence of an error the literature review revealed that main issues of interest could be organized around the following questions: *who* was involved in the detection and recovery of the error and/or its consequences; *when* was the error or its consequences detected; *how* was the error and/or its consequences detected and corrected; and finally *what* was the behavioral response and outcome*?* In addition, it should be possible to also give an answer to the *why*-question – namely why did the error occur and why was it successfully or unsuccessfully managed? This can be determined on the basis of so-called Performance Shaping Factors (PSFs) - that is, contextual factors that can have a positive or negative influence on the course of events. In the following pages the conceptual framework based on an error management model is presented.

A Model of Error Management

To be able to develop a classification system of the error management process it is useful to the have a model that can be used as an organizing principle. In our view, some of the most promising error management frameworks are to be found in Helmreich, Klinect, and Wilhelm (1999) and Kanse and Van der Schaaf (2001a). Each of these two models has advantages and disadvantages in relation to the current context. The model by Kanse and Van der Schaaf (2001a) is good at describing the underlying process of error management but it does not seem to be applicable to errors that do not lead to unwanted situations. Consequently, the large majority of errors that are caught before any consequences have occurred cannot be analyzed by this model. The Helmreich et al. model has the advantage that it describes in a rough but intuitively appealing manner the main stages of the error management process. However, the classifications included in the model are only of behaviors and outcomes and not of the underlying cognitive processes. That is, the classifications allow a description of *what* happened but not *how* it came about. Below we present a model that tries to incorporate the advantages of the former two models. Its main structure is closely related to the model presented by Helmreich et al., but more emphasis is made on the underlying processes of error management.

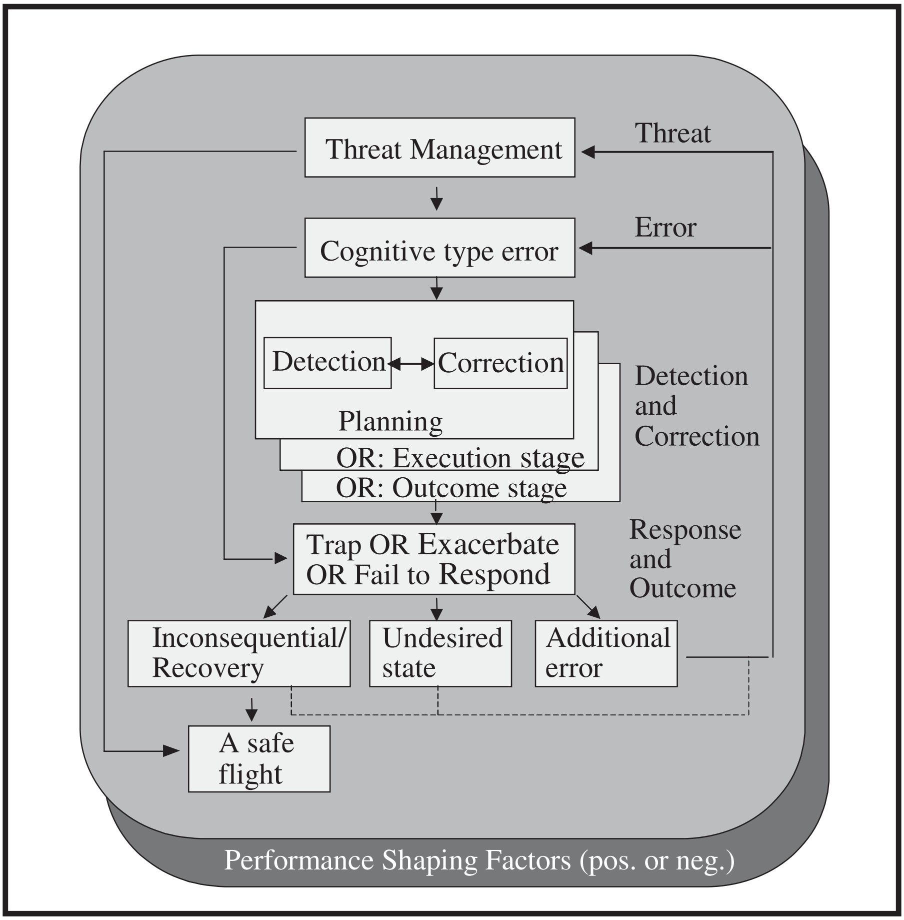


Figure 1. A model of error management.

The model starts out with a threat management section, which concerns the ATCO's awareness of aspects of the operational environment that might lead to errors and operational problems (e.g. a thunderstorm). If the threat is discovered the ATCO may try to avoid it, but if the threat is not discovered an error might result. The error, in turn, can be analyzed on the basis of the cognitive mechanisms behind it. The error might not be detected, but if it is, the detection and/or the recovery may happen at different stages in the evolution of the error. Different kinds of responses might be produced and the result may vary from being inconsequential to an undesired state or to a new error. In the case where the outcome is an additional error, a new error analysis can begin. Even if the outcome is inconsequential or involves an undesired state, new errors might still occur in the event sequence.

The unit of analysis within the framework is the individual acts by and cognitive process of the individual ATCO in the error production and management. Still, different actors may be associated with different stages in the model. Furthermore, a set of contextual factors – so called Performance Shaping Factors - constitutes an integrated part of the framework. These factors can be used to expand the analysis beyond the individual level to include team and organizational factors that are relevant to gain a comprehensive understanding of why the event occurred and how it was prevented from developing into an even more serious situation.

The Main Dimensions of the Framework

Based on the model presented above, an error management taxonomy has been developed. The dimensions and classifications associated with the individual actions are shown below and will be elaborated in the following.

<table>
<tr><th colspan="14">THREAT & ERROR</th></tr>
<tr><td rowspan="3"><i>Threat
&
error</i></td><td>Threat preparedness</td><td colspan="6">No anticipation</td><td colspan="6">Anticipation</td></tr>
<tr><td>Cognitive error type</td><td colspan="2">Perception</td><td colspan="3">Short-term memory</td><td colspan="3">Long-term memory</td><td colspan="2">Decision</td><td colspan="2">Response</td></tr>
<tr><td>Procedural violation</td><td colspan="6">Yes</td><td colspan="6">No</td></tr>
<tr><th colspan="14">DETECTION & RECOVERY</th></tr>
<tr><td rowspan="2"><i>Who:</i>
Actor</td><td>Error/state detected by</td><td colspan="2">No one</td><td colspan="3">Producer</td><td colspan="3">Co-actor in context</td><td colspan="2">Co-actor outside context</td><td colspan="2">System</td></tr>
<tr><td>Error/state corrected by</td><td colspan="2">No one</td><td colspan="3">Producer</td><td colspan="3">Co-actor in context</td><td colspan="2">Co-actor outside context</td><td colspan="2">System</td></tr>
<tr><td><i>When:</i>
Processes</td><td>Detection stage</td><td colspan="4">Planning</td><td colspan="4">Execution</td><td colspan="4">Outcome</td></tr>
<tr><td rowspan="2"><i>How:</i>
Processes</td><td>Detection source</td><td colspan="4">External communication</td><td colspan="4">System feedback</td><td colspan="4">Internal feedback</td></tr>
<tr><td>Error/state correction</td><td colspan="3">Ignore</td><td colspan="3">Apply rule</td><td colspan="3">Choose option</td><td colspan="3">Create solution</td></tr>
<tr><th colspan="14">RESPONSE & OUTCOME</th></tr>
<tr><td rowspan="2"><i>What:</i>
Behavior
&
outcome</td><td>Error/state response</td><td colspan="4">Trap/ mitigate</td><td colspan="4">Exacerbate</td><td colspan="4">Fail to respond</td></tr>
<tr><td>Error outcomes</td><td colspan="4">Inconsequential/ recovery</td><td colspan="4">Undesired state</td><td colspan="4">Additional error</td></tr>
</table>

*Figure 2.*The analysis framework.

Threat management

A threat can be seen as a part of the operational environment which might evolve into a problem if not handled in due time. Threat management means being prepared for the threats and is important insofar as by knowing in advance that certain problems might occur it becomes easier to respond in a timely and efficient manner. Below, two types of threat management are described.

- No Anticipation: In this case no recognition of any threat(s) was made by any of the involved ATCOs before it developed into a real problem.
- Anticipation: A threat (or several threats) in the environment is known by one or several ATCOs before it leads to a problem.

Cognitive domain

Below are shown the cognitive failure types that were chosen for the analysis of mechanisms behind individual errors (Isaac, Shorrock, Kirwan, Kennedy, Andersen, & Bove, 2000).

- Perception: This cognitive domain concerns issues related to picking up and understanding information.
- Short-term memory: This domain concerns short-term storage or retrieval of information.
- Long-term memory: This domain concerns long-term storage or retrieval of more permanent information based on the person's training and experience.
- Judgment & decision-making: This domain concerns making projection of trajectories, planning future actions and to making decisions.
- Response execution: Sometimes people carry out actions that they have not intended.

Procedural violation

Procedural violations are included as a part of the error section and they constitute within the framework a subgroup of the decision-making errors. That is, only in the case where a decision-making error has been made should the classifier determine whether it was also a procedural violation. Although procedural violations and errors within some frameworks are considered mutually exclusive (e.g. Reason, 1990) we have chosen to classify procedural violations, in the current context, as a subgroup of errors insofar as intentional violations usually are carried out as a short-cut to what is seen as unnecessary procedures and regulations (Helmreich, Wilhelm, Klinect, & Merritt, 2001).

Error discovery and recovery

The taxonomy for structuring the analysis of the error discovery and recovery is based on following questions, namely the "who"-, "when"-, "how"- and "what"-questions. More specifically, these questions concern:

1) *who* was involved in the detection and recovery of the error and/or its consequences;
2) *when* was the error or its consequences detected;
3) *how* was the error and/or its consequences detected and corrected; and finally
4) *what* was the behavioral response and outcome?

The “who”-question

The following are the different possible actors involved in the detection and correction of the error or its consequences:

- No one: No one discovered the problem while it was still possible to solve.
- Producer: The person who produced the error was also the one to discover and/or recover the error (or its consequences).
- Co-actor in context: An observer sharing all or most of the context, goals and actions (e.g. two ATCOs at the same position) discovered and/or recovered the error (or its consequences).
- Co-actor outside context: An observer sharing a significant part of the goals, but not the context (e.g. a neighboring ATCO) discovered and/or recovered the error (or its consequences).
- System: Any kind of automated defense – e.g. Traffic Collision Avoidance System (TCAS) or Short-Term Conflict Alert (STCA) - discovered and/or recovered the error (or its consequences).

The “when”-question

Kontogiannis (1999) suggested three different stages of performance during which error detection may occur.

- The outcome stage: The error is not caught until it has produced some consequences on the environment.
- The execution stage: An erroneous action has been carried out on the system, but the error is caught before any consequences have ensued.
- The planning stage: Detection at the planning stage is usually associated with information-pick-up necessary for later actions or discussions and deliberations about what to do (e.g. between radar and planner controller).

The “how”-question

The how-question covers both how the error and/or its consequences were detected and how it was corrected. In relation to detection, it is of interest to obtain knowledge about the cues or mechanisms of the detection. Various researchers (e.g. Rizzo, Ferrante & Bagnara, 1995; Kontogiannis, 1999) have suggested a number of partially overlapping classification systems. The mechanisms relevant in the current context can be subsumed under the following categories.

- External communication: Interaction with other people can provide information that serves to detect an error. That is, a problem is discovered because another person says something that is either wrong or that reveals the presence of a problem.
- System feedback: This kind of feedback relies on cues directly found in the operational environment. System feedback includes information from the radar screen and also visual sighting from either the tower or the cockpit.
- Internal feedback: This kind of feedback refers to error detection that

requires no direct feedback from the environment (e.g., the ATCO remembers that (s)he has to check a parameter).

In addition to identifying the processes underlying the detection, it is also of interest to understand the processes underlying the problem solving and decision-making associated with the recovery. Below are presented some distinctions, which are inspired by a classification system developed by Orasanu and Fischer (1997) to distinguish between different kinds of decision events.

- Ignore: Even though an error has been detected - while there still is a chance to do something about it - no response to correct it is chosen.
- Apply rule: In many situations there only seems to be one thing to do in order to resolve the problem.
- Choose option: In this case several options were considered before deciding on a specific solution and resources that are more conscious are required than the "Apply rule" category.
- Create solution: This group of recovery processes is concerned with situations where a completely new response has to be generated since no known options are applicable to the situation at hand. This is the most resource demanding of the possible recovery processes.

The "what"-question

The final question concerns what was the behavioral response and the outcome of the error. These issues are based on directly observable phenomena and do not require any inferences about the underlying cognitive processes. Based on a model of error management developed by Helmreich et al. (1999) the following classifications have been derived.

The response to the error can be divided into the following three groups:

- Trap/mitigate: Error is detected and managed before any consequences have developed or the consequences of the error are diminished.
- Exacerbate: The error is detected but the recovery action worsens the situation.
- Fail to respond: No response is produced because the error is either not detected, detected too late, or simply ignored.

The outcome of the error can be divided into the following categories.

- Inconsequential/recovery: No negative consequences were observed, possibly because recovery attempts were successful.
- Undesired state: The end result was a potentially critical situation, an incident, or accident. In the current context, the most frequent undesired state is violation of the prescribed aircraft separation standards.
- Additional error: Sometimes errors pave the way for new errors and this may be the beginning of a chain of errors. The general characteristic of these errors is that they negatively affect workload, situation awareness or other task related factors.

Performance shaping factors

Performance Shaping Factors (PSFs) are generic technical, psychosocial, and organizational factors that can have a positive or negative influence (possibly both) on the course of events. They can be used to give an answer to the why-question – namely why did the error occur and why was it successfully or unsuccessfully managed. The main groups of PSFs are shown below. The PSFs are based on frameworks that explicitly focus on the domain of the ATC (or aviation in general) and/or on factors influencing the recovery process (e.g. Isaac et al., 2000; O'Leary, 1999; Kanse & Van der Schaaf, 2001b). The main groups of PSFs have been selected on the basis of the previously described evaluation criteria.

1. Traffic, airport and airspace
2. Ambient Environment
3. Procedures and Documentation
4. Workplace design, Human Machine Interaction (HMI), and equipment factors
5. Training and Experience
6. Person Related Factors
7. Social and Team Factors
8. Company, Management and Regulatory Factors

Validation

The usefulness of the error recovery framework in relation to error management analyses will be explored in the following. For this purpose, the framework has been evaluated on the basis of different kinds of data material. In the current context, we decided to start out with incident reports, because this type of method exemplifies naturalistic data from operational activities over an extended time frame. At the same time, acquiring the data does not require a large amount of resources in the basic data collection process. However, since incident reports will not allow for additional inquiries concerning specific issues of interest, this source will have to be supplemented with a real-time study and some interviews based on the critical incident technique. In addition to these studies aimed at applying and developing the framework it was also decided to do a questionnaire study where the goal was to get some input from human factors experts concerning the relevance of the individual dimensions and the overall structure of the framework. Hereby it would be possible to get an indication of the usability of the framework. The studies used in the validation are described in detail below.

Study 1 - Swedish Incident Reports (Pilot Study)

The main goal of the pilot study was to get an indication of the reliability of the core of the classification system on the basis of analysis of incident reports. More specifically, the study focused on whether consistent classifications could be obtained by different judges (PSFs were not analyzed in this study).

Participants

Two human factors experts made the classification of error events.

Materials

Forty-five Air Traffic Management incident reports from the Swedish CAA (Civil Aviation Authorities) were used in the study.

Procedure

A total amount of 144 controller errors were identified in the reports. The classification procedure was divided into two phases: (1) a calibration trial where all incident reports from 1997 were coded by two judges (independently) and afterwards any problems and disagreements in the classification principles were clarified and resolved; (2) a test trial where the incident reports from 1998 (involving 81 events) were independently coded by the same two judges on the basis of clarifications made at the calibration trial.

Study 2 - Simulator Scenarios

The goal of this study was to do a small-scale study of how errors are captured in everyday operational situations. For this purpose, error events in simulator scenarios containing realistic everyday scenarios were explored.

Participants

The study was carried out on the basis of a simulator study of En Route Air Traffic Control with video recordings involving ATCO students at the end of their first year of their 3-year training. At this phase, the students had acquired a basic level of skills necessary for carrying out controller tasks and they would subsequently start on on-the-job-training. Recordings from 10 teams (pairs of students) were used. The classifications were made independently by two human factors experts.

Materials

The simulator recordings were obtained from an experiment that was conducted using a real-time ATC simulation facility at the Danish CAA in Kastrup. Each of the scenarios in the experiment began with the instructor giving a handover to the trainee team and finished about 30 minutes afterwards. The operator activities were recorded on video to provide a continuous record of events including recordings from head-mounted cameras on radar and planner controller,a video overview of the scene and audio recordings of radio communication. In addition, strips from the simulator scenarios were obtained. A total number of about 60 scenarios were initially available. Among these 10 scenarios were selected for further analysis. The scenarios were selected on the basis that for this subset of scenarios, an instructor was present and this was considered important because the instructor would often comment on things that were not clearly visible on the video recordings.

Procedure

The scenarios were segmented into a number of errors, each of which was a relatively self-contained description. For these errors, both the information related to the communication between the two trainee controllers and, when relevant, contextual information was added to the descriptions to give an understanding of what was going on. In order to get a more complete understanding of the scenarios – and in particular the individual error events - it was decided to recruit comments from two "ATC-experts" present (i.e. instructors) helping in reviewing the video tapes and providing insights about central episodes within each of the scenarios.

The classification part of the study was divided into three sub-phases: (a) *Initial classification*. This consisted in all the error events being classified twice by the first author with one month's interval and on this basis, a consensus classification was produced; (b) *Calibration phase.* In this phase, the second classifier was trained in the use of the taxonomy. In order to achieve reliable and valid categorizations the classifier received feedback about the "correct" or intended classification as a part of the training (based on the author's classifications) and potential misunderstandings in the use of the taxonomy were calibrated. The first four scenarios containing 98 error events were used for this purpose; (c) *Test phase*. On the basis of the transcripts from both the simulator episodes containing errors and instructors' comments about the errors the trained observer classified the remaining errors and error recoveries observed in the simulator trials.

Study 3 - Elicitation of Expert Evaluation

The purpose of this study was to get an expert evaluation of the framework. The focus was on obtaining input concerning the usability of the framework from human factors experts who have experience with developing and/or applying conceptual human factors frameworks. Furthermore, these researchers bring along experiences from many different domains, which is useful when considering the more general applicability of the framework.

Participants

A series of participants were selected based on the criterion that each expert had been involved in research that was highly relevant in relation to this project. This included other conceptual and empirical work related to human error and error management and, in particular, researchers who had been involved in development of comprehensive conceptual human factors frameworks were considered highly relevant for the current project. None of the researchers selected had been involved in or consulted during the development of the taxonomy presented here. A total of 21 researchers were identified and 11 of these responded to the questionnaire (response rate: 52.3%). A few additional responses were received in the form of informal comments.

Procedure

Each of the participants in the survey received two documents. The first contained a short description of the framework and its main components. The second contained a questionnaire asking the respondents to give their opinion about both the components and the overall structure of the framework. The main focus was on the relevance of the individual items within the framework. The elicitation of expert opinion was made in terms of a Likert-like rating scale. In addition, the subjects were encouraged to give free-text comments to the individual items.

Study 4 - The Critical Incident Technique

To be able to evaluate the comprehensive version error management taxonomy, it was necessary to obtain some descriptions of authentic episodes where the issue of error management was important. For this purpose, a series of interviews with controllers were made using the critical incident technique (Flanagan, 1954). Using this technique, it is possible to obtain descriptions of potentially critical episodes and to elicit information about recovery related aspects of these situations that are rarely described in the incident reports.

Participants

Experienced controllers were selected for the interviews because they were most likely to have encountered one or several critical situations during their professional career. No specific preferences concerning the operational positions were made. Twenty-five ATCOs were interviewed and all the participants were from Scandinavian countries. Participation was voluntary. For the classification of error events two human factors experts made an independent classification of error events.

Procedure

The ATCOs were interviewed and their narratives were recorded on tape. From the interviewees 43 episodes were elicited. Each of the episodes was, after the interview, carefully converted into a coherent written description of the event and only information obtained from the interviewee was included. In some cases, the interviewees were contacted again concerning information that was either missing or was unclear.

The procedure for classification of the error events was similar to the one in study 2 and consisted of three sub-phases: (a) *Initial classification*. All the error events (96) and PSFs (106) were classified twice by the author with one month's interval and a consensus classification was produced; (b) *Calibration phase*. The second classifier was trained in the use of the taxonomy. In order to achieve reliable and valid categorizations the classifier was given feedback concerning the "correct" or intended classification as a part of the training (based on the first authors' classifications) and potential misunderstandings in the use of the taxonomy were calibrated. The first eight scenarios containing 17 error events and

24 PSFs were used for this purpose; (c) *Test phase* - On the basis of the remaining incidents the trained observer was asked to classify the errors and error recoveries (79) as well as the PSFs (82).

Hypotheses

Diagnosticity is a crucial aspect of the validation and concerns whether the framework is able to produce insight into the area of interest. A way to explore this is by examining whether the framework is able to produce results that are in conformity with existing empirical results as well as theoretical inferences. Some a priori hypotheses about error and error management *and* the relationship between dimensions within the framework are given below. The empirical studies in this paper will be used to explore whether the framework can verify these statements.

- Hypothesis 1: *Long-term memory errors will be more frequent among novices*. The argument for this is basically that long-term memory (LTM) errors most frequently result from insufficient experience with a certain task. Therefore, experienced controllers should be less susceptible to LTM errors compared with novices.
- Hypothesis 2: *Response Execution errors are most frequently self-detected.* A series of studies have previously established this relationship. For example, in a study from an ATC microworld it was shown that the errors most frequently recovered were slips and to a far lesser extent rule- and knowledge-based mistakes (Wioland & Amalberti, 1998). The explanation for this is that the criteria for successful performance are to a large extent directly available in the head of the error perpetrator in the case of response execution errors. The chances of discovering response execution errors may, on the other hand, be more difficult for external observers because they do not have access to the intentions underlying the behaviour (Wioland & Doireau, 1995).
- Hypothesis 3: *Decision-making errors are either not detected at all or detected by others.* In an experimental study of emergency scenarios in a nuclear power plant by Woods (1984) it was determined that none of the diagnostic errors (i.e. a subgroup of decision-making errors) were noticed by the operators themselves. On the other hand, the diagnostic errors that were detected were discovered by external agents with "fresh eyes." In a study by Wioland and Doireau (1995) where pilots and instructors viewed video recordings of scenarios where actor pilots committed errors it was demonstrated that only a small part of the inserted errors were discovered by the observers. However, the subset of errors that were discovered had a tendency to be associated with problem solving and decision-making (i.e. rule- and knowledge-based) rather than slips. In short, both of these studies indicate that decision-making errors will be difficult to discover and do frequently happen through the assistance of others.

- Hypothesis 4: *Long-term memory errors are either not detected at all or detected by others*. LTM errors share some of the characteristics with decision-making errors and are therefore also expected frequently either not to be detected or to be detected by others. However, LTM errors are probably easier to discover (and agree upon) than decision-making errors by a trained observer insofar as a standard for determining successful performance might be more readily available.
- Hypothesis 5: *Error detection by others depends on the amount of context-sharing.* Several researchers have suggested that the amount of context-sharing is critical for the chances of discovering errors committed by others (e.g. Wioland & Amalberti, 1998).
- Hypothesis 6: *Response Execution will be more frequently detected at the execution stage*. Basically, if people have a clear expectation concerning what action they intended to carry out it should be easy to detect their execution errors by comparing the action carried out with what they felt, saw, or heard. Studies have shown that response execution errors are frequently caught and corrected by a direct feedback-checking (e.g. Rabbitt, 1966).
- Hypothesis 7: Errors found in incident reports will have a tendency to be more frequently detected at the outcome stage compared with errors committed in normal operations. Errors that are detected at the planning or execution stage will tend to be omitted from incident reports, because they are not considered relevant for the investigation. That is, since the focus is on factors that directly or indirectly affected the incident – and not factors that could have affected the situation if not caught at such an early stage – they will not be described in the incident report. Instead, these fast and effective corrections will only be apparent when observing normal operations.
- Hypothesis 8: *The problem-solving process associated with error recovery will vary in such a way that 'Ignore'/'Apply rule' will be most frequent and 'Choose option'/'Create solution' the least frequent*. The reason for this expectation is that the taxonomy is here very similar to Rasmussen's Skills-Rules-Knowledge-model (Rasmussen, 1984). Within the SRK-framework it is postulated that the behaviour of experienced operators most of the time will be controlled at the lower resource demanding levels (skill- and rule-based level) and only rarely is it required to move up to the resource intensive level (knowledge-based level). In the current framework the "Ignore" and "Apply rule" are the cognitive processes that require the least mental resources – that is, a straightforward recovery solution is available in the situation. On the other hand, the categories "Choose option" and "Create solution" are associated with increasingly more cognitively demanding recovery situations.
- Hypothesis 9: *The errors that are ignored and tolerated are more inconsequential than the errors that are responded to.* In a study based on an ATC microworld by Wioland and Amalberti (1998), it was demon-

strated that with increased expertise – and thereby better knowledge of the system and its risks – the subjects tolerated a larger degree of errors without consequences. This is most likely related to the fact that the subjects learn that certain errors are without consequences and consequently they can save resources by not correcting them.

- Hypothesis 10: *Response execution errors (including speech or action errors) are more often detected than other errors (e.g. lapses and mistakes).* This is related to the fact that there is a direct discrepancy between intention and corresponding action or outcome and, consequently this should be easy to discover.
- Hypothesis 11: *Decision-making errors are more often associated with undesired states*. This is, for example, supported by a study by Klinect et al. (1999) that showed that operational decision errors as well as proficiency errors were the most difficult for the flight crews to manage and, consequently, were the ones that most often had consequences. The explanation for this is that for decision-making errors (including reasoning, judgment and diagnosis), the criterion for detection is not directly available in the head of the individual, but instead the correct solution is only available in the external world and is often not clearly recognizable in advance (Reason, 1990).
- Hypothesis 12: *Most errors in everyday-life situations will be inconsequential*. For example, in an observational study of pilot crew errors during normal operations it was found that about 85% of the crew errors were inconsequential (Klinect et al. 1999). Therefore, a larger amount of consequential errors is expected to be found in incident reports compared with real-time observation.
- Hypothesis 13: *Procedural violations will frequently be inconsequential.* A study by Klinect et al. (1999) based on real-time observation of pilots' behavior showed that procedural violation errors were the most frequently committed errors and also the least consequential. It may be speculated that the reason for this is that people develop a meta-knowledge based on experience concerning which violations are consequential and which are not. This would be in agreement with studies indicating that people develop natural risk-taking abilities and that their main goal is not to avoid errors, but instead to maintain cognitive control (Wioland & Amalberti, 1996). Consequently, many "minor" violations might be accepted, because the risk is considered small or absent.

Results

The empirical studies described previously can be used to shed some light on the utility of the conceptual framework. To determine the utility it is reasonable to take a look at how it satisfies the previously mentioned product criteria - namely reliability, comprehensiveness, diagnosticity, and usability.

Reliability

A critical measure of the utility of the framework was the intra- and inter-rater reliability of the classifications in the three studies based on incident reports, a simulator study, and interviews based on the critical incident technique. Kappa is a statistical measure that is commonly used to determine the reliability of classifications made by independent judges and which is corrected for chance agreement (Cohen, 1960). Kappa-results for all dimensions from these three studies are shown in the table below (two raters were used in the inter-rater analysis). The interpretation of the level of agreement (above chance) obtained by independent raters is that the closer the Kappa value is to 1 the better is the agreement. Furthermore, levels below 0.40 corresponds to poor or merely fair agreement, and this figure remains a conventional cut-off point (rather as the interpretation of a p-value at or below 0.05). All except two of the analyses provided p-values below 0.001 (namely "procedural violation" and "detection" in study 4 where the p-values were 0,004 and 0,119, respectively).

Table 1
Kappa Coefficients from Study 1, 2, and 4

		Empirical studies				
		Incidents (study 1)	Simulator study (study 2)		Critical incidents (study 4)	
Main dimension	Sub-dimension	Inter-rater	Intra-rater	Inter-rater	Intra-rater	Inter-rater
Threat	Anticipation				1.00	0.80
Error	Producer		0.90	0.95	0.94	0.95
	Cognitive domain	0.81	0.86	0.69	1.00	0.52
	Procedural violation				0.81	0.72
Who	Detector	0.64	0.94	0.81	0.90	0.89
	Corrector	0.62	0.88	0.69	0.85	0.54
When	Detection	0.56	0.89	0.60	0.66	0.27
How	Detection	0.62	0.84	0.68	0.87	0.71
	Correction				0.97	0.60
What	Response	0.45	0.94	0.80	0.96	0.76
	Outcome	0.51	0.74	0.50	0.79	0.69
Why (PSFs)	Main groups				0.97	0.64
	Individual categories				0.88	0.61

As can be seen in the table the overall picture from the empirical studies is that the framework did in fact support highly reliable classifications throughout all of the empirical studies (the empty cells mean that no classifications have been made because the available data material did not provide the necessary information or the issue was not explored). Only once did the reliability results get below the critical 0.40 cut-off point, namely for the when-dimension. This dimension is associated with some difficulties applying the categories consistently to real-world cases and should perhaps be removed from the framework.

Comprehensiveness

Comprehensiveness is related to the extent to which the framework is able to cover all the main categories and issues associated with the area of error management. A way to determine the extent to which the categories adequately reflect the natural variation of a phenomenon is to examine the amount of unknown classifications used within the individual dimensions. An analysis of the amount of unknown classifications in the simulator study and the critical incident study has been done. In the simulator study, the highest rate of unknown classifications was found for Cognitive Domain and was 4.4%. In the critical incident study, 12 out of 17 dimensions did not contain any unknown categories. The dimension receiving the highest amount of unknown classifications was procedural violations (23.5%). Hence, the categories within the framework were able to account for almost all of the error events analyzed in the two empirical studies. In the few cases where a relatively high level of unknown classifications was found this was mainly related to insufficient information available in the data material (in particular from the pilot's perspective). In the case with the high amount of unknown classifications for procedural violations, this was basically a reflection of the fact that the classifiers did not possess a sufficient degree of domain knowledge to make the classification in these cases. In the analysis of the PSFs, similar positive results were obtained insofar as only known categories were used.

Diagnosticity

To establish the diagnosticity 13 hypotheses concerning how the framework should "behave" have been formulated. The question is whether the results from the pattern analysis corresponded with theoretical expectations and previous research. Different statistical measures will be used. For the statistical analysis of main effects, we use an exact Goodness-of-Fit test based on a uniform distribution (i.e. the observed distribution is compared with a distribution where each of the categories has an equal likelihood of occurring). For the analysis of interaction effects, Pearson's Exact Test will be used to determine whether different dimensions are independent or not. In addition, to determine which specific cells contribute most to the interaction effect, adjusted residuals (AR) can be used. According to Agresti (1996), the adjusted residual follows a standard normal distribution (which is similar to the z distribution). Agresti (1996) stated: "Adjusted residuals larger than about 2 in absolute value are worthy of attention, though one expects some values of this size by chance alone when the number of categories is large" (p. 91).

The results from the analysis of the a prior hypotheses are shown below.The table shows which comparisons were made (and the studies used) to explore

Table 2

The Framework's diagnosticity

Hypotheses & Comparisons	Significance	
	Chi-Square (table)	AR (cell)
Hypothesis 1: Long-term memory errors will be most frequent among novices. Comparison: Dependence between cognitive domain and data material (Study 2 vs.4)	Study 2 vs. 4: $X^2(4, N=391)=84.93$, $p<0.001$	$z=5.20$, $p<0.001$
Hypothesis 2: Response Execution errors are most frequently self-detected. Comparison: Dependence between cognitive domain and error detector (Study 2)	Study 2: $X^2(16, N=202)=129.35$, $p<0.001$	$z=7.64$, $p<0.001$
Hypothesis 3: Decision-making errors are either not detected at all or detected by others. Comparison: Dependence between cognitive domain and error detector (Study 1, 2 and 4)	Study 1: $X^2(9, N=64)=36.12$, $p<0.001$ Study 2: $X^2(16, N=202)=129.35$, $p<0.001$. Study 4: $X^2(4, N=48)=32.83$, $p<0.001$	$z=5.32$, $p<0.001$ 1z=2.10, $p=0.036$ $z=-2.57$, $p=0.010$
Hypothesis 4: Long-term memory errors are either not detected at all or detected by others. Comparison: Dependence between cognitive domain and error detector (Study 2)	Study 2: $X^2(16, N=202)=129.35$, $p<0.001$	$z=6.28$, $p<0.001$ $z=2.80$, $p=0.005$
Hypothesis 5: Error detection by others depends on the amount of context-sharing. Comparison: Dependence between the errors committed by ATCOs and the errors detected by pilots and ATCO colleagues, respectively (Study 1,2 and 4)	Study 1,2 and 4: $X^2(1, N=90)=25.13$, $p<0.001$	
Hypothesis 6: Response Execution will be more frequently detected at the execution stage. Comparison: Dependence between cognitive domain and detection stage	Study 2: $X^2(8, N=156)=99.17$, $p<0.001$	$z=7.98$, $p<0.001$
Hypothesis 7: Errors found in incident reports will have a tendency to be more frequently detected at the outcome stage compared with errors committed in normal operations. Comparison: Dependence between cognitive domain and data material (Study 2 vs.4)	Study 2 vs. 4: $X^2(2, N=346)=87.58$, $p<0.001$	
Hypothesis 8: The problem-solving process associated with error recovery will vary in such a way that 'Ignore'/'Apply rule' will be most frequent and 'Choose option'/'Create solution' the least frequent. Comparison: Distribution of error correction/problem-solving (study 4)	Study 4: $X^2(3, N=63)=55.41$, $p<0.001$	
Hypothesis 9: The errors that are ignored and tolerated are, to a larger extent, inconsequential than the errors that are responded to. Comparison: The dependence between the decision-making associated with the correction of errors (ignore vs. respond) and the outcome (consequential vs. inconsequential)	Study 4: $X^2(1, N=76)=3.60$, $p=0.067$	$z=1.90$, $p=0.057$
Hypothesis 10: Response execution errors (including speech or action errors) are more often detected than other errors (e.g. lapses and mistakes). Comparison: Dependence between cognitive domain and data material (Study 2 vs.4)	Study 2 vs. 4: $X^2(4, N=391)=84.93$, $p<0.001$	$z=3.9$, $p<0.001$
Hypothesis 11: Decision-making errors are more often associated with undesired states. Comparison: Dependence between cognitive domain and data material (Study 2 vs.4)	Study 2 vs. 4: $X^2(4, N=391)=84.93$, $p<0.001$	$z=8.7$, $p<0.001$
Hypothesis 12: Most errors in everyday-life everyday situations will be inconsequential. Comparison: Dependence between cognitive domain and data material (Study 2 vs.4)	Study 2 vs. 4: $X^2(2, N=438)=150.87$, $p<0.001$	$z=11.9$, $p<0.001$
Hypothesis 13: Procedural violations will frequently be inconsequential. Comparison: Dependence between procedural violation and outcome (study 1 and 4)	Study 1 and 4: $X^2(2, N=68)=6.35$, $p=0.037$	$z=2.30$, $p=0.021$

the a prior hypotheses. For each of the comparisons the overall statistical significance level is given for the associated cross-tabulation. Since most of the hypotheses were related to specific cells in the cross-tabulations the ARs are included to determine whether the hypotheses could be confirmed or not. As can be seen, 12 out of the 13 hypotheses were confirmed (the only one that could not be confirmed was Hypothesis 9). Hence, the framework showed a high level of diagnosticity.

Usability

Even though the framework is comprehensive, attempts were made to ensure that it still maintained its usability. This included both employing categories that were easy to understand and at the same time were practically relevant.

Are the categories easy to understand and use? Deliberate attempts were made to optimize the usability of the framework. For example, the categories associated with the detection source were chosen because they seemed to make intuitive sense. In principle, it was possible to split this dimension into much finer-grained categories as suggested by some researchers. However, it was expected that this would jeopardize the reliability and usability of the taxonomy by introducing very subtle distinctions. It was therefore chosen to use a more coarse-grained but also much more easy-to-use categorization. Similar deliberations were made in relation to other dimensions such as the "how-correction dimension" where the problem-solving terms were chosen on the basis of being easy to grasp. Even though considerations have been given to developing a framework with a high degree of usability, there is a trade-off between diagnosticity and usability. Therefore, the framework will require some degree of familiarization and training before it is possible to apply the concepts in a consistent manner (e.g. in study 4 approximately 10 hours were spent on the training and calibration of the second rater who initially was unfamiliar with the taxonomy, though not with ATC and ATCO tasks).

Are the categories practically relevant? In the questionnaire study human factors experts were asked to give their opinion about the relevance of the framework. The results showed that both the overall structure of the framework as well as the individual dimensions received a high level of expert acceptance: On a four-point scale containing "Irrelevant" (1), "Somewhat irrelevant" (2), "Somewhat relevant" (3) and "Highly relevant" (4) all average ratings were somewhere between "Somewhat relevant" and "Highly relevant". Furthermore, several comments also indicated that framework could be relevant in error management studies in other contexts – such as the maritime domain and process control.

Table 3

Average Respondent Ratings from Questionnaire

Framework core	Rating	PSFs	Rating
1. Threat management	3,8	1. Traffic, airport, and airspace	3,8
2. Cognitive domain	3,4	2. Ambient environment	3,3
3. Detector/corrector	3,8	3. Procedures and documentation	3,7
4. Detection stage	3,2	4. Workplace design, HMI, and equipment	3,8
5. Detection source	3,7	5. Training and experience	3,8
6. Correction process	3,8	6. Person related factors	3,4
7. Response	3,4	7. Social and team factors	3,9
8. Outcome	3,4	8. Company, management and regulatory factors	3,7

Conclusions

The purpose of this paper was to present and validate an error management framework. Four different empirical studies were conducted for the validation. The empirical studies showed that reliable classifications could be obtained by the use of the framework. Both on the basis of intra- and inter-rater analyses a high level of reliability was obtained. In addition, analyses of patterns of both main effects and interactions between dimensions provided interesting insights. In particular, the analysis of interaction between dimensions was useful in supporting the analysis of the criterion validity of the framework. Out of the 13 a priori defined hypotheses - based on theoretical expectations and previous research – all but one were confirmed by testing the framework against the data material from the empirical studies. Since studies within this area have been somewhat hampered by the lack of a comprehensive framework, we believe it is an important quality to the present framework. The results, therefore, seem to demonstrate the usefulness of the framework. Furthermore, the framework includes several dimensions that have been scantily explored in the literature and might therefore pave the way for new studies of error management processes.

Future Research Directions

The air traffic control system is undergoing many challenging changes in the near future that might have implications for both the human errors that will occur and the chances of discovering and recovering from these errors. The impetus for these changes grows out of the fact that the system is currently stretched to its capacity limit and rapid increases in traffic levels are envisaged for the near future. To be able to accommodate this development it is necessary to implement new equipment as well as considering new procedures for regulating the air traffic. With the advent of new technologies and new operating philosophies there is a risk that new types of errors will emerge and at the same time the chances of recovery might be diminished. Since human errors cannot be completely avoided, it is of paramount importance for safety to be maintained so that the human operators' powerful human recovery abilities are not undermined, but actively supported. Hence, the development and evaluation of future initiatives aimed at safely enhancing the capacity of the air traffic system will require careful consideration of error and error management profiles.

References

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: Wiley.

Amalberti, R. (2001). The paradoxes of almost totally safe transportation. *Safety Science*, *37*, 109-126.

Amalberti, R. & Wioland, L. (1997). Human error in aviation. Invited paper to the International Aviation Safety Conference (IASC-97) Rotterdam Airport, The Netherlands. In H. Soekkha (ed), *Aviation Safety*, VSP, Utrech, 91-108.

Bove, T. (2002). *Development and Validation of a Human Error Management Taxonomy in Air Traffic Control*. Ph.D. dissertation. Risoe-R-1378(EN). Risoe National Laboratory, Roskilde, Denmark. Retrieved March 2003 from http://www.risoe.dk/rispubl/SYS/ris-r-1378.htm.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327-358.

Frese, M. (1991). Error management or error prevention: Two strategies to deal with errors in software design. In H.J. Bullinger (Ed.), *Human aspects in computing: Design and use of interactive systems and work with terminals* (pp. 776-782). Elsevier Science Publishers.

Helmreich, R. L., Klinect, J. R., & Wilhelm, J. A. (1999). Models of threat, error, and CRM in flight operations. In *Proceedings of the Tenth International Symposium on Aviation Psychology (pp. 677-682)*. Columbus, OH: The Ohio State University.

Helmreich, R. L., Wilhelm, J. A., Klinect, J. R., & Merritt, A. C. (2001). Culture, error and crew resource management. In E. Salas, C.A. Bowers, & E. Edens (Eds.), *Improving Teamwork in Organizations: Applications of Resource Management Training* (pp.305-331). Hillsdale, NJ: Erlbaum.

Isaac, A., Shorrock, S. T., Kirwan, B, Kennedy, R, Andersen, H. B., & Bove, T. (2000). Learning from the past to protect the future – the HERA Approach. Presented at the *24th Conference of European Association For Aviation Psychology (EAAP) 2000*, Crieff (Scotland).

Kanse, L. & Van der Schaaf, T. (2001a). Recovery from failures in the chemical process industry. *International Journal of Cognitive Ergonomics, 5* (3), 199-211.

Kanse, L. & Van der Schaaf, T. (2001b). Factors influencing recovery from failures. In R. Onken (Ed.), CSAPC '01, *8th European Conference on Cognitive Science Approaches to Process Control* (DGLR-Bericht 2001-04) (pp. 123-136). Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt.

Kinney, G. C., Spahn, M. J., & Amato, R. A. (1977). *The human element in air traffic control: Observations and analyses of the performance of controllers and supervisors in providing ATC separation services*. METRIEK Division of the MITRE Corporation: MTR-7655.

Kontogiannis, T. (1999). User strategies in recovering from errors in man-machine systems. *Safety Science*, 32, 49-68.

O'Leary, M. (1999). The British Airways Human Factors Reporting Programme. In D. Javaux (Ed.), *3rd Workshop on human error, safety, and system development (HESSD '99)*, Liège (BE), 7-8 June 1999. University of Liege, Belgium.

Orasanu, J., & Fischer, U. (1997). Finding decisions in natural environments:

The view from the cockpit. In Zsambok, C.E. & Klein, G (Eds.) *Naturalistic Decision Making* (pp. 343-357). Lawrence Erlbaum Associates, Publishers. Mahwah, New Jersey.

Rabbitt, P.M. A. (1966): Errors and error-correction in choice-response tasks. *Journal of Experimental Psychology, 71*, 264-272.

Rasmussen, J. (1984): *Human error data. Facts or fiction?* Invited paper presented at a seminar on Accident Research in Rovaniemi, Finland, April 1984. Risoe-M-2499, 1984. Risoe National Laboratory, Roskilde, Denmark.

Reason, J. (1990). *Human Error*. Cambridge University Press. Cambridge.

Rizzo, A., Ferrante, D., & Bagnara, S. (1995). Handling human error. In J-M Hoc, P.C. Cacciabue, E. Hollnagel (Eds.) *Expertise and Technology* (pp. 195-212). Hiilsdale, NJ: Lawrence Erlbaum.

Shorrock, S. T. & Kirwan, B. (1998). The development of TRACEr: a technique for the retrospective analysis of cognitive errors in ATC. Paper presented at the *2nd Conference on Engineering Psychology and Cognitive Ergonomics*. Oxford: 28 - 30 October.

W iegmann, D. A. & Shappell, S.A. (2002). Human error perspectives in aviation. *The International Journal of Aviation Psychology*, *11* (4), 341-357.

W ioland, L. & Amalberti, R. (1996). When errors serve safety: towards a model of ecological safety. Paper presented at *CSEPC '96, Cognitive Systems Engineering in Process Control*, Kyoto, Japan, November 1996, 184-191.

W ioland, L. & Amalberti, R (1998): Human error management: towards an ecological safety model. A case study in an Air Traffic Control microworld. Paper presented at the *9th European Conference on Cognitive Ergonomics* - Limerick University, Ireland August 24-26.

W ioland, L. & Doireau, P. (1995): Detection of human error by an outside observer: a case study in aviation. *5th European conference on cognitive science approaches to process control*. Espoo, Finland, Aug. 30 - Sept. 1, 1995. Norros, L. (Ed.)

Woods, D. D. (1984). Some results on operator performance in emergency events. *Institute of Chemical Engineers Symposium on Ergonomics Problems in Process Operations*. Birmingham, July 11-13, 1984.

International Journal of Applied Aviation Studies, 3,1
Copyright © 2003, FAA Academy, Oklahoma City, OK

CRM training in the New Environment: Challenges for flight crew training after September 11, 2001

Clint Bowers
University of Central Florida
College of Arts & Sciences
CAS 190Y
Orlando, FL 32816-1990
Email: bowers@mail.ucf.edu

Florian Jentsch
University of Central Florida
Orlando, FL 32816-1990
Email: fjentsch@pegasus.cc.ucf.edu

and

Eduardo Salas
University of Central Florida and
Institute for Simulation and Training (IST)
Human Systems Integration Research Department
Institute for Simulation and Training
3280 Progress Drive
Orlando FL 32826-0544
Email esalas@pegasus.cc.ucf.edu

Abstract

The events of September 11, 2001 have created a new set of security challenges for the aviation industry, including airports, airlines, pilots, passengers, and regulators. Flight deck crews have formed one of the linchpins in a system of prediction, prevention, and response. The first need for new training therefore was to focus on giving flight crews the declarative, procedural, and strategic knowl-

---

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

edge to look for, spot, report, assess, and respond to threats. In studying how to best train flight crews for the new security challenges, it might be advantageous to leverage the already existing concepts and training in Cockpit/Crew Resource Management (CRM). Towards that end, the current paper examines the content and competencies engendered by traditional CRM programs and compares those to the needs of the new environment after September 11, 2001. In so doing, we offer additions to training contents and procedures that should make CRM training programs more effective in responding to the needs described above.

Introduction

On September 11, 2001, four teams of radical-Islamic terrorists hijacked four U.S. domestic commercial airline flights. Using knives and box cutters, the terrorists succeeded in getting access to the airliners' flight decks, then took control of the aircraft, and crashed two of the aircraft into the World Trade Center Twin Towers in New York City and one into the Pentagon in Washington, DC. The fourth aircraft, United Flight 93 from Newark, NJ, to San Francisco, CA, changed course towards Washington, DC, but crashed in a remote area of Pennsylvania after what appeared to be a struggle between the passengers and the hijackers. The human toll of these four incidents of more than 3,000 killed or missing and many thousands more injured, as well as the material damage to aircraft and structures, made this the worst terrorist attack in U.S. history (AbuKhalil, 2002).

The events of September 11, 2001 created a new set of security challenges for the aviation industry, including airports, airlines, pilots, passengers, and regulators. Immediately following the attacks, the Federal Aviation Administration (FAA) took the unprecedented step of closing the entire U.S. airspace for non-military flights. All aircraft in the air had to land immediately at the nearest airport; when the system was closed down on September 11, FAA air traffic controllers landed 2,800 aircraft within 54 minutes (Federal Aviation Administration, 2002). Furthermore, certain flight activities were prohibited or severely limited for an extended period of time, and certain airspace regions (i.e., the Washington, DC airspace including Washington's Ronald Reagan National Airport) were temporarily closed. Other responses to the security threats posed by domestic and international terrorism initiated after September 11 included the creation of a new Transportation Security Administration (TSA), the concurrent federalization of airport security screeners, a revival and expansion of the Federal Air Marshal program, and the reinforcement of cockpit doors. More far-reaching proposals have included arming cockpit crews for their protection, the use of "stun gas" to incapacitate would-be hijackers, and the development of technology to remotely control aircraft to a landing if the crew is overpowered by attackers. Finally, there are looming issues regarding the potential use of aircraft to conduct terrorist attacks with weapons of mass destruction, specifically biological and chemical agents. In all, the White House Commission on Aviation Safety and Security that was created after September 11, 2001, has made 57 recommendations, 31of which pertained to aviation security (Federal Aviation Administration, 2002). Thus, since September 11, a number of new challenges have come to the forefront of attention with respect to the knowledges, skills, and abilities (KSAs) required of flight crews.

However, even before September 11, there were a number of new security-related challenges for flight crews. Specifically, the increased incidence of "air rage" put pilots and cabin crews into situations for which they had previously not been trained (Anonymous & Thomas, 2001): Between 1995 and 2000, the number of reported air rage incidents in the FAA data base increased from 146 in 1995 to 314 in 2000 (i.e., one major incident almost every day). Similar increases in air rage incidents were reported elsewhere in the world, making air rage a trans-national problem (airWise, 2000). In fact, according to one source (airWise, 2001), air rage incidents worldwide quadrupled between 1996 and 1999, up to over 5,000 incidents worldwide in 1999. Thus, in addition to the issues of the prediction, prevention, and response to air rage incidents (Berkley & Mohammad, 2001; Bor, Russell, Parker, & Papadopoulos, 2001), there have been trans-national legal and multi-cultural questions on which pilots must be briefed (Huang, 2001). Thus, there has been a need for comprehensive steps towards improved measures to ensure the safety and security of the flying public, specifically as far as the training of cockpit and cabin crews is concerned.

Responding to the Need

The new threats after September 11, 2001 have been a system-wide phenomenon and could only be responded to through a concerted effort among all participants in the aviation system (Baldwin, 2000; Wilkinson & Jenkins, 1999). Flight deck crews thus have formed one of the linchpins in a system of prediction, prevention, and response. The first need for new training therefore was to focus on giving flight crews the declarative, procedural, and strategic knowledge to look for, spot, report, assess, and respond to threats. For example, flight crews need to know what information is necessary and helpful in predicting and responding to threats. Crews thus need to be provided with training that helps them to seek out, categorize, and report information that may assist in the prediction and prevention of emerging threats.

To respond to these new need areas, the FAA has developed a list of training issues that should be addressed as part of what has become known as "Common Strategy II" (Reiley, 2002). However, the FAA did not issue specific guidelines as to the detailed training content and training media. While this allowed airlines to customize their training to their specific situation, it did not provide much guidance as to what the critical knowledge, skills, and abilities (KSAs) for security training were, nor how these KSAs could be imparted in the most effective and efficacious way.

In studying how to best train flight crews for the new security challenges, it might be advantageous to leverage the already existing concepts and training in Cockpit/Crew Resource Management (CRM). CRM was developed to respond to incidents in which human error was responsible for aviation mishaps. The goal of these programs has been to assist crews in using all of the intellectual resources available to them in dealing with crises. CRM has been a very successful and well-accepted program in the industry (General Accounting Office, 1997). However, it was not designed to address the complexities of the current environment. Towards that end, the current paper examines how the content and competencies engendered by traditional CRM programs should be expanded to address the needs of the new environment after September 11, 2001. In so doing, we offer several additions that might make CRM training programs more effective in responding to the needs described above. We will refer to the revised training as "Enhanced CRM" for the New Environment to differentiate it from traditional CRM programs. In the following sections, we will describe the knowledge required by this type of new CRM training, as well as its skill competencies, and its tools and methods.

Training Content

Traditional CRM training programs typically do not emphasize declarative knowledge, that is, the knowledge of facts in the domain of coordination and cooperation. Instead, it is presumed that aviators know most of the flight-related facts required to coordinate and communicate. Instead, the goal of traditional CRM is (a) to create more positive attitudes to increase the likelihood of coordination behaviors and (b) to practice the skills necessary for their application. However, as described above, the new environment after September 11, 2001 has generated the need for modern flight crew members to know a host of things that may not be within their general sphere of knowledge. For example, knowledge regarding potential threats includes such issues as a basic understanding of the types of persons and groups that pose a threat to commercial aviation, their tools and methods, and of appropriate responses. In the following sections, we discuss a number of these areas.

Skyjacker types. In our recent efforts to support flight crew training for the new environment whose contents we describe in more detail in the Appendix, we discussed different categories of aircraft hijackers. Based on research from areas as diverse as clinical (Hubbard, 1973), social (Crenshaw, 1988), and small group psychology (Hudson, 2002), law enforcement (e.g., Snow, 1996), and aviation security (Wilkinson & Jenkins, 1999), we were able to describe five different types of skyjackers; we labeled these as (a) the "Desperate Loser," (b) the "Terrorist," (c) the "Suicidal Candidate/Mental Patient," (d) the "Fugitive," and (e) the "Ordinary Criminal." Although many of the behaviors that some of these categories of skyjackers (e.g., the "terrorist" and "desperate loser") would demonstrate are similar, they have vastly different motivations for their actions. Additionally, all types of skyjackers typically have completely different motivations from a disruptive passenger who may present himself in a similarly threatening way (cf. Beeks, 2000; James & Nahl, n.d.; Sheffer, n.d.). The ability to identify a type of skyjacker and to know his/her motivations thus might be very helpful in assisting crews to choose the correct course of action.

Terrorist groups. By the same token, it is important that crews have a basic understanding of larger groups that might threaten the flight. Up to the events of September 11, 2001, it was the common belief that passive resistance to hijackings was the most promising strategy to avoid human casualties (Choi, 1994; Wilkinson & Jenkins, 1999) regardless of the motivations of the hijackers. However, on September 11, the hijackers replaced the flight crews with trained pilots, then used the aircraft for suicide attacks on highly visible and populated targets. In this case, a strategy of passive resistance was doomed to failure, and one of active resistance, as demonstrated by the success of the passengers of United Airlines Flight 93 in avoiding further casualties on the ground, would have been more promising. This is, obviously, a difficult decision to make "on the fly." However, understanding the motivations of the terrorist group might have helped the crew with this difficult decision.

In order to choose the best possible response to an attack, it is important that crews understand the methods employed by potential terrorists. This includes understanding the weapons that are most likely to be used in an attack, either because they are effective, favored by terrorists, or less difficult to hide from airport security screenings (cf. Gero, 1997; National Research Council, 1996). This includes such "unconventional" devices as the bombs that were hidden in "shoe-bomber'" Richard Reid's hiking boots. As was demonstrated in the Reid case, crew members need to understand the risks and weaknesses of each weapon and the best methods to respond to each. Finally, crews need to understand the implication of each weapon to the likely impact on the airworthiness of the aircraft. A number of publications are available which describe different bombs and explosive devices (e.g., Brodie, 1996; Lenz, 1976; Pickett, 1998), but these books are typically not aimed at flight deck and cabin crew. An effective sifting and translation of the information in these documents is therefore needed to create good training.

Chemical/biological threats. Because cockpit and cabin crews would likely be the first to come in contact with an attacker infected with viral or bacteriological agents who uses a commercial airliner as his/her means of transport, crew members need to be well informed about how to detect these threats and how best to respond to them. In addition to possibly spotting the infected attacker and/or the agent, crew members must know how to detect illness symptoms on themselves so that they can report them to a physician. This is a difficult subject matter, however, as even trained physicians must receive special training on detecting and responding to biological and chemical agents (Frist, 2002; Osterholm & Schwartz 2000). While relatively well-designed handbooks for the recognition of different biological and chemical agents exist (e.g., the *FR-CBH First Responder Chem-Bio Handbook*; Anonymous, 1998), the content in these documents is easily overwhelming and thus must be adapted carefully to present only the critical aspects needed by flight crews.

Electronic threats. Yet another content area regards detection of ground and airplane systems degradation (e.g., GPS and other navigation aids) by flight crews. Future threats against the U.S. civil aviation system will likely include attempts to degrade the air traffic control system and navigation aids (cf. Wilkinson & Jenkins, 1999), such as the Instrument Landing System (ILS) or the Global Positioning System (GPS). Thus, the threat in this area is similar to that of military electronic warfare (cf. De Landa, 1991; Dickson, 1976; Gordon, 1981). Pilots will need to know how to monitor the integrity of the aircraft's navigation systems and how to detect any system degradation. Crews will also need to have declarative (what?), procedural (how?), and strategic (under which circumstances?) knowledge on system characteristics and behaviors, the likely problems resulting from unlawful manipulations, and the indications that can be expected.

Legal questions. Finally, crew members must be informed about the legal implications of operating in the new environment. Recent events have demonstrated very clearly that pilots, flight attendants, and ground crew are all likely to come into contact with irate, even threatening, passengers. It is important that crew members understand their rights and the limitations on their behavior. This is important not just to limit liability, but to reduce hesitancy to act in critical situations. Congress has set favorable conditions for this by creating the equivalent of a "good Samaritan" law for well-meant responses to perceived aviation threats (Section 125 of S.1447ENR, the Aviation and Transportation Security Act of 2001).

Summary. In summary, the security challenges of the new environment after September 11, 2001, have created new and expanded knowledge requirements for flight crews. First, since de-escalating cooperation with skyjackers is no longer a universally appropriate response to hijacking threats, flight crews need to know and be better able to distinguish among different types of skyjackers and terrorist groups. Second, crew members need to know more about the types of threats and weapons used by terrorists. This includes a basic knowledge of explosives, biological and chemical weapons, and of electronic threats. Third, there are a number of new legal issues that flight crew members must be aware of, so that they can provide an appropriate, measured, and legally defensible response to any real and perceived threats. Thus, CRM training for the future will require much more focus on declarative knowledge than before. Finally, since the contents in all three knowledge areas will be continually changing, a one-time training such as that required by "Common Strategy II" will not be sufficient. Instead, regulators and trainers alike will have to periodically review and refresh the Enhanced CRM training to address new and emerging threats.

Skill Competencies

Over and above simply "knowing things," however, the modern flight environment requires that crew members possess skill competencies that are subtly different from those engendered in existing CRM programs (see General Accounting Office, 1997, and Salas, Burke, Bowers, & Wilson, 2001, for reviews of existing CRM programs). In order to illustrate this point more fully, we'll review changes that are required in several competency areas that are often used in team training.

Communication. In traditional CRM programs, communication emphasized the use of standard phraseology, "closing the loop," and other basic communication behaviors. The behaviors are just as important in the new environment, but must be practiced among a much larger range of participants. The FAA has stated that "airports, airlines, and a host of Federal, state, local, and even international agencies and organizations that provide intelligence and enforcement are all security partners" (Federal Aviation Administration, 2002, p. 15). Specifically, the development of training for improved communications among large teams with varied backgrounds, such as those made up of flight deck crews, cabin crews, Federal Air Marshals, and ground security/gate personnel is more important than ever. As the events of September 11, 2001, have demonstrated, there is a need to identify a common vocabulary among those that may be affected by terrorist threats or unlawful acts. Past research has shown that one cannot assume that even groups with daily, close contact (such as cockpit and cabin crews) have a common vocabulary and are able to communicate effectively (Chute & Wiener, 1996). This has been the motivation behind joint crew resource management training for cockpit and cabin crews and CRM training for dispatchers. Providing this type of training to other groups, especially law enforcement personnel such as Federal Air Marshals, thus seems to be a logical and necessary extension.

Finally, cooperation between flight crews, cabin crews, and law enforcement/anti-terrorist personnel is a worldwide challenge and thus has an important cross-cultural component. This includes knowing and understanding differences in approaches among likely hostage-rescue teams (HRTs) and anti-terrorism units worldwide (cf.; Katz, 1997; Micheletti, 1999; Thompson, 1986). While there are many similarities among the units and their tactics, units from different backgrounds have different approaches to solving similar situations. For example, a military anti-terrorism units, such as the British SAS (Brown, 1986), has a slightly different approach than, for example, a specialized police hostage rescue team, such as the German GSG-9 (Scholzen & Froese, 1999), or a urban police emergency response team, such as the LAPD SWAT team (Halberstadt, 1994; Tomajczyk, 1997). In addition, even among emergency response units from the same occupational background and similar cultures (such as the HRTs of the German and Austrian police) there are significant differences (cf. Hufnagl, 1999; Scholzen, 2000). Applying the research on cross-cultural aspects of communication, coordination, and cooperation to security-focused CRM training is therefore an important additional concern (Helmreich, Wilhelm, Klinect, & Merritt, 2001).

A special form of communication competency involves negotiation and conflict resolution. Following the events of September 11, 2001, crews and passengers in future threat situations will have to carefully weigh the situation to determine what strategy is most appropriate in response to the threat. Not only negotiation skills and methods of de-escalation, but also information and practice regarding the use of force may be required to make these types of decisions. This type of training also would extend to other critical decision-making situations, specifically those involving unruly passengers and/or air rage incidents.

Coordination. Coordination refers to the group of behaviors required to share resources effectively among crew members. One resource that is often shared is information. Teams are said to demonstrate good coordination when they provide information before it is needed. Again, the new "players" in the aviation environment both possess and require information. In many instances, there is a requirement to share information, but there has been little training to assist crew members in understanding the needs of one another. Coordination often requires "backing up" behaviors. That is, when individual crew members are over-taxed they are often assisted by other teammates. However, with the recent changes to the cockpit, specifically the installation of an impenetrable barrier in form of the armored and locked cockpit door, these types of behaviors are more difficult to perform. As such, there is a requirement for further research to determine how best to assist crew members with these changes.

Situation awareness. Situation awareness refers to the ability to use data to create an accurate representation of one's condition. The most specific application of this ability is to know where one is in three-dimensional space and to project one's position in the future. In large part, this is done by collecting information from one's instruments to create a mental picture of the situation. However, it is also important that pilots create and maintain such a mental representation based on other data such as time, visual cues, and so forth. In this fashion, crews are best able to detect anomalies in flight systems. As described above, navigation systems are a potential target for future terrorist activity (Wallis, 1993). As such, it is important that crews frequently check the system against other data sources to ensure that the displayed data are valid. This will require crews to change their behavior rather than being reliant on one system. Because many crew members have substantial experience, this new pattern will be difficult to acquire. Thus, training (and job aids) must be developed to assist with this transition.

Decision making. Training for improved decision making among distributed teams is another important consideration. Not unlike other abnormal events in aviation, the new threats in the aviation environment require a response that features the coordination and cooperation of widely distributed, heterogeneous teams. In the case of threats to commercial airline flights, those who have to coordinate and make decisions include flight deck crews, cabin crews, in-flight law-enforcement personnel, and ground units such as air traffic control, dispatch and maintenance facilities, and military/law enforcement agencies. Research on distributed decision making by multidisciplinary teams has made some progress, but especially the impact of modern information systems on the ability of large, dispersed groups of diverse individuals to make fast, high-risk decisions requires more attention.

Leadership/followership. One of the important aspects of traditional CRM training has been to provide guidance about how to play one's role in the overall flight. One intervention that has been especially effective in this regard is pre-flight briefings. By taking advantage of the relatively quiet time before the flight, crew members can discuss likely problems, their preferences and expectations, and even plan for emergencies. This type of planning might be equally effective for "both sides of the locked door." Pilots should continue this practice, and add to it consideration of their response to cabin emergencies. Cabin crew might also benefit from a similar briefing, especially if there is a current state of alert or if a difficult passenger has been identified.

Furthermore, the locked cockpit door may require even more highly developed skills for cabin crews in the area of appropriate assertiveness (Jentsch & Smith-Jentsch, 2001). In fact, there is now a need to conduct research to study the effects of the changes in the cockpit-cabin interface on crew interactions and leadership roles.

Adaptability. Adaptability refers to that set of behaviors that allows aircrews to realize that they cannot follow their original plan. They must use their resources to create a new plan that is responsive to the change in their environment. This set of behaviors is likely to grow increasingly important as a variety of threats confront aircrews. Until recently, it was presumed that most flight disruptions were done with the purpose of attracting attention. It was further assumed that the perpetrator of such actions was motivated to remain alive. However, we now know that terrorists may have the goal of destroying the aircraft or even using it as a weapon. Thus, there is a need for crew members to do their best to discern the intentions of attackers and react accordingly. Further, however, there is also a need to continue to collect information and alter the plan if need be.

Performance monitoring and feedback. Finally, there is a need to develop an environment where all members of the flight community are encouraged to observe one another and to give feedback about how best to improve the safety of the flight. We often tend to compartmentalize security issues as relevant to one only one part of the aviation operation. However, recent events have demonstrated that terrorists will exploit even subtle weaknesses in the system and that we might all notice these shortcomings and/or suggest ways to improve them. Further, all members of the flight operation should receive feedback about their behaviors. This is especially important in those cases where the employee's behavior might have contributed to a weakness in the system.

Training CRM in the New Environment

The need for revised training has not gone unnoticed. These issues are discussed in the recent Aviation and Transportation Security Act, the Homeland Security Act, and in FAA guidance. However, there is little guidance about what, specifically, should be trained or how best to train it. In the following paragraphs, we will provide some suggestions about how to modify CRM training based on the recent scientific literature as well as on practical experience.

Declarative knowledge. Historically, CRM programs have not had to impart much new knowledge; any new knowledge was related to team coordination behaviors, shared mental models of teamwork, etc. (Cannon-Bowers & Salas, 1997). However, this is likely to change in response to the new threats. Crew members need to know the backgrounds and motivations of the most likely attacking groups in order to select the best course of action. Further, they need to be especially aware of subtle behaviors that might help them distinguish between similar-appearing groups. This information is required to enable the adaptability competency required above.

Above and beyond knowing about potential attackers, it is important that crew members also understand their reactions to an attack. These possible reactions place the crew member at risk for poor decision making. For example, aircrews should probably understand phenomena such as the "Stockholm Syndrome," which might make them feel sympathetic towards their captors (Fuselier, 1999). Worldwide, psychologists, in particular, have extensively studied this phenomenon (e.g., Auerbach, Kiesler, Strentz, & Schmidt, 1994; Cremniter, Crocq, Louville, & Batista, 1997; Favaro, Degortes, Colombo, & Santonastaso, 2000; Slatkin, 1998). Consequently, they should be at the forefront of assisting flight crew members by transitioning research results into usable training products.

Finally, it is important that crews are familiar with the host of weapons that might be used against them, the appropriate reactions, and the implications for the aircraft. It is especially important that crew members understand the symptoms of a chemical or biological attack so that they can respond as quickly as possible.

Attitude change. Early CRM programs emphasized changing attitudes as a way of improving teamwork (Helmreich, 1997). The data suggested that this emphasis was useful in helping pilots to improve their team performance (Helmreich, Wiener, & Kani, 1993). In the case of improving security, there is a much larger team to be considered. There is every reason to think that this intervention will be equally effective. The issues related to safety share some similarity to the teamwork issues originally targeted by CRM training. Most people believe that they know all of the safety behaviors and therefore may not be motivated to learn. An attitude intervention might be effective in improving motivation for training. Further, these behaviors are prone to complacency. Again, improving attitudes might be helpful in this regard.

Behaviors. As described above, the new environment imposes a demand for a new behavioral repertoire. Our experience has been that a critical element in acquiring these behaviors is practice and feedback. We suggest that these training opportunities be created to reinforce the behavioral element of training. Typically, this is done through simulator training. A typical complaint about simulator training is that it is expensive. However, we would point out that many of these skills do not need extremely high fidelity in order to be trained effectively. In fact, several researchers have reported considerable success using low-fidelity simulations. Security training would appear to be a good candidate for this approach.

An especially important element of simulator training is the scenario. It is important that the scenario be realistic, target the specific behaviors of interest, and satisfy the various practical elements of training (e.g., length, simulator database, etc.). We developed a product to create scenarios for coordination training (the Rapidly Reconfigurable Line Oriented Events software; Jentsch, Bowers, Berry, Dougherty, 2001; Jentsch, Bowers, Berry, Dougherty, & Hitt, 2001). In our work with one airline, in particular, we developed four different, security-related scenarios so that cockpit and cabin crews could practice their responses to different security threats in well-developed and documented scenarios. A similar application could be easily developed for security training for other aviation personnel.

Finally, it is important that security behaviors become proceduralized. As has been shown by Degani and Wiener (1997), as well as our colleagues at George Mason University (Seamster, Boehm-Davis, Holt, & Schultz, 1998), proceduralizing CRM behaviors can significantly improve the consistency and frequency with which newly trained CRM behaviors are applied. Thus, transforming security knowledge, skills, aptitudes, and behaviors into a set of simple, yet effective procedures is another step that promises large payoffs for aviation security.

Summary

The aviation environment following the attacks of 9/11 has changed dramatically. Current training approaches are no longer sufficient to operate in this environment. We no longer can assume that crew members possess adequate knowledge to deal with the events in flight. Rather, we must be proactive in identifying deficits in crew member knowledge and developing revised training regimens to respond accordingly. We have identified some of these areas above.

However, a change in the training system is not likely without government and industry intervention. We hope that this paper has called attention to these potential training needs in order to start a dialogue that will bring these issues to light.

References

AbuKhalil, A. (2002). *Bin Laden, Islam, and America's new "war on terrorism."* New York: Seven Stories Press.

airWise. (2000, July 07). *World wide air rage campaign*. Retrieved on 10 March 2003 from http://news.airwise.com/stories/2000/07/962970550.html.

airWise. (2001, July 30). *Air rage incidents in South Africa increase in number and violence*. Retrieved on 10 March 2003 from http://news.airwise.com/stories/2001/07/996496230.html.

Anonymous. (1998). FR-*CBH First Responder Chem-Bio Handbook: Practical manual for first responders*. Alexandria, VA: Tempest Publishing.

Anonymous, & Thomas, A. R. (2001). Air *rage: Crisis in the skies*. Amherst, N Y: Prometheus Books.

Auerbach, S. M., Kiesler, D. J., Strentz, T., & Schmidt, J. A. (1994). Interpersonal impacts and adjustments to the stress of simulated captivity: An empirical test of the Stockholm Syndrome. *Journal of Social & Clinical Psychology, 13*(2), 207-221.

Baldwin, R. (1998). Developing *the future aviation system*. Aldershot, England: Ashgate.

Beeks, D. W. (2000). *A study to determine the specific contributors to U.S. airline passenger air rage*. Retrieved from http://www.skyrage.org/PDF/ERAU/dwbeeks.pdf

Berkley, B. J. & Mohammad, A. (2001). Identifying and controlling threatening airline passengers. *Cornell Hotel and Restaurant Information Quarterly*. Cornell University.

Bor, R., Russell, M., Parker, J., & Papadopoulos, L. (2001, September). Handling disruptive passengers. *Air Line Pilot*, 17-20.

Brodie, T. G. (1996). Bombs *and bombings: A handbook to detection, disposal and investigation for police and fire departments* (2nd ed.). Springfield, IL: Charles C Thomas.

Brown, A. (Ed.). (1986). Undercover *fighters: The British 22nd SAS regiment*. New York: Villard Books.

Cannon-Bowers, J. A., & Salas, E. (1997). Teamwork competencies: The interaction of team member knowledge, skills, and attitudes. In H. F. O'Neil, Jr. (Ed.), *Workforce readiness: Competencies and assessment* (pp. 151-174). Mahwah, NJ: Lawrence Erlbaum Associates.

Choi, J.-T. (1994). Aviation *terrorism: Historical survey, perspectives, and responses*. New York: St. Martin's Press.

Chute, R. D., & Wiener, E. L. (1996). Cockpit-cabin communication: II. Shall we tell the pilots? *International Journal of Aviation Psychology, 6*(3), 211-231.

Cremniter, D., Crocq, L, Louville, P., & Batista, X. (1997). Posttraumatic reactions of hostages after an aircraft hijacking. *Journal of Nervous & Mental Disease, 185*(5), 344.

Crenshaw, M. (1988). The subjective reality of the terrorist: Ideological and psychological factors in terrorism. In R. O. Slater & M. Stohl (Eds.), *Current perspectives on international terrorism* (pp. 12-46). New York: St. Martin's Press.

De Landa, M. (1991). *War in the age of intelligent machines*. New York, NY: Zone Books.

Degani, A., & Wiener, E. L. (1997). Procedures in complex systems: The airline cockpit. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, 27*(3), 302-312.

Dickson, P. (1976). *The electronic battlefield*. Bloomington: Indiana University Press.

Favaro, A., Degortes, D., Colombo, G, & Santonastaso, P. (2000). The effects of trauma among kidnap victims in Sardinia, Italy. *Psychological Medicine, 30*(4), 975-980.

Federal Aviation Administration (FAA). (2002). *FAA Strategic Plan Update – January 2002*. Retrieved on 10 March 2003 from http://www.mop-shop.org/plans/FAA_02_SP.pdf.

Frist, B. (2002). When *every moment counts: What you need to know about bioterrorism from the Senate's only doctor*. Lanham, MD: Rowman & Littlefield.

Fuselier, G. D. (1999). Placing the Stockholm syndrome in perspective. *FBI Law Enforcement Bulletin, 68*(7), 22-25.

General Accounting Office. (1997). *FAA's guidance and oversight of pilot crew resource management training can be improved* (Report No. GAO/RCED-98-7). Washington, DC: Government Printing Office.

Gero, D. (1997). *Flights of terror – Aerial hijack and sabotage since 1930*. Sparkford, UK: Patrick Stephens Limited.

Gordon, D. (1981). *Electronic warfare*. New York: Pergamon Press.

Halberstadt, H. (1994). *SWAT TEAM – Police special weapons and tactics*. Osceola, WI: MBI.

Helmreich, R. (1997, May). Managing human error in aviation. *Scientific American*, 62-67.

Helmreich, R.L., Wiener, E., & Kanki, B.G. (1993). The future of Crew Resource Management in the cockpit and elsewhere. In E. Wiener, B. Kanki, & R. Helmreich (Eds.), *Cockpit Resource Management* (pp. 479-501). San Diego, CA: Academic Press.

Helmreich, R. L., Wilhelm, J. A., Klinect, J. R., & Merritt, A. C. (2001). Culture, error and Crew Resource Management. In E. Salas, C.A. Bowers, & E.

Edens (Eds.), *Applying resource management in organizations: A guide for professionals (pp.305-331).* Hillsdale, NJ: Erlbaum.

Huang, J. (2001, September). Legal issues related to unruly airline passengers. *Air Line Pilot*, 21-23.

Hubbard, D. G. (1973). *The Skyjacker – His flights of fantasy*. New York: Collier.

Hudson, R. (2002). *Who becomes a terrorist and why – The 1999 government report on profiling terrorists* (by Rex Hudson and the staff of the Federal Research Division of the Library of Congress). Guilford, CT: Lyons Press.

Hufnagl, W. (1999). *Spezialeinheiten der österreichischen Polizei und Gendarmerie* [Special units of the Austrian police and gendarmes; in German]. Stuttgart, Germany: Motorbuch.

James, L. & Nahl, D. (n.d.). *Air rage prevention*. Retrieved on 10 December 2002 from http://www.aloha.net/~dyc/rages/index.htm.

Jentsch, F., Bowers, C., Berry, D., & Dougherty, W. (2001). Improving the generation of line-oriented scenarios through computerized tools. *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus: The Ohio State University.

Jentsch, F., Bowers, C., Berry, D., Dougherty, W. & Hitt, J. M., II. (2001). Generating line-oriented flight simulation scenarios with the RRLOE computerized tool set. *Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

Jentsch, F., & Smith-Jentsch, K. (2001). Assertiveness in teams: More than "just say no." In E. Salas, C. Bowers, & E. Edens, (Eds.), *Resource Management in Organizations* (pp. 73-94). Mahwah, NJ: Lawrence Erlbaum Associates.

Katz, S. M. (1997). *LAPD – Patrol, SWAT, Air Support, K-9, CRASH, and Homicide*. Osceola, WI: MBI.

Lenz, R. R. (1976). *Explosives and bomb disposal guide*. Springfield, IL: Charles C Thomas.

Michelletti, E. (1999). *Spezialeinheit GIGN - Frankreichs Elitegendarmerie im Einsatz* [Special unit GIGN - France's elite gendarmes in action, in German; French original: *Le GIGN en action*]. Stuttgart, Germany: Motorbuch.

National Research Council. (1996). *Airline passenger security screening – New technologies and implementation issues* (Publication NMAB-482-1). Washington: National Academy Press.

Osterholm, M. T. & Schwartz, J. (2000). *Living terrors: What America needs to know to survive the coming bioterrorist catastrophe*. New York: Delacort/ Random House.

Pickett, M. (1998). *Explosives identification guide*. New York: Delmar.

Reiley, M. (2002, July). *National hotline; Teamsters Local 2000 Flight Attendant*

Union. Retrieved on January 22, 2003 from http://www.local2000.org/hotlines/hotline020719.htm.

Salas, E., Burke, S. C., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors, 43*(4), 641-674.

Scholzen, R. (2000). *SEK - Spezialeinsatzkommandos der deutschen Polizei* [SEK – Special operations units of the German police; in German]. Stuttgart, Germany: Motorbuch.

Scholzen, R. & Froese, K. (1999). *GSG 9* [in German]. Stuttgart, Germany: Motorbuch.

Seamster,T. L., Boehm-Davis, D. A., Holt, R. W., & Schultz, K. (1998). *Developing Advanced Crew Resource Management (ACRM) training: A training manual*. Washington, DC: FAA AAR-100.

Sheffer, M. P. (n.d.). *The problem passenger: A history of airline disruption: 1947 to present*. Retrieved from http://www.skyrage.org/PDF/SKYRAGE/scsi.pdf

Slatkin, A. A. (1998). The Stockholm Syndrome and situational factors related to its development. *Dissertation Abstracts International: Section B: The Sciences & Engineering Univ. Microfilms International, 58*(7-B), 3970.

Snow, R. L. (1996). *SWAT teams – explosive face-offs with America's deadliest criminals*. Cambridge, MA: Perseus.

Thompson, L. (1986). *The Rescuers – The world's top anti-terrorist units*. New York: Dell.

Tomajczyk, S. F. (1997). *U. S. Elite counter-terrorist forces*. Osceola, WI: MBI - Motor Books International.

Wallis, R. (1993). *Combating air terrorism*. Washington: Brassey's.

Wilkinson, P. & Jenkins, B. M. (Eds.). (1999). *Aviation terrorism and security*. Portland, OR: Frank Cass/ISBS.

Appendix

Description of a Sample Training Program for CRM in the New Environment

We developed a resource document as part of a program of research to support the FAA's response to the new environment in the U.S. National Airspace System (NAS). We focused on three research thrusts in this area, each centering on one aspect of improved CRM training for flight crews. The approach combined findings from modern behavioral and cognitive sciences, specifically from human factors, with state-of-the-art technological solutions. The resulting training materials are intended to give trainees the necessary knowledge to detect, assess, report, and counteract new threats to the safety and security of flight; it provides them with information and strategies for team communication, coordination, and cooperation; and it provides them with practice opportunities to apply the knowledge and strategies.

We have collected and analyzed information on the new threats in the aviation environment after September 11, 2001, and have studied their effects on CRM. Additionally, we have created event-based scenarios for this training. This task focuses on the training of pilots in the procedures and maneuvers that allow them to respond most effectively to new challenges in the aviation environment.

In accordance with these goals, the document contains resource material in the following areas:

- Air rage
- Psychology of aircraft hijackers
- Motivations of terrorist and criminal organizations
- Psychology of interactions between hostages and their kidnappers
- Goals and techniques of hostage negotiators
- Simulation scenarios to practice CRM in the new environment

In addition to chapters on these critical areas, the document also contains a bibliography, a glossary or related terms, and additional resource material from various open sources to support the chapters.

Training Modules

Ten training modules (in the form of instructional goals and supporting training media, mostly PowerPoint slides) were developed to support security training for flight crews. In the following sections, we briefly describe the modules and their contents.

Module I provides the overview of the training tool set and gives introductory examples for each of the modules.

Module I-A gives the history of air terrorism throughout the world. Two case studies provide insight into the evolution of terrorism over time.

Module II thoroughly covers the area of explosive devices. This includes the types of explosive devices that are on the market as well as hand-made devices and the different types of triggers.

Module III provides detail on biological and chemical threat. First, this module covers the components of chemical threats and the effects on the environment and population. Second, biological threats are explained, categorized, and the related symptoms are discussed. Last, the effects of a nuclear weapon threat are covered.

Module IV discusses the types of electronic threat to civil aviation and countermeasures to resist or avoid this type of risk.

Module V presents information regarding five general types of skyjackers (which fall into three categories, i.e., criminal, terrorist, or suicidal) and discusses ways to react if confronted with a skyjacking situation.

Module VI covers three types of terrorist groups as well as their motivations. First, the rational terrorists think through decisions to create the most advantageous situation for themselves. Second, the psychologically motivated terrorist uses terrorism to maintain their identity and self-esteem. Last, the culturally motivated terrorist perceives a threat against their culture, religion, or some type of group. The module delves into the various international terrorist groups in detail and gives examples of domestic terrorist groups.

Module VII presents detailed information on the psychology of hostage situations. First, information is given regarding the use of intimidation. Second, the phenomenon of Stockholm Syndrome is addressed as well as the opposite effect of the London Syndrome. Finally, the skill of hostage negotiation is covered including many recommendations and cues for how to react in certain situations.

Module VIII discusses the internal threats relevant to the airline industry. Two case studies are included as well as procedures and tips to prevent workplace violence.

Module IX provides information regarding the issue of air rage including the possible causes, profiles, examples, and the impact of air rage on the airline industry. Various types of prevention strategies are suggested for traveling customers.

Module X presents the susceptibility of aircraft to various dangers. This includes the interior and exterior high-threat areas of caution throughout the aircraft.

Module XI gives examples of incognito terrorists that were successful in completing a terrorist act by dressing and behaving in inconspicuous ways. Strategies are given to prevent this type of terrorist from completing their mission.

Author Notes

Clint Bowers and Florian Jentsch, Department of Psychology, University of Central Florida. Eduardo Salas, Department of Psychology, University of Central Florida, and Institute for Simulation and Training (IST), Orlando, Florida.

Preparation of this article was supported by FAA Grant 99-G-047 to the first and third authors. Eleana Edens, Ph.D., was the FAA contract monitor. All opinions expressed herein are those of the authors and do not necessarily reflect the position of the University of Central Florida, the Federal Aviation Administration, or the U.S. Department of Transportation.

Correspondence regarding this paper should be directed to Clint Bowers, Department of Psychology, University of Central Florida, P. O. Box 161390, Orlando, Florida 32816-1390. Electronic mail can be sent to bowers@mail.ucf.edu.

International Journal of Applied Aviation Studies, 3,1

A Longitudinal Study of the Learning Styles of College Aviation Students

Craig A. Kanske
Skyviews L.L.C.
905 N.E. 20th Street
Moore, OK 73160-6310
Email: skyviews@swbell.net

L. Tom Brewster
Oklahoma State University-Tulsa
517 Ridge Oak Lane
Sapulpa, OK 74066-8934
Email: tombrew@okstate.edu

and

Richard O. Fanjoy
Purdue University
Department of Aviation Technology
Purdue Airport, AT 103
West Lafayette, IN 47907-3398
Email: rofanjoy@tech.purdue.edu

Abstract

The authors have initiated a five-year study to investigate the learning styles of collegiate aviation students and track them by year group (freshman, sophomore, junior, senior and graduate). The study is based on earlier work by Brewster and Kanske (2001) that suggested significant differences between lower and upper division aviation students. This paper reports on data collected at the end of the first year from four hundred twenty (420) students sampled across eight aviation programs. The participating institutions were: Oklahoma State University-Stillwater, Oklahoma State University-Tulsa, Tulsa Community College, Southeastern Oklahoma State

---

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

University on the Durant campus and at Tinker Air Force Base, Rose State College, University of Oklahoma and Purdue University. Kolb's Learning Style Inventory (LSI) was used to analyze the learning styles of the responding students. Findings from first-year data suggested converger or assimilator learning styles predominate within the college aviation student population (67.2 %), comparable to earlier findings with United States Air Force pilot populations (67.8%). In addition, first-year data indicated that freshmen aviation student learning style distribution is similar to that found within the general student population. Such findings seemed to suggest a shift in student learning styles over the course of a 4-year college aviation experience. Data collected over the course of the five-year study may reveal whether such a shift results from learning style modification or self-elimination of students with a non-predominant learning style. The authors believe that study findings will inform faculty and program administrators how best to meet the needs of all enrolled students.

A Longitudinal Study of the Learning Styles of College Aviation Students

A learning style is a student's consistent way of responding to and using stimuli in the context of learning (Clark, 2000). Learning styles are points along a scale that help us to discover the different forms of mental representation. They describe how people learn without judging. A measurement of learning style places the individual along a continuum; it does not define a limitation in learning, but identifies a preference or comfort zone in learning environments (Clark, 2000).

Understanding the predominant learning style is not just an aid in course development. Helping the student understand his or her personal learning style can directly affect their learning. Sharp, an associate professor of technical communications in the chemical engineering department at Vanderbilt University, "found that teaching students about learning styles helps them learn the course material because they become aware of their thinking processes. More importantly, she says, it helps them develop interpersonal skills that are critical to success in any professional career" (Felder, 1996).

Identifying and measuring learning styles has taken many forms. Torrance and Rockenstein (1988) catalogued 23 different learning styles measurable through various survey instruments. Additional research studies focused on learning modes (visual, verbal, etc), left hemisphere/right hemisphere brain bias, field dependence/field independence, and adult learning preferences. The common thread in this field of study is a focus on the individual learner. Recent studies in aviation education have begun to identify individual preferences held by aviation students (Quilty, 1996; Brady, et. al., 2001; Kanske & Brewster, 2001). In this study, we continue the process, looking at preferred learning styles of college aviation students within the context of the experiential learning model.

Experiential learning, or learning from experiences, is described as a cycle that arbitrarily begins with a single learning event. After the event, the learner reflects on the experience, relating it to other events both similar and dissimilar. Next, the learner builds theories that determine actions during future events, and conceptualizes possible outcomes. Finally, the learner experiments with these theories, testing them for accuracy and effectiveness. In this way, each learning event has its basis in past experience and provides a background against which the learner prepares for future events (Kolb, 1984).

Over time, each learner becomes more comfortable with specific parts of the learning process. Schmeck (1988) attributed this learning style development to positive reinforcement during early learning situations. Continued success with a style of learning, and the positive feelings of achievement this success brings, lead to a preference for a specific style even when other styles may be more appropriate for specific subject matter. Identifying the learner's preferred style is the basis of learning style research. Kolb (1984), in developing a learning style model based upon the experiential learning model, described the four basic steps in the learning cycle by using two sets of opposing choices. Abstract conceptualization, or theory building, opposes participation through concrete experience. The reflective observation of the introvert is opposed by the active experimentation of the extrovert. The dichotomy between the abstract and the concrete thinker has entered the popular culture with the images of a concrete thinking, pocket protector wearing, left-brained engineer and the longhaired, abstract-thinking artist. The reflectively observing introvert and the actively experimenting extrovert are also readily identifiable stereotypes. Kolb (1984, 1993) applied the concept of preferred learning styles to these two dichotomies to identify four different styles. It should be noted that the general population distributes evenly across these four styles.

The Accommodator is a concrete-thinking extrovert who combines concrete experience and active experimentation. The accommodative learner is action-oriented and prefers facts to theories. This style of learner is open to new experience but fixates on insignificant tasks.

Almost a mirror image of the Accommodator, the Assimilator is the abstract-thinking introvert combining abstract conceptualization and reflective observation. The assimilative learner is one who builds logical and precise theories about an event from seemingly unrelated facts. Idea and concept focus, as opposed to a people focus, is another trademark of the assimilative learner.

The Diverger is a concrete-thinking introvert combining concrete experience and reflective observation. The Diverger looks at the many sides of an event to develop a "big picture" analysis of what occurred. This is the brainstormer of the group, offering many possible approaches to a problem, actively searching for support and gauging how others in the group feel about these alternatives.

On the opposite side of the scale from the Diverger is the Converger, an abstract-thinking extrovert combining abstract conceptualization and active experimentation. The convergent learner offers practical methods of applying what might be possible, without having to live through the experience first. This is the learner who will sift through alternatives, select what he or she sees as the most technically appropriate, and put into practice that alternative.

As each individual will have all of the learning traits to differing degrees, it is appropriate to return to Schmeck's (1988) statement that dominant learning styles are developed from past successes in using that approach. Positive outcomes and reinforcements from authority figures during learner development tend to produce a predisposition toward a specific style, even when a different style could produce a better result that is within the learner's capability. It is this dominant learning style, which is identified using the multitude of available survey devices.

The Kolb Learning Style Inventory (LSI), was developed in 1976, and revised in 1985 and 1993 to identify where in the learning cycle an individual's preferences fall. To complete the LSI, the subject ranks four possible endings for 12 sentence stems. This forced-ranking, where each ending identifies one of the four steps in the learning cycle, produces a score between 12 and 48 for each mode of learning. Two combination scores are derived to identify a preferred location along each of the two learning dichotomies: abstract/concrete and active/reflective. Subtracting the concrete experience score from the abstract conceptualization score, and subtracting the reflective observation score from the active experimentation score, provides these two combination scores. Plotting these two combination scores on a learning-style grid will identify which quadrant best describes the preferred learning style of the subject (Kolb, 1985, 1993).

Kolb's Learning Style Inventory (LSI) has several advantages for research. In addition to being revised twice since its development in 1976, the LSI has been validated over the years in such studies as a comparison of learning among European management training students (Jackson, 1995) and Willcoxson and Prosser's (1996) validation study in Australia. The LSI also is relatively easy to understand and administer, and with only 12 forced-choice responses to complete, the instrument is quickly answered.

Kanske (1998, 1999) used Kolb's Learning Style Inventory to examine the pilots who had completed undergraduate pilot training with the United States Air Force. When compared to the general population's even distribution, the study group showed a significant ($p < 0.0001$) preference for the Converger learning style. This prompted the question of when this distribution occurs. A follow-up study by Brewster and Kanske (2001) also used Kolb's Learning Style Inventory to examine the learning styles of college aviation students in an effort to determine if there was a change in learning style during the college years. That study found a significant difference ($p < 0.0013$) in learning styles from freshman to graduate student, with an apparent step between freshman and sophomore years. While the observed differences appeared to be significant, the small sample size dictated that caution be taken when trying to describe the specific change at the sophomore year. The current longitudinal study will expand the sample size and hopefully provide further support for previous data. Additionally, it will begin the process of finding out if the difference in learning style shown between class levels is a result of changing learning styles among students, or if students who have dissimilar learning styles leave the program.

Current aviation curricula are designed with a building block approach to prepare students for operations in a professional aviation environment. This methodology does not consider the impact of sequential, pedagogical presentations on the learning style distribution within the aviation student population. Previous studies indicated a significant population of students has a learning style preference that is well suited to the building block approach. The longitudinal portion of this study is intended to determine if individual students within aviation programs change their learning styles to match the material and presentation, or if those students with non-predominant learning styles leave the program. The answer to this question will provide insight to course design decisions within an aviation education environment. The first phase of the longitudinal study (year one) is reported here.

Methodology

The population for this study consisted of students enrolled in the aviation programs at Oklahoma State University-Stillwater campus, Oklahoma State University-Tulsa campus, Tulsa Community College, Rose State College, University of Oklahoma, Southeastern Oklahoma State University on the Durant campus and at Tinker Air Force Base and Purdue University. Kolb's Learning Style Inventory was used as the survey instrument. The demographic survey form was modified to include student identification information, in order to track students through subsequent surveys.

Surveys were distributed to students during the Spring 2002 semester. Survey packages, including a cover letter, the demographic form, and the Learning Style Inventory were distributed and collected by classroom instructors at each location.

Results and Analysis

Responses were received from 88 students at Oklahoma State University-Stillwater, 42 students at Oklahoma State University-Tulsa, 30 students at Tulsa Community College, and 45 students at the Southeastern Oklahoma State University Durant and Tinker Air Force Base locations. A further 28 students from Rose State College responded, as did 47 students from Oklahoma University and 140 students from Purdue University. The grade level breakdown of this group of students is shown in Table 1. Students from Oklahoma State University-Tulsa were enrolled in a cooperative education program with Tulsa Community College. Southeastern Oklahoma State University-Tinker students primarily were enrolled in upper division courses. Lower division courses for Tinker Air Force students were obtained from either a local junior college or schools previously attended and were transferred into the Southeastern Oklahoma State University program. Some responses were unusable due to errors such as no demographic data, missing responses on the Learning Style Inventory, and responses on the Learning Style Inventory, which violated scoring criteria. Only useable survey responses are included in Table 1.

Table 1

Responses by Grade Level and School

School	**OU**	**Purdue**	**Rose State**	**OSU - Tulsa**	**OSU - Stillwater**	**TCC**	**Southeastern**	**Total**
Grade Level								
Freshman	9	53	10	0	30	8	7	117
Sophomore	12	16	5	1	22	17	15	88
Junior	11	56	2	9	22	3	15	118
Senior	8	15	1	22	12	0	8	66
Grad Student	6	0	9	10	2	2	0	29
Non-credit	1	0	1	0	0	0	0	2
Total	47	140	28	42	88	30	45	420

Instruments were scored using the methods outlined in the Learning-Style Inventory Self-Scoring Inventory and Interpretation Booklet (Kolb, 1993). This method produced scores for concrete experience, reflective observation, abstract conceptualization, and active experimentation for each subject. From these four raw scores, two combination scores, abstract conceptualization minus concrete experience (AC-CE) and active experimentation minus reflective observation (AE-RO) were derived. The final step in the analysis was to plot the intersection of the two combination scores on a grid using AE-RO as the X-axis and AC-CE as the Y-axis. The quadrant on the grid in which the intersection falls was used to define the subjects learning style.

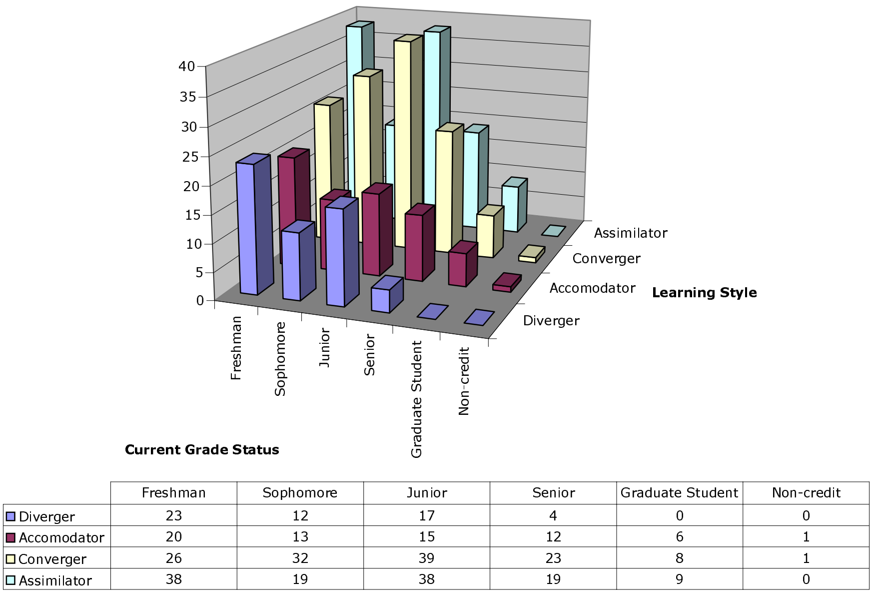
The predominant learning style displayed by 129 college students, using this scoring method, was Converger. This represents 34.4 percent of the surveys with valid learning styles. The Diverger learning style was least represented among the responding students with only 56 students, 14.9 percent. Assimilator style was the second most prominent with 123 students or 32.8

percent, followed by the Accommodator style with 67 students and 17.9 percent. These results are shown in Table 2 and Figure 1.

Table 2

Style by Current Standing (frequency)

Grade Level	**Freshman**	**Sophomore**	**Junior**	**Senior**	**Graduate Student**	**Non-credit**	**Total**
Style							
Accommodator	20	13	15	12	6	1	67
Assimilator	38	19	38	19	9	0	123
Converger	26	32	39	23	8	1	129
Diverger	23	12	17	4	0	0	56
Total	107	76	109	58	23	2	375

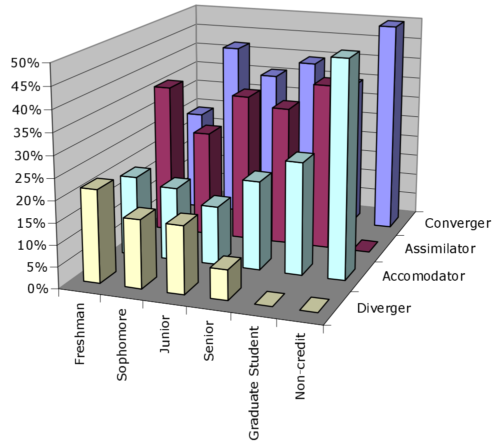


	Freshman	Sophomore	Junior	Senior	Graduate Student	Non-credit
🟪 Diverger	23	12	17	4	0	0
🟥 Accomodator	20	13	15	12	6	1
🟨 Converger	26	32	39	23	8	1
🟦 Assimilator	38	19	38	19	9	0

Figure 1. Learning Style by Current Grade Status.

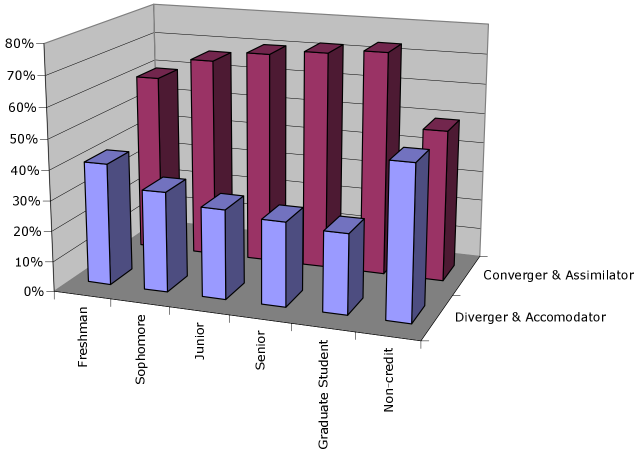
Figures 2 and 3 summarize the learning styles of the study group broken down by grade level. Among the freshman class, 21% percent were Divergers, 19% percent Accommodators, 24% percent Convergers, and 36% percent Assimilators. Convergers and Assimilators represented 67% percent of the sophomore class, and the remaining 33% percent were Accommodators or Divergers.

The junior class had 71% percent Convergers/Assimilators and 29% percent Accommodators/Divergers, while the senior class had 72% percent Convergers/Assimilators and 28% percent Accommodators/Divergers. Finally, 74% percent of graduate students were Convergers/Assimilators and only 26% percent were Accommodators/Divergers.



	Freshman	Sophomore	Junior	Senior	Graduate Student	Non-credit
■ Diverger	21%	16%	16%	7%	0%	0%
■ Accomodator	19%	17%	14%	21%	26%	50%
■ Assimilator	36%	25%	35%	33%	39%	0%
■ Converger	24%	42%	36%	40%	35%	50%

Figure 2. Style by Class (Percentage).



	Freshman	Sophomore	Junior	Senior	Graduate Student	Non-credit
■ Diverger & Accomodator	40%	33%	29%	28%	26%	50%
■ Converger & Assimilator	60%	67%	71%	72%	74%	50%

Figure 3. Style Groupings by Class (Percentage).

The results for group learning styles, derived with the calculations of AE-RO and AC-CE used to define learning styles, are shown in Table 3. Means for the total sample of AE-RO equal 5.20 and AC-CE equal 7.50 plots in the Assimilator style. Freshmen, with an AE-RO of 4.07 and an AC-CE of 6.01 also plot in the Assimilator style. The Sophomores AE-RO of 6.66 and AC-CE of 6.71 plots in the Converger style. An AE-RO of 4.61 and AC-CE of 8.07 for juniors results in the Assimilator style. Means of 6.08 for AE-RO and 9.40 for AC-CE falls on the vertical axis (AC-CE) between the Assimilator and Converger styles for seniors. Finally, Graduate students also plot very near the AC-CE axis in the Assimilator/Converger styles with means for AE-RO of 5.87 and AC-CE of 12.13.

Table 3

Mean Values for Active Experimentation minus Reflective Observation (AE-RO) and Abstract Conceptualization minus Concrete Experience (AC-CE)

Grade Level **Style**	**Freshman**	**Sophomore**	**Junior**	**Senior**	**Graduate Student**	**Non-Credit**	**Total**
Active Experimentation minus Reflective Observation (AE-RO)	4.07	6.66	4.61	6.08	5.87	8.00	5.20
Abstract Conceptualization minus Concrete Experience (AC-CE)	6.01	6.71	8.07	9.40	12.13	1.00	7.50

Conclusions

Findings from this data set were similar to earlier findings with both United States Air Force pilots and college aviation students (Kanske & Brewster, 2001). Perhaps most striking was the similarity in the percentages of respondents with either the Assimilator or Converger learning styles. These two learning styles made up 67.8 percent of the U.S. Air Force study group (Kanske, 1998, 1999), 61.5 percent of the college study group and 67.2% percent of the current study group.

The Kolb learning style inventory, for a random population, will produce an equal distribution among the four learning styles. The total sample of this study showed a significant deviation ($p<0.0001$) from equal distribution with a tendency toward abstract-conceptualization. The distribution of freshman learning styles matched that of a random population ($p=.073$). At the sophomore level, a strong distribution ($p<0.004$) toward Assimilator and Converger was observed. Junior level student responses skewed toward Converger and Assimilator ($p<0.0004$). Converger was the dominant style among seniors at 38% ($p<0.004$) and graduate students were Assimilator at 39% ($p<0.0232$).

The study of college aviation students by Kanske and Brewster (2001) showed an initial random distribution of learning styles, with a significant shift toward the Assimilator or Converger style by the sophomore year. Martin (2000) found a shift in style after the sophomore year, and suggested that this shift deserved further study. Kanske and Brewster (2001) found the shift to occur even earlier, after the freshman year. The continuing longitudinal data collection should further validate the current findings. In addition to highlighting learning style trends for individual students, this data should provide insights to learning style aspects of the entire aviation student population.

Certain styles of learning appear to be dominant throughout the aviation experience. All United States Air Force pilots are required to have a four-year college degree, and closely match the characteristics of graduate students sampled in this study. It should be noted that the United States Air Force study (Kanske, 1998, 1999) percentage 67.8% closely matched the 74% result for graduate students from this study.

By tracking the learning styles of aviation students, the authors hope to determine if changes of individual learning styles occur during the college experience, or if individuals with "non-predominant learning styles" tend to self-eliminate from aviation programs. Identification of a predominant aviation student learning style also can be used as an aid in course instructional design.

Demographic data may provide additional insights to student learning styles. While the data from the current study represented a relatively large sample, it was difficult to generalize the results due to limited demographic information. The study surveyed eight institutions with varied program size, location, and student profiles. There was insufficient data, however, to determine if this data can be considered a representative sample of any group other than the eight institutions involved. General demographic information about collegiate aviation programs would provide a statistical background against which to analyze these study results. Over the course of the longitudinal study, efforts are being made to obtain demographic information from participating institutions. Additional demographic information from the population of aviation programs will allow generalization of results across the entire aviation student body. The eight sampled institutions were chosen based on access. A random sample chosen from all aviation programs based upon the general demographics would produce significantly better data.

References

Brady, T., Stolzer, A. J., Muller, B., & Schaum, D. (2001) A comparison of the learning styles of aviation and non-aviation students. *Journal of Aviation/ Aerospace Education & Research (II)* 1. 33-41.

Clark, D. (2000, November 5). *Learning Styles. Or how we go from unknown to known.* Retrieved January 2003 from http://www.nwlink.com/~donclark/hrd/learning/styles.html

Felder, R.M. (1996). *Matters of Style.* Retrieved January 2003 from http://www2.ncsu.edu/unity/lockers/users/f/felder/public/Papers/LS-Prism.htm

JMP-IN Edition 3.1.5 [computer software]. (1995). Cary, NC: SAS Institute, Inc.

Kanske, C. A. (1998/1999). The learning styles of pilots currently qualified in United States Air Force aircraft (Doctoral dissertation, Oklahoma State University, Aviation and Space Program 1998). *Dissertation Abstracts International*, 9918799.

Kanske, C. A., & Brewster, L. T. (2001). The learning styles of college aviation students. *Collegiate Aviation Review* (19), 1. 62-70.

Kolb, D. A. (1985). *Learning-Style Inventory: Self-Scoring Inventory and Interpretation Booklet*. Boston: Hay/McBer Training Resources Group.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, N.J.: Prentice-Hall.

Kolb, D. A. (1993). *Learning-Style Inventory: Self-Scoring Inventory and Interpretation Booklet (Revised scoring)*. Boston: Hay/McBer Training Resources Group.

Martin, R. A. (2000). The relationship of document and quantitative literacy with learning styles and selected personal variables for aviation university students. *Collegiate Aviation Review* (18), 1. 26-40.

Quilty, S. M. (1996). Cognitive learning bias of college students in an aviation program. *Journal of Air Transportation World Wide*

Schmeck, R. R. (Ed.). (1988). *Learning strategies and learning styles.* New York, N.Y.: Plenum Press.

Statistica Release 5.1 [computer software]. (1997). Tulsa, OK: StatSoft, Inc.

Torrance, E. P. & Rockenstein, Z. L. (1988) Styles of thinking and creativity. In Schmeck, R. R. (Ed.), *Learning Styles and Learning Strategies* (pp 275-290). New York, NY: Plenum Press.

International Journal of Applied Aviation Studies, 3,1
Copyright © 2003, FAA Academy, Oklahoma City, OK

Developing An Aviation Safety Culture: Utilizing Databases To Promote Accident/Incident Prevention Programs

Brian G. Dillman
Purdue University
Department of Aviation Technology
Hangar 6, 128
West Lafayette, IN 47907-3398
Email: dillman@purdue.edu

Jang R. Lee
Purdue University
Department of Aviation Technology
1 Purdue Airport
West Lafayette, IN 47907-3398

and

Don Petrin
Purdue University
Department of Aviation Technology
Purdue Airport, Hanger 5
West Lafayette, IN 47907
Email: dapetrin@tech.purdue.edu

Abstract

Developing a safety culture in collegiate flight training is often a critical step in aviation safety that must be approached from several different directions. A key element in the development of a safety culture is the implementation of an accident/incident prevention program. To aid in the development of such a program requires numerous information sources with maintenance data-

Requests for reprints should be sent to Beverly Laughead, FAA Academy,AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

> *bases being just one such source. Maintenance databases can be utilized not only in the necessary analysis after an accident or incident, but also prior to such an occurrence.*

Safety Culture

Before developing programs to create a safety culture, the concepts of "culture" and "safety culture" first must be defined. Since the actual measurement of culture is problematic, it is imperative that these guidelines be established. Kluckhohn and Strodbeck (1961) identified culture as a mega-variable from which all contingencies originate.

> We think the way we have been socialized to think. The universal value orientations are to be found in all cultures. Each of the variations has implications for values, behavior, and emotions. Kluckhohn and Stodtbeck's five orientations are: (1) human nature, (2) relationships between man and nature, (3) time, (4) activity, and (5) relationships.

Each variable is interwoven such that one cannot be changed without the others being affected as well as the culture as a whole. Kornblum identified culture as a total system of values, beliefs, attitudes, traditions, and behavior norms regulating life among a particular group of people (Kornblum, 1991). It is the goal of safety individuals to mold the characteristics of aviation professionals so they become part of the culture and the culture becomes a part of them. As Reason (2000) pointed out, culture is something that an organization "is" rather than something it "has."

Now that the overall concept of "culture" has been discussed, how does the organizational culture impact the establishment of a "safety culture"? One of the steps in establishing a safety culture is to ensure that individuals are actually aware or informed of the need for a safety culture. "Informed culture" equates with the term "Safety Culture" – one in which those who manage and operate the system have current knowledge about the human, technical, organizational, and environmental factors that determine the safety of the system as a whole (Reason, 2000). Simply realizing that a safety culture is something of importance is a step in the correct direction. One of the ways that an awareness of a safety culture can be promoted is by placing the idea of safety at the forefront from the beginning of training all the way through the certification process. Safety should be taught as a continuous process beginning in primary schools with the objective of creating a society which understands the basic tenets of risk and which encourage individuals to form opinions based on real knowledge of hazards (Thorburn, 1990). Furthermore, the teaching at the primary stage should relate to social behaviors and to safeguarding health by taking precautionary and avoidance measures (Thorburn, 1990). The establishment of accident prevention data is paramount in the effort of precautionary and avoidance measures. Analyzing and preparing for situations that have a reasonable probability of occurrence can significantly increase the likelihood of breaking the error chain.

Accident & Incident Databases

Now that the concept of a safety culture has been discussed, what is the ideal way of proceeding to establish such a mindset? Without first identifying the problem, there is no possible way to proceed. Accurate and permanent records of observed behavior are essential for a job safety analysis (Geller, 2001). Concrete measurements and detailed observations are required to determine where there are weaknesses in the safety culture so that appropriate remedies can be devised.

> In the absence of bad outcomes, the best way – perhaps the only way – to sustain a state of intelligent and respectful wariness is to gather the right kind of data. This means creating a safety information system that collects, analyses, and disseminates information from incidents and near-misses as well as from regular proactive checks on the system's vital signs (Reason, 2000).

It is imperative that databases be created such that organizations and individuals can reap the rewards of past occurrences. Aviation professionals need to have the skill set required to learn from past mistakes and prevent them from happening again.

Accident databases are not unique to the aviation industry. Numerous institutions see the validity in identifying and tracking accidents and incidents and the key elements that lead to their occurrence. The sports industry has determined that an accident data system will lead to the systematic collection of valid and reliable data and material evidence which, if utilized properly, can eliminate most potential accident hazards and minimize the consequences of mishaps which do occur (Damron, 1977). Elementary education professionals have determined that to prevent bus accidents involving students and motor vehicles, an accident database must be developed to increase safety. If we cannot compile basic accident data which is necessary for basic accident analysis, we cannot develop basic safety programs which are the lifeblood of all accident prevention programs (Comeau, 1986). Summer camp organizers realize that an organizational culture that supports the reporting of near miss incidents is an important foundation to successful risk management (Leemon, Schimelpfenig, Gray,Tarter, & Williamson, 1998). Each of these institutions have concluded that without the collection and analysis of accident data, events are destined to repeat themselves until preventive steps are taken. Even with near-miss situations, it is imperative that proper analysis is completed to reduce the total number so that an accident will not develop. For every 600 near hits, there will be 30 property damage incidents, 10 minor injuries, and 1 major injury (Geller, 2001). If we simply focus on the major accidents, there is a multitude of information that is being overlooked. Information about near hits and other types of incidents must be analyzed for future reference and trend analysis.

Besides preventing future accidents and incidents from occurring, databases have additional uses for the safety professional as outlined by summer camp organizers. According to Lemmon et al. (1998), accident databases:

- Provide a tool to educate interested parties about the risk management.
- Provide a significant and comprehensive incident database upon which statistical analysis can be done to provide a collective knowledgebase and reliable information source.
- Provide a central reporting forum for incidents.
- Enhance the collective judgment by examining trends within it.

However, one of the most valuable aspects to gathering incident data is it provides the ability to share the narratives of each incident with staff and colleagues (Leemon et al., 1998). The sharing of narratives places a real world backdrop to a situation, which is often removed to the point that it only happens to "other people." This mentality is exactly what accident and incident databases can help to prevent.

Safety Programs

The process of creating a safety program and incorporating an accident and incident database is no small endeavor. Because of the interwoven properties of a safety culture, it is imperative that careful consideration be placed on the entire process instead of creating a program simply for the sake of creating a program.

A recent review of a number of safety process measures identified five broad clusters in the creation of a safety program (Reason, 2000).

- Safety-specific factors (for example, incident and accident reporting, safety policy, emergency resources and procedures, off-the-job safety and so on)
- Management factors (for example, management of change, leadership and administration, communication, hiring and placement, purchasing controls, incompatibilities between production and protection and so on)
- Technical factors (for example, maintenance management, levels of automation, human-system interfaces, engineering controls, design, hardware and so on)
- Procedural factors (for example, standards, rules, administrative controls, operating procedures and so on)
- Training (for example, formal versus informal methods, presence of a

training department, skills and competencies required to perform tasks and so on)

A thorough and accurate maintenance database can facilitate the incorporation of safety-specific management and technical factors. Specifically, Leemon et al. (1998) stated that a formula for accident potential can be boiled down to environmental hazards (aircraft) + human hazards (pilots) = accident potential (incident databases). A maintenance database encourages open communication between all three areas specifically relating to the status of the aircraft. By actively engaging all areas of the safety equation, it is possible to improve the overall safety culture. As identified by researchers involved in analyzing pedestrian safety, students learn to practice safe behavior most effectively when they are involved in the process of identifying and controlling hazards (Wojtowicz & DesLauriers, 1995). By placing the student in the center of the safety culture and actively engaging them in the safety process, they become more aware of the requirement for safety.

A maintenance database also will have a profound effect on the relationship between maintenance technicians and pilots. By creating a strong link of communication between environmental hazards (aircraft) and human hazards (pilots) accident potential (incident databases) can be significantly reduced.

Historically, the communication link between these two parties has left something to be desired. Young, Mattson, and Petrin (2001) discovered an area of concern in their research on the communication between pilots and mechanics that showed 95% of pilots felt mechanics were their colleagues while 69% of mechanics felt the same. When respondents were asked to identify problem areas in aircraft "write-ups," 31% wrote that a face-to-face aspect of the policy was not working well while 46% wrote that written or electronic logbook entries were problematic (Young et al., 2001). After completion of their research, Young et al. (2001) determined that the most common advantage noted by students was "learning to work with other disciplines." The maintenance database and its integration into the safety process is a key ingredient to improving the communications between pilots and mechanics. If the process is utilized efficiently there will be series of continuous communications instead of discussions only when an aircraft is broken.

Aircraft Discrepancy Analysis Matrix (ADAM)

As maintenance, databases and their integration into the safety process have been identified in previous studies as a key element for communication between pilots and maintenance technicians as well as for enhancing the safety culture. Because the Department of Aviation Technology at Purdue University is dedicated to enhancing safety, it has developed a matrix for observing, measuring, and analyzing the current safety status of Purdue aircraft. As an initial step, aircraft operators in conjunction with maintenance technicians developed the "Airplane Discrepancy Analysis Matrix (ADAM)," and have utilized it in many ways. A few examples are: safety management of flight operations, pre-flight instruction, management of aircraft maintenance, strategic planning involving budgeting and insurance, and aircraft life cycle.

Basically, ADAM displays the discrepancies of all Purdue University airplanes in a hierarchy of aircraft type (Figure 1). Currently, ADAM consists of a database that includes aircraft and maintenance historical data from the past three and half years of Purdue aircraft activities. The basic components of the ADAM database include airplane flight time, age (months), number of discrepancies, description of discrepancies, and level of discrepancies. These data are collected on a monthly basis and updated by the fifth day of the following month.

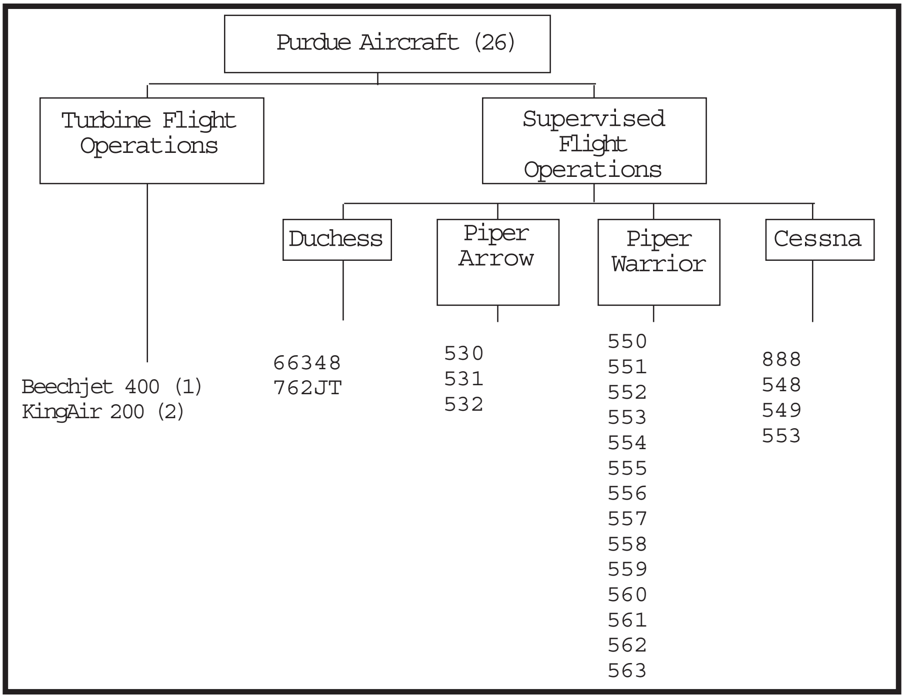


Figure 1. Composition of the Matrix

Once raw data from each airplane are collected and entered into the matrix, ADAM automatically calculates the history of discrepancy reporting for each airplane. It includes the cumulative number of discrepancies and cumulative probability of discrepancies. It is the probability of discrepancy that holds significant promise. If a student reviews a printout of the matrix and discovers that there is a history of discrepancies involving the alternator/generator or the aircraft radios then that student can review the appropriate emergency procedure to better prepare for the intended flight. To aid in the overall understanding of the matrix, all of the historical data are shown with respective graphic charts that create visually illustrated trends.

As soon as the updating of discrepancy data for individual airplanes is completed, the data are automatically transferred to the first categorized airplane groups, which represent each type of airplane such as Piper Warriors, Piper arrows, or the Duchess. Data in the categorized airplane group will summarize all the discrepancy data from each airplane belonging within the respective category.

Next, the first airplane group data goes to the second categorized airplane groups. These groups represent Student Flight Instructions (SFI) and Turbine Flight Operations (TFO). At this level, all the data from each airplane are summarized into two categories: discrepancies of the flight instruction airplanes and discrepancies of university flight operations.

Finally, all data that are summarized by the first and second category will be summed and analyzed at the level of all Purdue aircraft. This level shows the overall picture of all Purdue airplanes' current historical data including flight time, number of discrepancies, type of discrepancies, probability of discrepancies, visual tracking graphs, and so on.

Data, which has been created for demonstration purposes, has been included in (Chart 1) of the appendix and is a display of possible maintenance discrepancies. The information that can be gathered from this chart includes the various types of maintenance discrepancies that are tracked, the total number of class 1, 2, or 3 discrepancies and the probability of the class 1, 2, or 3 discrepancies occurring at any given point. This information can be derived for a given month, months, or for the entire year. Once the data has been entered, it is automatically graphed to display the trends of the total discrepancies and the probability of the discrepancies for the given months (Figures 2 & 3).

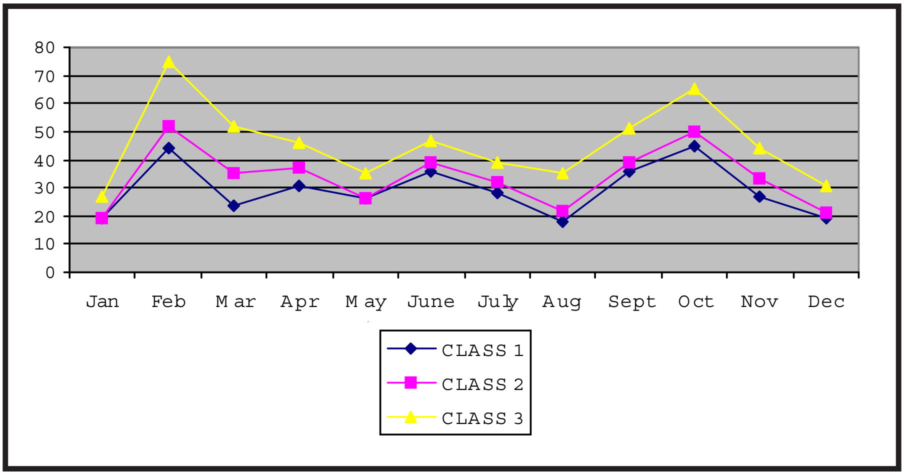


Figure 2. Total Number of Discrepancies

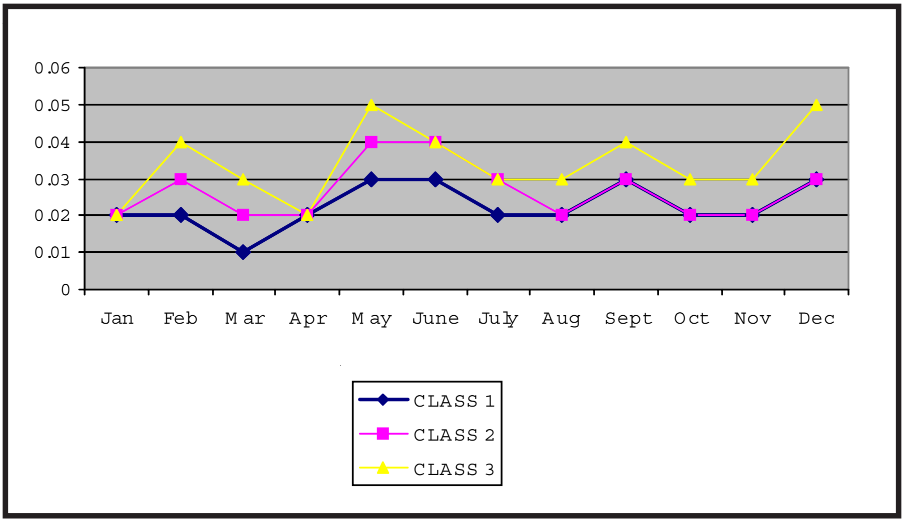


Figure 3. Probability of Discrepancies

Another unique feature of ADAM is that each of the collected discrepancy data are designated into a division of systems which were categorized through intensive discussions between management of flight operations and airplane maintenance personnel. As mentioned earlier the divisions are named as class one, class two, and class three discrepancies. Class 1 discrepancies are directly related to flight operations such as power plant, electrical, or navigation aid systems. Class two discrepancies add airplane lighting systems to class one discrepancies. Class three discrepancies describe all airplane discrepancies. For example, if an airplane had a magneto problem during a flight, this discrepancy would belong to the power plant category and designated as a class one discrepancy.

Some discrepancies tracked by ADAM are considered more important than other discrepancies. Engine failure would be considered more important than a discrepancy of the passenger safety belt locking system. However, we cannot ignore either of those discrepancies and, in both cases, the item would be fixed prior to the next flight.

On the maintenance side, some amount of effort will be required to fix a discrepancy, even if it is the change of a bulb or an engine replacement. Data from ADAM also will be useful in evaluating how many work force hours are required, in average, for the service of air operations.

For flight management, data from ADAM will help to evaluate the level of safe operations relating to the various aircraft. The data also can be used to predict how reliable and serviceable the airplanes have been over a given period of time, as well as provide budgeting insight.

For flight instruction purposes, flight instructors and students can review the data of ADAM for the assigned airplane, and acquire preflight knowledge of repeated or recent major discrepancies of the airplane. For authentic flight emergencies ADAM can help pilots determine possible causes of the problem and therefore assist in pilot decision-making.

Airplane Incident Analysis Matrix (AIAM)

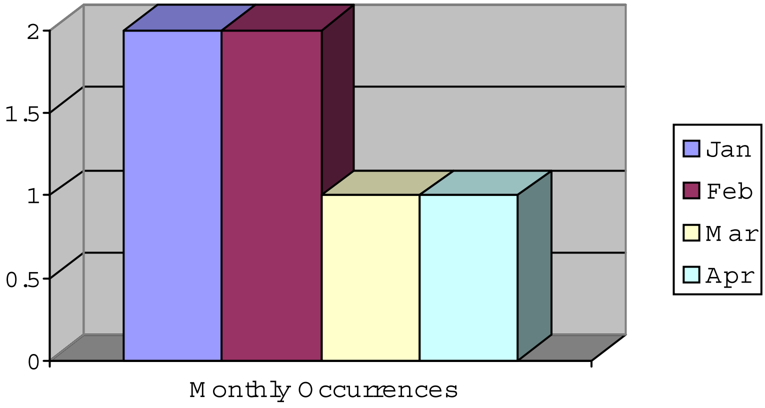
The ADAM database approaches the idea of the safety culture with a proactive method of preventing maintenance-related discrepancies from becoming a problem in flight. A system to track airplane incidents that are nonmaintenance-related also is required to achieve a holistic approach to risk management. As a secondary step to ADAM, the Airplane Incident Analysis Matrix (AIAM) was created to track these occurrences.

Basically,AIAM displays the discrepancies for an entire fleet of training aircraft and allows the information to be displayed in a graphical format so that trends and focus areas can be identified for targeted intervention. The information that is gathered for the database includes basic information such as the date, time, student and instructor (if dual) names, type and N-number of aircraft, the type of flight operation and basic weather observed at the time of the occurrence. More detailed information is gathered such as a basic description of the occurrence and a detailed "walk-through" of how the event transpired as viewed from the pilots involved and any observers. Both the basic and more detailed information is collected and a determination is made as to what type of occurrence category would apply. The possible categories are:

A – Accident
B – Aircraft Damage
C – Unauthorized Departure
D – Runway or Taxiway Incursion
E – Communication Error
F – Judgment Error

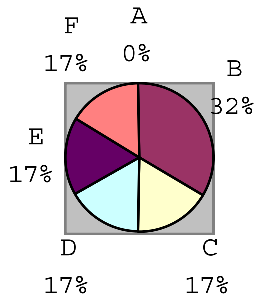
Data, which has been created for demonstration purposes, has been included in (Chart 2) of the appendix and is a display of possible aircraft problem occurrences.

After all of the data has been collected and inputted into AIAM, the database automatically displays the trends of occurrences in graphical form for various variables. Figure 4 displays the overall monthly occurrences and trends for all of the occurrences; while Figure 5 displays the type and percentage of occurrences for the whole database. These two graphs can be utilized to determine the effectiveness of a particular method of risk management that has been implemented or to determine focus areas for particular problems.



	Jan	Feb	Mar	Apr
CUMMULATE	2	4	5	6
MONTHLY	2	2	1	1

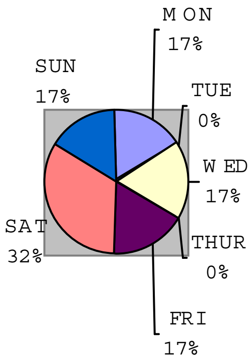
Figure 4. Monthly Occurrence and Trend



DESCRIPTION	ACCIDENT	A/C	Unauthorized	WY/TWY	COMM	JUDGEMENT	
		DAMAGE	DEPARTURER	INCURSION	ERROR	ERROR	TOTAL
CLASS	A	B	C	D	E	F	
NUMBER	0	2	1	1	1	1	6

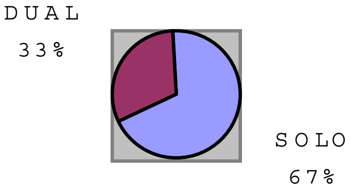
Figure 5. Type and Percentage of Occurrence

Figure 6 displays the number and percentage of occurrences that happen on various days of the week. Figure 7 displays whether the flight was dual or solo, Figure 8 & 9 display what type of operations were being conducted. Information from these four graphs can be utilized to illustrate the importance of increased vigilance for a particular day or operation if a trend is observed.



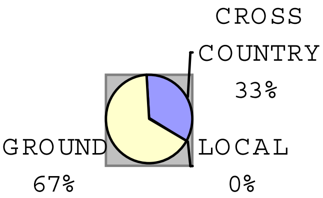
DESCRIPT	MON	TUE	WED	THUR	FRI	SAT	SUN	TOTAL
NUMBER	1	0	1	0	1	2	1	6

Figure 6. Weekday of Occurrence



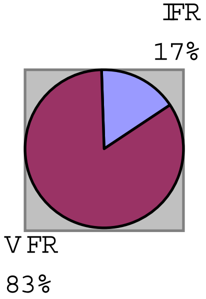
DESCRIPT	SOLO	DUAL	TOTAL
NUMBER	4	2	6

Figure 7. Occupancy



DESCRIPT	CROSS COUNTRY	LOCAL	GROUND	TOTAL
NUMBER	2	0	4	6

Figure 8. Location



DESCRIPT	IFR	VFR	TOTAL
NUMBER	1	5	6

Figure 9. IFR / VFR

All of the information gathered by AIAM can be utilized by safety conscious individuals to target and eliminate problem areas in a flight training situation where the possibility of occurrence is higher.

References

Comeau, L. (1986). Compiling data prevents school bus accidents. *School Business Affairs*, *52* (4), 42-44.

Damron, C. (1977). *Accident surveillance systems for sports*. Washington D.C.: American Alliance for Health, Physical Education, and Recreation. (ERIC Document Reproduction Service No. ED147267)

Geller, E. S. (2001). *The psychology of safety handbook.* New York, NY: Lewis Publishers.

Kluckhohn, F. R. & Strodbeck, F.L. (1961). *Variations in value orientations.* Westport, CT: Greenwood Press.

Kornblum, W. (1981). *Sociology in a changing world.* New York, NY: Harcourt, Brace, Jovanovich.

Leemon, D., Schimelpfenig, T., Gray, S., Tarter, S., & Williamson, J. (1998). *Adventure program risk management report: 1998 Edition*. Boulder, CO: Association of Experiential Education. (ERIC Document Reproduction Service No. ED443653)

Reason, J. (2000). *Managing the risks of organizational accidents.* Hamphire, England: Ashgate Publishing Limited.

Thorburn, S. (1990). Safety through education and training. *European Journal of Engineering Education*, *15* (1) 13-19.

Wojtowicz, G. & DesLauriers, K. (1995). T*eaching about hazard identification and injury control: A Student-based project focusing on pedestrian safety.* Charlotte, NC: Department of Health and Promotion and Kinesiology. (ERIC Document Reproduction Service No. ED389674)

Young, J., Mattson, M., & Petrin, D. (2001). Integrating the aviation system toward safety: Implications of interdepartmental human factors training for pilots and maintenance technicians. *Journal of Air Transportation World Wide*, *6*(1), 37-49.

Appendix

YEAR		TOTAL	AVERAGE	REMARK												
TRACKED MONTHS		12.0			Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
PWR PLANT		102	8.50	PER MONTH	2	7	9	13	8	11	8	3	11	14	9	7
FLT CONT		12	1.00		0	2	2	1	2	0	1	0	1	1	2	0
STRUCTURE		16	1.33		0	1	0	1	0	1	0	1	3	6	2	1
L/D GEAR		57	4.75		5	16	4	1	2	5	5	3	4	7	1	4
ELECTRICAL		37	3.08		3	5	1	5	7	5	2	3	1	2	2	1
COMM / NAV		56	4.67		3	4	2	3	1	7	8	3	2	13	6	4
INSTRUMENT		55	4.58		3	7	4	6	2	6	3	5	12	1	4	2
FUEL		15	1.25		3	2	2	1	1	1	1	0	2	1	1	0
LIGHTING		55	4.58		0	8	11	6	3	3	4	4	3	5	6	2
OTHERS		142	11.83		8	23	17	9	9	8	7	13	12	15	11	10
FLT TIME		17059.4	1421.62		1215.2	1944.5	1798.3	2036.8	701.2	1044.5	1199.6	1140	1435.5	2185.2	1700.1	658.5
NUMBER	CLASS 1	350	29.70	PERMONTH	19	44	24	31	26	36	28	18	36	45	27	19
	CLASS 2	405	33.75		19	52	35	37	26	39	32	22	39	50	33	21
	CLASS 3	547	45.58		27	75	52	46	35	47	39	35	51	65	44	31
PROBABILITY	CLASS 1	0.26	0.02	PERMONTH	0.02	0.02	0.01	0.02	0.03	0.03	0.02	0.02	0.03	0.02	0.02	0.03
	CLASS 2	0.30	0.03		0.02	0.03	0.02	0.02	0.04	0.04	0.03	0.02	0.03	0.02	0.02	0.03
	CLASS 3	0.41	0.03		0.02	0.04	0.03	0.02	0.05	0.04	0.03	0.03	0.04	0.03	0.03	0.05

Chart 1. Maintenance Discrepancies (Overall)

NO.	DATE & TIME			TIME VALUE	CATEGORY	SUMMARY
	DATE	DAY	TIME(L)			
1	Date	SAT	12:45		E	Comm Error
2	Date	MON	3:20		C	Event Depart
3	Date	FRI	4:00		F	Judge Error
4	Date	SUN	1:30		B	Aircraft Damage
5	Date	WED	9:30		D	Rwy Incursion
6	Date	SAT	12:05		B	Aircraft Damage

DESCRIPTION	DETAILS
Entered Class C airspace without 2-way communication	*Hyperlink to details
Used unauthorized airplane for solo flight	*Hyperlink to details
Flew to wrong airport	*Hyperlink to details
A/C rolled and hit another A/C	*Hyperlink to details
Failure follow GND instruction & entered unauthorized R/W	*Hyperlink to details
A/C wingtip hit a truck parked in grass while taxi back	*Hyperlink to details

PILOT		AIRPLANE	
SP	IP	TYPE	TAIL No.
Student 1		Piper	N***
Student 2		Piper	N***
Student 3		Piper	N***
Student 4	Instructor 1	Piper	N***
Student 5		Piper	N***
Student 6	Instructor 2	Piper	N***

IFR/VFR	LOCATION	OCCUPANCY	WEATHER		
			CEILING	VISIB	WIND
VFR	CC	SOLO	Clear	7	Calm
VFR	GND	SOLO	Clear	6	070 @ 10
VFR	CC	SOLO	Clear	7	120 @ 13
IFR	GND	DUAL	2000 Bkn	4	130 @ 9
VFR	GND	SOLO	Clear	5	Calm
VFR	GND	DUAL	Clear	6	260 @ 16

Chart 2. Airplane Accident/Incident Analysis Matrix (Overall)

International Journal of Applied Aviation Studies, 3,1

Adaptation-Innovation Theory of Cognitive Style in Aviation

R. Kurt Barnhart
Indiana State University
Department of Aerospace Technology
Terre Haute, IN 47809
Phone: (812) 237-2665
Email: r-barnhart@indstate.edu

Abstract

This article lends further credence to the important role psychology plays generally in education, and particularly in curriculum design. One basic dimension of personality is cognitive style, which influences virtually every aspect of human life. This article deals with Adaption-Innovation (A-I) theory of cognitive style, which is one of many ways to measure style and does so on a continuum from highly adaptive to highly innovative style preferences. In this aviation education application of A-I theory, it was found that aircraft pilots and pilot trainees had a distinct cognitive style preference in that they were a fairly adaptive group when compared to population norms. This fact has implications for curricular design which are discussed herein.

Introduction

As an aviation professional, having been involved in aviation education for the last several years, this author has found it to be one of the most rewarding and challenging experiences in the field of aviation. Those involved with education know the importance of continual curriculum design that often evolves along with the instructor's knowledge of the subject. Two of the other drivers of the curriculum design should be the needs and uniqueness of the learner. One way learners can be distinguished is by means of their cognitive style.

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

Cognitive Style

Cognitive style, which can be measured in a variety of ways, while related to learning style, is broader in scope. Each individual has their own unique cognitive style, a basic dimension of human personality; a means by which that individual organizes information, solves problems, manages change, and communicates with others. Effective instructional designers at least should consider the potential impact cognitive style might have on the teaching/learning process.

The fact that instruction that is tailored to the learner's cognitive style produces more effective learning is beyond question. As one example, research conducted by Carrell (1992) at the University of Wisconsin-Oshkosh examined the relationship between cognitive style and communication style among secondary education students and teachers. The most significant finding was that there was a significant main effect of cognitive style on student evaluation of teaching indicating, in her analysis, that "a match between patterns of thinking in teacher and student may actually affect learning; causing the student to believe that instruction which comes from his/her own perspective is more effective."

Using an example from aviation, Bowling Green State University researcher Quilty (1999) noted that "the ability to use different cognitive processing styles has implications on instructional techniques related to [CRM and new pilot training]." He went on to state, "it is suggested that communication, coordination and task completion can be optimized if cognitive preferences or biases are understood and appropriately considered in teaching CRM concepts."

A-I Theory

One powerful measure of cognitive style was described by A-I theory (and its associated inventory). This theory and its associated thirty-three item inventory (the KAI) was developed by Kirton (personal communication, September 2, 1999) while working in the field of organizational behavior in the mid-1970's. Kirton described preferred cognitive style in terms of a single dimension, with high adaption on one end and high innovation on the other. Those who have more adaptive preferences preferred a greater amount of structure as they approached and dealt with problems while those who have more innovative preferences prefer less structure. Within the theory this preference for structure lies along three basic dimensions: Style of originality, style of efficiency and style of rule/group conformity.

Style of Originality

This relates to an individual's preference to generate original ideas in problem solving. Individuals whose style is more adaptive prefer to generate a limited number of novel or original ideas and to focus on those that they consider to be the most effective. Those with a more innovative style prefer to proliferate ideas until the most novel approach is found.

Style of Efficiency

This relates to an individual's preference for organizational or system structure as it relates to accomplishing a task. The more adaptive preference here is to work within a given structure to attempt to solve problems within that structure while the more innovative preference here is to work outside or push the limits of the existing structure is order to achieve problem solution.

Style of Rule/Group Conformity

This relates to operating within the confines of organized rules, norms, and group consensus. The more adaptive preference is to attempt to solve problems within the confines of existing rules and procedures, while the more innovative preference is to go outside the rule boundaries to achieve resolution. As individuals interact and solve problems in their daily environment, they operate within their preferred cognitive style and tend to seek out groups and interact with people who exhibit the same style (Kirton, 1994b). If an individual is required to operate outside of the preferred style by interacting with an individual or groups who have a different style, the individual may need to employ "coping behaviors" that require an amount of effort related to the degree of difference, and become a source of stress for the individual. Thus it can be said that coping behavior is relatively expensive from a psychological standpoint (Kirton, 1994a). It has been shown that if this difference in mean KAI score differs by at least one standard deviation or more, coping behaviors will have to be "turned on" causing the potential for communication difficulty and interpersonal conflict (Kirton & De Ciantis, as reported by Kirton, 1999).

The differences in cognitive style preferences are known as cognitive gaps and must be managed in order for effective problem resolution. Cognitive gaps that are not recognized and effectively managed will often lead to frustration of the original effort and, at times, the complete disfunctionality of the group. Cognitive gaps can exist between two people, two groups of people, between a person and a group of people, or between a person and the requirements for a particular task.

A-I theory is based on several assumptions. First, all individuals have a preferred cognitive style which is not necessarily always linked to their observed behavior. An individual often may be required by their situation or environment to behave in ways that are contrary to their preferred style. This process is, again, known as coping behavior and all individuals must engage in coping behavior at different times in their lives, the degree to which depends upon their own preferred style and the requirements of the particular situation.

The second assumption concerns the important distinction between cognitive style and cognitive capacity. Cognitive capacity is often described in terms of "high" and "low," relating to one's cognitive ability such as I.Q. level. Cognitive capacity often is measured along a vertical scale with high considered as preferable and visa versa. Cognitive style frequently is thought of as being measured on a horizontal continuum from left to right in non-pejorative terms as it relates to one's preferences. This is similar to a left-handed person preferring to work with the left hand as opposed to a right-handed person.

Thirdly, change is a constant process to which all individuals must be attuned and in this process individuals will bring their own preferred cognitive style to bear in response to that change.

Fourth, A-I theory embraces the concepts that creativity, problem solving, and decision making are all concepts which are closely related to cognitive style and all human beings engage in, and are adept at, all three according to their own preferred cognitive style.

Fifth, all human-driven change implies some degree of structure. An absence of all structure is chaos. The distinction in A-I theory is the amount of structure preferred by a given individual in order to function.

A-I theory was selected for this study for two reasons. First, this theory's accurate description of human interpersonal dynamics is well documented by the fact that, when measured by the KAI inventory, people report seeing the described adaptive/innovative characteristics played out in their own lives (Barnhart, 2002). This is in addition to the fact that the validity and reliability of the KAI are well established. Secondly, little work has been done using the A-I theory in the field of aviation and the researcher felt there was a need to tie in what is known about A-I theory to aviation. In the only known application of KAI theory to aviation "Cognitive Style and Learning Performance of Adaptors and Innovators," Pounds and Bailey (1999) discovered that high adaptors performed better than high innovators on air traffic control tasks requiring strict adherence to rules of a given scenario, a fact that fits nicely with known KAI theory.

Associated characteristics of A-I theory are listed in Appendix A. In addition, Appendix B contains a list of average KAI scores by profession based on the literature.

Purpose of the Study

The purpose of this study was to determine the influence of cognitive style, as measured by A-I theory, in an aviation flight-training environment.

Research Question

What is the KAI cognitive style score profile for this sample of aviation flight professionals and trainees?

Data Collection and Methodology

Resource and access factors necessitated the use of a purposeful sample. Pursuant to that and to the above, current students and alumni (having graduated within the past five years) of Indiana State University's Professional Pilot degree program had the KAI administered to them. The Indiana State University (ISU) professional pilot program represented an average size undergraduate pilot education program in the United States with approximately 200 students (UAA, 1999). Data collection progressed from October 2001 through early March 2002. Current students had the KAI administered in a classroom setting while alumni were handled through the mail. All were volunteers. The Data Collection yielded 164 responses of 347 surveys distributed for a response rate of 47 percent. Of these respondents, the average age was thirty years. Ninety-two percent of the respondents were male and seven percent were female. Approximately 100 of the student responses were current students; the remainder were alumni. Of the alumni, 87 percent were employed actively as a pilot in industry; the remainder were from other career fields.

Results

The KAI score profile for all of the respondents was normally distributed and the results are summarized in Table 1. Note that the mean KAI score for the total respondents, which is the overall KAI score, was more innovative than that of the alumni, the majority of which are employed pilots. The average overall unpaired student score was relatively unchanged from that of the general population. This data suggested that most employed pilots had a markedly more adaptive cognitive style preference than the overall U.S. population and that students in general had a more innovative preference than their industry-employed counterparts. Also, note that the range of scores was narrower than for the overall U.S. population. Table 2 is a breakdown of alumni KAI scores by gender.

Table 1
Respondent KAI Score Detail

	Total Respondents	Paired Respondents	Unpaired Students	Unpaired Alumni	Overall US Population[1]
N	164	78	42	44	214
Mean KAI Score	92.21	92.18	94.52	90.24	94.98
Standard Deviation	15.21	14.44	14.45	17.24	15.90
Range	62 - 135	62 - 135	65 - 127	65 - 130	44 – 147

Table 2

Alumni Score Detail- all are employed pilots

	Alumni (Employed Pilots Only)	Male Alumni (Employed Pilots Only)	Female Alumni (Employed Pilots Only)	Females-General Population[1]	Males-General Population[1]
N	40	37	3	242	290
Mean KAI Score	90.66	90.16	97.00	90.84	98.12
Standard Deviation	16.10	16.42	11.53	17.82	16.75
Range	65 - 130	65 - 130	84 - 106		

[1] Kirton Manual, p. 69

Discussion

The findings contained herein were significant because this was the first study to examine the cognitive styles, as defined by KAI theory, of those involved in collegiate aviation education, as well as of those aviation flight professionals currently in the field. This information can now serve as the basis for the application of some of the other important aspects and implications of KAI theory in the future.

It is important to note that this group of current and future professionals was more adaptive, and had a smaller range of scores than the general population. Of the 38 (out of 44) Indiana State University Aerospace Technology alumni respondents who were actively employed as airline pilots, it was noted that their average overall KAI score of 90.24 was even more adaptive than any of the samples taken for the entire survey. This was an expected phenomenon as given by numerous conversations with Kirton regarding these results.

For the aviation environment, this information provided a valuable link to the larger body of KAI and cognitive style literature that might ultimately have implications for aviation and flight-training safety. As educators, it was important for those involved to be aware that, as a whole, they were dealing with a fairly adaptive group, certainly in terms of students, and most probably faculty as well. As further evidence, the range of scores showed that in aviation training we were dealing with individuals with a more narrow style preference than when compared to the overall U.S. population especially when considering that the range of scores in aviation training was somewhat skewed on the innovative end by two or three individuals, as can be seen by examining the mean from Table 1.

Implications for Curriculum Design

Implications for curriculum design in collegiate education were that the overall preferences for the students in this study were courses and assignments that were highly structured in nature due to their cognitive style. Individuals who have KAI scores in the range of the students in this sample prefer assignments that are more tightly managed and contain clear examples and guidelines as well as a clear set of expectations for satisfactory completion of these assignments. Assignments that are loosely structured will be a source of stress for students with these scores.

Interestingly, many of the students in ISU's collegiate aviation program reported that their general education requirements were burdensome and many reported disliking them altogether. Anecdotal evidence suggested that the average overall KAI score for college professors across the board is a more innovative score (on the order of 110 – 120). This is a large cognitive style difference from the average collegiate aviation education student score and the "cognitive gap" that occurs here might help to explain some of reported tension and dislike.

Also of interest was the average KAI score of all employed pilots. Individuals with a higher Adaptive style had a preference for a fairly high degree of task structure on the job. This is certainly the case with flying aircraft. A pilot's job very much involves following checklists, procedures, and operating within the confines of a tightly controlled and highly structured air traffic system. The researcher has spoken with pilots who score on the Innovative end of the scale and who feel very much confined and stifled by such a system. Concerning air safety (an area of expertise for the researcher), as the aviation safety record bears, these adaptive preference pilots were very well suited to dealing with foreseen difficulties (for which there is already a prescribed procedural remedy), which describe the vast majority of aviation safety-related occurrences. Conversely, pilots as a group who have these KAI scores may not be as well suited for dealing with situations for which there is no prescribed procedure. Such situations do occur, although rarely. Of practical significance here, is for those in aviation training and curriculum design to specifically concentrate on training which helps pilots cope with the vague environment of an occurrence for which there is no prescribed procedure.

Gender Differences

Another item of note that is seen in Table 4 of Chapter 4 was the score disparity between male and female pilots. Although the small sample here precluded any generalizations to a larger population, it would be interesting to determine if the score difference in this sample would hold true in a larger sample. These preliminary findings suggested that female pilots tend to have a more innovative style preference as a group than do their male counterparts. Reasons for this could be the subject of another project. On the other hand, the average score for U.S. males was 98, somewhat more innovative than the general population average (KAI Feedback Booklet, 5). Comparing the male pilot alumni overall score of 90.66, it can be seen that males who choose this profession have a markedly greater tendency towards an adaptive preference than do males in the general population.

To conclude, the researcher hopes that this small bit of information can aid the aviation educator to better design instruction as they consider the cognitive diversity of their students.

References

Barnhart, R. K. (2002). *The Application of Kirton Adaption-Innovation Theory to the Flight Instruction Environment of Higher Education.* (pp. 1-84). Ph.D. Dissertation Indiana State University.

Carrell, L. J. (1992). *The relationship between cognitive style and communication style in the secondary school context* (Doctoral Dissertation, University of Wisconsin-Oshkosh 1992), US Department of Education, ERIC reference number, ED 355565.

Kirton, M. J. (1999). *Kirton Adaption-Innovation Inventory Manual, 3rd. ed.* Berkhamsted, UK: Occupational Research Centre.

Kirton, M.J. (1999) *Kirton Adaption-Innovation Inventory Feedback Booklet* – M. J. Kirton, 1999.

Kirton, M. J. (1994a). A theory of cognitive style. In M.J. Kirton (Ed.) *Adaptors and innovators, styles of creativity and problem solving*, 1-33. New York, NY: Routledge.

Kirton, M. J. (1994b). Adaptors and Innovators at Work. In M.J. Kirton (Ed.) *Adaptors and Innovators, Styles of Creativity and Problem Solving*, 51-71. New York, NY: Routledge.

Pounds, J. J. & Bailey, L. L. (1999). *Cognitive Style and Learning: Performance of Adaptors and Innovators in a Novel Dynamic Task.* FAA Report abstract given via email with the author, Pounds, (1 March 2001).

Quilty, S. M. (1999). Analyses and comparisons of cognitive learning preferences among airline pilots, corporate pilots, and aviation students. *Journal of Air Transportation World Wide, 4* (1), 1-13.

University Aviation Association, Telephone conversation, Spring 1999.

Appendix A

Characteristics of Adaptors and Innovators

Characteristics of adaptors and innovators
(Adapted from the KAI feedback booklet, M.J. Kirton, 1999)

	ADAPTORS	INNOVATORS
Perceived Behavior	Sound, conforming, safe, predictable, inflexible	Glamorous, exciting, unsound, impractical, risky
In Problem Definition	Accept as defined, prefer to limit disruption, need to see short term benefit	Reject generally accepted definitions, look at long term gains
In Solution Generation	Prefer a few novel, relevant and acceptable solutions aimed at improving what's existing	Prefer numerous ideas possibly not appearing relevant, prefer to do things differently
In Policy Formulation	Prefer well established, structured situations- good at incorporating new ideas into established situations	Prefer less structured situations- use new ideas to create new structures- accept greater risk
In Organization	Essential for ongoing functions but have difficulty with change in moving out of existing functions	Essential in times of change but have trouble applying themselves to ongoing organizational demands

Adaptors and Innovators in Collaboration

The Principle	Groups need both adaptors and innovators to be effective over time	Small Gaps	The narrower the thinking diversity range, the more limited the range of problem solving potential; within this range high efficiency is the norm
Problem of Large	The larger the gap between people's scores on the KAI, the greater the problem communicating and collaborating even if both are adaptors or innovators, it is the gap size which is the problem not location on the scale	Bridgers	Those who happen to have an intermediate score within a group- may be helpful in bridging the gap between two sides
The Advantage of Large Gaps	The wider the difference the more effort and tolerance is needed to stay together but the greater is the group's breadth of problem solving	Coping Behavior	Allows people to play successfully a role to which they are not naturally suited- creates stress and is difficult over long periods
		Leadership	Good leaders ask for minimum coping behavior most of the time and get offered maximum coping behavior in a crisis

Appendix B
Occupational KAI Norms

Occupational Group Mean KAI Score Studies (Kirton, 1999)

Occupational Group	Country	N	Mean	S.D.	Author
Apprentices	U K	624	83.6	9.8	Flegg (in Kirton 1994)
Bankers	US/UK/Italy	217	91.3	14.0	Gryskiewicz et. al., 1986; Holland 1987, Prato Previde, 1984
Clerical Staff	UK/Italy	205	89.2	16.4	McCarthy 1988, Prato Previde, 1984
Managers	Singapore	75	95.0	12.6	Thomson, 1980
Managers	Singapore	695	96.3	11.3	Thomson, 1985
Managers	U K	79	96.9	16.4	Kirton, 1980
Managers	U K	88	97.1	16.9	Kirton, 1980
Managers	Italy	207	99.3	17.4	Prato Previde, 1984
Managers	U K	93	98.5	14.9	Lowe & Taylor, 1986
Managers	U K	192	102.2	14.2	Davis (in Kirton & Pender, 1982)
R&D Prof.	USA	256	100.9	?	Keller & Holland, 1987a
Members: Committees for Community Adult Ed. Program	USA	208	101.9	15.8	Cutright & Martorana, 1989
Education Personnel	U K	79	103.0	17.1	McCarthy, 1988
Bank Management Trainees	US/UK	127	97.6	16.4	Holland, 1987
Teachers	USA	430	95.0	12.8	Pulvino, 1979
Teachers	USA	202	97.0	14.0	Dershimer, 1980
Teachers	USA	80	101.4	14.4	Jorde, 1984
Teachers	U K	182	94.5	18.2	Kirton et al., 1991
Nurses	USA	77	92.2	14.9	Ligman, 1991
Nurses	USA	60	92.3	12.0	Pettigrew & King, 1993
Nurse, Chief Administrators	USA	147	108.9	12.6	Adams, 1988
Doctors- Gen. Prac.	U K	180	91.9	16.1	Salisbury et. al. 1998
Engineers		800	96.81		Kirton, 1999

International Journal of Applied Aviation Studies, 3,1
Copyright © 2003, FAA Academy, Oklahoma City, OK

Training Development Papers

The Critical Components of Aviation English

Marjo Mitsutomi
University of Redlands
School of Education
University Hall North, P.O. Box 3080
2100 E. Colton Avenue
Redlands, CA 92373-0999
Email: marjo_mitsutomi@redlands.edu

and

Kathleen O'Brien
Federal Aviation Administration Flight Standards
Long Beach Flight Standards District Office
5001 Airport Plaza Drive, Suite 100
Long Beach, CA 90815Abstract

Abstract

Aviation safety depends on accurate pilot-controller dialogue. Both must be able to negotiate meaning through language at all times. Communicative competence in aviation English means that airmen have common and standardized proficiency levels in their use of the English language. This paper defines the critical components of aviation English as air traffic control (ATC) phraseology, English for Special Purposes (ESP) and English for General Purposes (EGP). The Aviation English Model also is introduced to serve as a framework for subsequent discussions on language issues as they relate to the global aviation context.

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

The Critical Components of Aviation English

International air operations continue to increase in 2003 as industrial powers expand their customer base and emerging countries move aggressively into the marketplace. In the United States alone, according to the Wilber Smith Associates (as cited in Heimerman, 1997), the aviation industry contributes about five percent of the annual U.S. gross domestic product, and there are strong economic incentives to maintaining a safe and healthy civil aviation industry. Along with the increase in air commerce activity, there is the push to expand system capacity by increasing air traffic density. As the runways and airways get busier, the need to communicate ever more efficiently and accurately also is multiplied. The issue of efficient communication in the aviation context is complicated by the fact that the participants in the complex global system represent multiple ethnicities, languages, and cultures. According to Day (2002), "Language is an imperfect medium for communication, but with awareness of basic linguistic principles, operating personnel can be motivated to adhere more closely to standard phraseology in all air-ground radio exchanges, thus enhancing safety" (p. 24).

It is now becoming common knowledge that pilots and air traffic controllers must be communicatively competent in English as a language, not just the specialized jargon used in air traffic control (ATC) communications. This new realization has caused an urgent need for English language proficiency standards both in domestic as well as global aviation. Since discussions about proficiency in aviation English will only intensify in the near future, it is important that all interested parties have a common understanding of the underlying terms and concepts that relate to the development of minimum proficiency standards in aviation English. This applies equally to native and nonnative English speakers.

This article will examine the role of the English language within the aviation context, particularly as it relates to the pilot-controller dialogue. Our goal was to explain the role and function of English in this dialogue. We will identify the three distinct areas of language use that, when taken together, form the basis of English proficiency for safe communications:

1. ATC phraseology
2. English for Specific Purposes (ESP)
3. English for General Purposes (EGP)

We will discuss the role of English as a lingua franca of today's business world and its implications on global aviation. Finally, we will offer a framework, a common frame of reference, for further explorations and discussions on language and aviation.

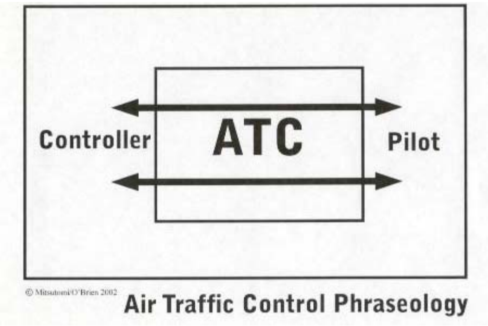


Figure 1. Air Traffic Control Phraseology – Assumption: ATC phraseology contains most, if not all, the phrases needed for standard routine aviation procedures. The phraseology is expected to suffice in most emergencies.

Background

The ATC is in existence to resolve conflicts between participating aircraft. Elaborate systems define procedures and the system works well. When pilots and air traffic controllers speak to one another in the professional context, it usually takes place in prescribed, coded language, called ATC phraseology (also known as radiotelephony) as shown in Figure 1. This phraseology is used routinely all over the world allowing pilots to fly across linguistic as well as national boundaries and still be understood by their foreign peers.

Nature of ATC Communications

Although differences between the International Civil Aviation Organization (ICAO) and the Federal Aviation Administration (FAA) phraseology do exist, the similarities far outweigh the discrepancies. It is important to note that the type of language which is supposed to be used in the ATC context is not tied to any particular culture or local variety of English. It is the mutually agreed upon and studied variety of language that uses English as its basic structure but focuses solely on communicative needs in aviation.

Another feature that separates ATC phraseology from general (also known as natural or plain) language is that the usage is standardized and non-idiomatic. All aircraft flying in controlled airspace adhere to certain standard procedures. These operations have accompanying standard phraseology, thereby allowing all the parties in the air and on the ground at any given time to stay informed about the progress of the flight. All the prescribed and predetermined expressions used in this context are self-contained and limited to the set sanctioned by the appropriate aviation authority. The Air Traffic Controllers' Handbook, 7110.65 (Air Traffic Control Services, The Federal Aviation Administration), contains extensive listings of words, phrases, and sentences to be spoken in a myriad of situations. As the airspace is increasingly busy, there is little time for friendly chatter or conventional politeness and niceties, although they do occur.

Most veterans of the skies seem to be able to understand one another's intents with amazingly few miscommunications. As reported by Mell (http://www.icao. int/anb/sg/pricesg/background/OotB.htm), this communicative success is largely due to three factors: the use of internationally recognized phraseology, a restricted number of topics, and the predictable and repetitive nature of the communications.

The phrases used in the radiotelephony context are designed to make the communicative function between the ground and aircraft as concise and brief as possible, with the emphasis on accurate content as opposed to linguistic form. The brevity and conciseness of the communication is accomplished partly by using formulaic and predetermined sentence fragments as opposed to complete sentences. Typically, grammatical markers, such as determiners (the/a) and auxiliary verbs (be) are deleted, this feature making the ATC communications markedly different from natural language. Example:

> *American Airlines flight 54, turn left heading 100, intercept the localizer and proceed inbound, cleared for the ILS approach to 13 Right, maintain 2,200 until established. Contact tower on 120.6 at NOLLA.*

The beauty of the ATC communication is that all parties know what is expected of them in terms of their performance-related procedures, both technical and linguistic. Thus, those who undertake aviation studies with the goal of either flying or directing traffic will memorize this standard phraseology in English, whether they speak the language as a native or not. However, air traffic communications outside the United States do not always take place in English .

Limitations of ATC Phraseology

Within national borders of other countries, when pilots and ground station personnel share the same language, the communication may take place in their native language. However, when pilots or controllers do not share the language of the ground station, both parties are expected to communicate in English, using ATC phraseology. Using this practice, all information relayed between air traffic and the aircraft use language that is comprehensible to all those intimately involved in the operations. To further assure the safety and efficiency of international air operations, ICAO members in March of 2003 directly addressed the importance of this issue by elevating the use of English from a recommended practice to the standard practice (B. Day, personal communications, March 5, 2003).

One obvious challenge in pilot-controller dialogue for both native speakers and nonnative speakers is the memorization and mastery of the ATC phraseology. To improve communication between both groups, Mathews (2001) observed that "the need for closer conformity to standard phraseology and for greater care in communication on the part of native and non-native speakers alike becomes readily apparent" (p. 26). The frequent use of informal language in place of standard phrases is an area of concern. This practice relates to personal usage of language rather than personal ability in language and requires discussion outside of the proficiency issue.

All novice airmen (in this context we mean pilots and air traffic controllers) begin at the same place, not knowing what the phrases are and how to use them. With practice, they eventually become fluent in the use of ATC phraseology. The differences and limitations that ultimately emerge in a particular airman's ability to communicate correctly and efficiently with ATC can partially be attributed to frequency of practice and quantity of experience. Other factors such as timidity, fear, and anxiety can also have a negative effect on communicative success. The goal at all times is the proper use of ATC phraseology and that goal, excluding any unusual circumstances, is achievable.

English for Specific Purposes

Background

For pilots and controllers to successfully send and receive messages from one to the other, they are expected to use and understand their common, shared, and standardized medium of communication: the ATC phraseology. The already mentioned ATC Controller's Handbook covers, as Villaire (1994) noted, "almost every situation, instruction and communication request imaginable, and each paragraph is based on information bought and paid for by someone's disaster." Much memorization and recall is required to be able to produce the appropriate phrase each time.

Besides mastery of the phraseology, pilots and controllers also need an intimate understanding of their area of study with the related technical and practical applications. When pilots and air traffic controllers first embark on their respective studies, they enter a highly specialized and technical world with its own language, the ATC phraseology being a subset of the larger whole. Orr (2002) defined this language, ESP, as follows: "specific subsets of the English language that are required to carry out specific tasks for specific purposes" (p. 1). ESP consists of vocabulary and concepts which are "unfamiliar to most native and nonnative speakers and thus require special training" (Orr, 2002, p. 1). In the following examples, common English words such as *base*, *three o'clock*, and *clear* have aviation-specific meanings:

> Turn base now, follow traffic at your three o'clock, cleared for the option.
> Remain clear of Class Charlie airspace, contact approach on one two three point six five.

For pilots and controllers to be successful in their careers, they must possess this type of specialized literacy. Their studies share many of the same topics and themes such as weather, emergency procedures, radio calls, etc. It is exactly this common core of shared knowledge that allows pilots and controllers to speak to one another; they understand the specialized world of flying each from their own perspective. They send messages to each other which are primarily related to their immediate situation, and they expect those messages to be received and understood as they were originally intended (see Figure 2).

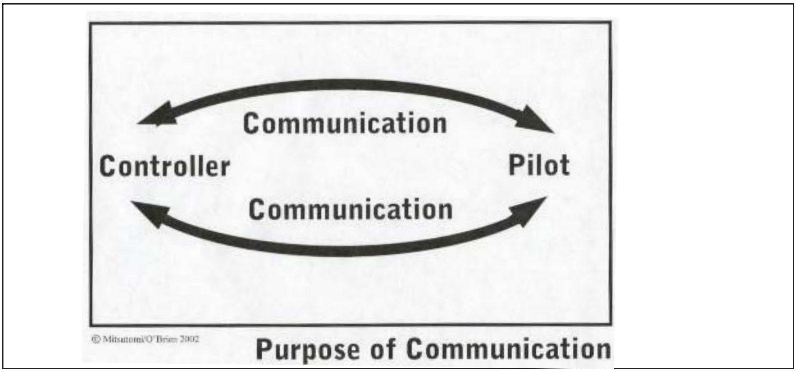


Figure 2. Purpose of Communication - Communication is the process of sending and receiving messages. Communication is successful when the sent and the received messages are similar enough to trigger the expected response.

Nature of ESP

Like the ATC phraseology, the ESP component of aviation English does not exist in a vacuum. It is rooted in general language. It is the use of the English language itself that gives life to the specialty area. EGP carries the specialty language within its sentence structure, vocabulary, and paragraph organization. The ability to understand text (written or spoken) in a specific subject is referred to as content area literacy. Expertise in related fields and shared experiences make the communication between pilots and air traffic controllers contextually meaningful.

As with ATC phraseology, airmen must master the ESP jargon. All airmen begin at the same place with the goal of becoming fluent in its use through practice. Here again, some of the differences and limitations that ultimately emerge in a particular airman's ability to discuss topics related to the ESP can be partially linked to practice and experience. Excluding unusual circumstances, competency in the use of aviation-specific ESP is an achievable goal for most airmen.

However, it needs to be mentioned that the vast repertoires of linguistic and technical information place a heavy burden on the brain's capacity to handle several pieces of information at once. In discussing the role of controllers to keep aircraft separated from one another, Villaire (1994) warned that the issue of cognitive workload also comes into play. Although outside the scope of this study, cognitive workload must be factored into the entire equation.

When recall of discrete linguistic points (including the phrases themselves and the conventions controllers have agreed on) is added to the multifaceted list of tasks, successful communication becomes even more demanding. To avoid a "linguistic stall" (Mitsutomi, 1999), the pilot-controller dialogue requires language readiness that goes beyond the current assumption that ATC phraseology is sufficient. It is not.

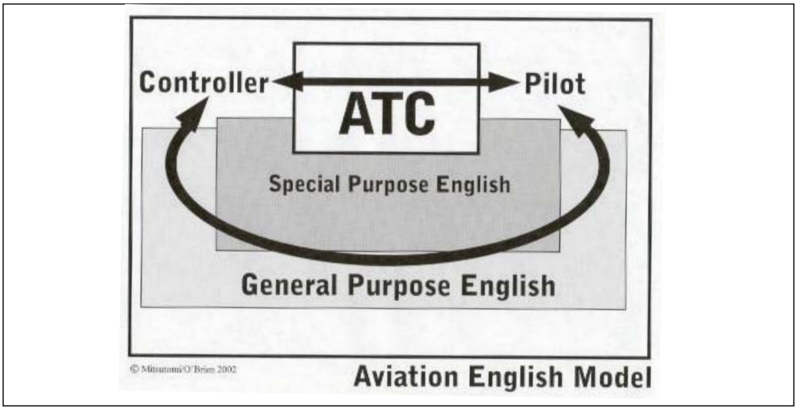


Figure 3. Aviation English Model – ATC phraseology remains central to aviation communications. However, when the phraseology does not suffice, airmen will possess a standard of proficiency in the EGP that will be sufficient communicate in all possible situations.

Background

Aviation accidents have always fascinated the public, and flying is feared by thousands of people. For many years it has been recognized that communication problems are implicated in many aviation accidents and in runway incursions. One of the most dangerous places for aviators is on the ground. In 1977, the world was shocked by the collision of two, new, giant Boeing 747 aircraft at Tenerife. A Pan Am 747 missed or misunderstood taxi instructions which required a turn off the active runway at taxiway three. At the same time, a KLM 747 initiated a fog-shrouded takeoff in the opposite direction. The two aircraft met on the active runway with the KLM at approximately 160 Knots. Five hundred eighty two (582) died in the crash.

O'Hare and Roscoe (1994) pointed out that although the vast majority of flights operate smoothly and without incident, misunderstandings between air traffic controllers and pilots, or between pilot to pilot, have played a major role in a number of accidents. A familiar example of ambiguity in communication is the instruction "takeoff power" issued by the pilot to initiate a missed approach procedure. In several cases this phrase has been interpreted by the first officer as an instruction to reduce (take off) power. Such misunderstandings have led to the replacement of this phrase by the potentially less ambiguous "go-around power".

In all of life, unusual and unexpected things happen and aviation is no exception. Emergencies crop up, inexperience causes havoc, and other unpredictable things occur routinely. This is when the pilot-controller communication is put to test. When working together using language, ATC or otherwise, they must be able to address the emerging situation quickly, accurately, and precisely. Recently, the aviation community, being influenced by cross-disciplined information, has turned its attention to language competency itself.

The ability to communicate when there is no prescribed script (ATC phraseology) is critical to safety. In practice this means that pilots and air traffic controllers must have the ability to achieve mutual understanding through the use of their general language ability to get their messages heard and understood. As Figure 3 illustrates, this ability to negotiate meaning at all times is the key to communicative competence.

Communicative Functions of Aviation English

Examining the communicative functions in pilot-controller dialogue has revealed the following four categories to be the most dominant (J. Mell, personal communication, July 5, 2002):

1. Triggering actions
2. Sharing information
3. Managing the pilot-controller relationship
4. Managing the dialogue

As Mell explained it, the ability to trigger actions is at the core of the pilot-controller exchanges. Giving orders and requesting actions or permission to do something are speech acts that trigger specific responses. Information sharing, on the other hand, is related to one's intentions, actions, readiness, or availability of something, etc. Relationship and dialogue management include greetings, complaints, paraphrases, checking for understanding or confirmation of information, etc. These actions need to be accomplished in pilot-controller dialogue.

Each category described by Mell contains all the pertinent ATC phrases which are always to be used first. However, when the situation so requires, general language (GL) in addition to the ATC phraseology must be accessible to the speaker as well. It is this issue of GL use that has been problematic in aviation contexts. Strict adherence to phraseology is always preferred, but situations arise for which there is no phrase or the phrase needs to be expanded upon with real-time information.

Proficiency in EGP

The problem with communication particularly in global aviation is that pilots' and air traffic controllers' abilities to use general English varies considerably. Some can only parrot the memorized ATC phrases, other are comfortable functioning in English in any situation. This vast discrepancy in airmen's English ability is precisely the reason for concern especially of international flight crews. As a matter of fact, the capability of flight crews and controllers in the recommended common language, English, is an unexamined area. It is safe to assume that everyone in the cockpit and tower has mastered at least the basics of ATC communication as it is part of one's course of study. What cannot be assumed, however, is that these same people have basic conversational ability in general English since it has not been commonly regulated in many training facilities.

Recognizing then, even if reluctantly, the need for competency in general English to complement the use of ATC has been the driving force behind the ICAO PRICE (Proficiency Requirements in Common English) Study Group and the FAA PEC (Pilot English Competency) Working Group for the last few years. The task of these groups has been to define the minimum level of proficiency in English needed to communicate safely at all times. ICAO has pioneered the way by already describing this minimum level of proficiency that facilitates speaking and understanding English in usual and unusual aviation-related contexts. The scale applies to native and nonnative speakers of English. This ICAO scale is published and currently available to be viewed on the ICAO website. (For a brief description of one milestone study regarding setting English proficiency standards, refer to Chatham and Thomas, 2000).

The Aviation English Model

"Communicative competence specifically in aviation means that pilots, air traffic controllers, mechanics, and ground crews can exchange important information in mutually intelligible messages" (Mitsutomi and O'Brien, 2001, p. 4). For aviation communication to be successful, the following is assumed: mastery of professional jargon or phraseology, including standard situations and standard procedures. The ATC phraseology contains expressions for all of the above functions, and they work very well most of the time. But for those times when the ATC phraseology does not "do the job," the call is out to use general English (EGP) which, in the aviation context, will consist mostly of aviation-specific topics and vocabulary (ESP). Unlike ATC phraseology, general English is not tied to a prescribed code (as illustrated in Figure 3) but is flexible, allowing the speaker to manipulate it to get the desired message across. It also allows the speaker to produce novel utterances that satisfy the communicative needs of the moment. In other words, general English facilitates thinking in English, outside of the "ATC box," and that can be the difference between life and death. Excluding unusual circumstances, competency in the use of EGP at the proficiency level specified in the ICAO scale is an achievable goal for most airmen.

The Role of the English Language in Global Aviation

Background

Within global aviation, pilots fly in and out of countries where the controllers speak English in distinctly different ways. The ICAO PRICE Study Group in its efforts to establish English proficiency guidelines for a very diversified membership has been keenly aware of the challenges created by the many varieties of English. The recently developed English proficiency description, therefore, in its very definition places the burden of successful communication on native (NSs) and nonnative (NNs) speakers alike. The scale states that proficient speakers of English "use a dialect, accent or variety which is intelligible to an international community of English language users."

It is of no small significance to note that the ICAO proficiency scale does not measure NNSs against NSs, using the latter group as the norm. As Mathews (2002) cautioned, "The proficiency requirement and scale were developed with both native and nonnative speakers in mind and are applicable to both groups." It puts all English speakers in the same category of "English users," requiring all to be careful and considerate users of the shared language. In the recent International Aviation English Association Seminar in Warsaw, Day (2002) noted that in the arena of international aviation, "English sheds all connection to political agendas, real or perceived, and becomes simply another tool for increased safety and efficiency of aviation operations."

English as the World's Lingua Franca

Indeed, there are many users of English today; according to some conservative estimates, approximately 1.6 billion use it to some extent daily (Geary, 1997; Fishman 1999). These English users can be divided into three major categories. Kachru (1988), in his groundbreaking work in the field of linguistics, divided English use into three circles, ranging from the extended circle to the inner circle. He labeled native speakers of English as being in the inner circle, the "insiders" as it were, those with the original rights to English as their native language (for example, the United Kingdom, the United States, Australia). This group consists of approximately 400 million people. The outer circle includes countries where English has an official position, such as India, Nigeria, and the Philippines. The speakers of English in the outer circle use it as their second or third language. In many of the countries in the expanding circle, English has no official purpose nor is it the language typically spoken on the streets. Yet, the expanding circle is the fastest growing of the three as it absorbs all those nations where English is being studied as a foreign language in addition to any other native or second languages. The list of the countries in the expanding circle is long, indeed, and includes small and large nations alike.

Native and Nonnative Speakers

English no longer belongs just to the native speakers. Nayar (1994) questions to whom English belongs and concludes the following: "It is ours and everyone's: the English language is truly a world possession" (p. 4). Because of the vast numbers of Englishes (established varieties of English) in the world today, many students of English are introduced to the major native speaker varieties but are not expected to adhere to any one of them as *the correct* variety. In fact, in many parts of the world, nonnative speakers of English never meet "authentic" native speakers of English but communicate solely with other NNSs whose accents may be distinctly different from their own. Incidentally, not all NSs of English understand one another always either. Not all varieties of English are mutually intelligible. It is not uncommon for native speakers of different varieties of English to experience difficulty understanding one another.

In order to communicate in English globally, the speakers of this same language must be aware of the major features of the other varieties. Accent is only one of those features although it may be the most salient one. Speakers of various varieties of English, native and nonnative, must also be willing to adjust their own speech so it becomes more understandable to those outside of their own community. The reality is that, "Communicating across cultures is no longer a goal; it's a mere starting point" (Anthony, 2000, p. 4).

Since English usage in the world is so commonplace, the NSs can no longer dictate to the rest of the world how it is to be properly spoken. The rules of cross-cultural communication must be defined by those who represent the various cultures, together forming an international community of English users.

World Englishes

Within the field of linguistics itself, identifying and analyzing the role of English in the world has been the major focus of a branch of linguistic study. Conceived and founded by Kachru some two decades ago, the premise of World Englishes (WE) is to recognize and examine all of the many native and nativized varieties of English as legitimate. Besides engaging in this analytical study, WE promotes the use of English as an International Language (EIL) serving as the medium for cross-cultural communication.

The official recognition of a type of English that belongs to no particular group is an exciting development for the global aviation community. EIL is a generic form of general purpose English, if you will, which is capable of accommodating the many differences stemming from local cultures and varieties of speech, including accent. While independent from regional variations and peculiarities, EIL uses the basic structure of English syntax and lexicon to make communication among English speakers possible. Global aviation requires effective communication among all the participants. This is certainly a crucial consideration in an age where English, although not everywhere, is "just closer to being everywhere than any language in history" (Anthony, 2000, p. 7). It is certainly the language of the skies.

Conclusions

Regardless of current and yet-to-be invented technology, language will always remain central to communication between humans. The communicative context in aviation requires pilots and air traffic controllers to send and receive messages primarily through the medium of ATC phraseology. There are times, however, when general language ability is necessary, times when the limited ATC phraseology fails to suffice the needs of the communicative situation. The ATC phraseology is useful but only in limited situations and lacks the dynamic energy of a "living" language.

Since aviation safety depends on accurate pilot-controller dialogue, both must be able to utilize all that language has to offer. In the ATC context, this means the ability to ask and answer questions, follow instructions, narrate events, describe situations, and paraphrase information; i.e., to be a full participant in a two-way dialogue. Pilots and controllers must be able to negotiate meaning through language at all times and under all circumstances. Communicative competence in aviation English means that airmen have common and standardized proficiency levels in the following three critical components: highly specialized ATC phraseology, ESP as it applies to aviation, and the foundational EGP. The three together form the linguistic safety cushion that will significantly enhance safe communications in the aviation context world-wide.

References

Anthony,T. (2002). English: 1 tongue for the new global village. *The Associated Press*. Retrieved February 26, 2002, from http://wire.ap.org

Chatham, R. L. & Thomas, S. (2000, March/April). Proposed English standards promote aviation safety.*ESL Magazine*, 20-23.

Day, B. (2002). Safe radiotelephony demands good discipline from all pilots and controllers. *ICAO Journal*, *57*, 24-25 & 33-34.

Day, B. (2002, September). Proposed ICAO proficiency requirements in common English. Paper presented at the 8th International Aviation English Association Seminar, Warsaw, Poland.

Geary, J. (1997, July 7). Sowing the seeds of speech. *Time*. Retrieved February 26, 2002, from http://www.time.com.

Heimerman, K. T. (1997). Air traffic control modeling. *Frontiers of Engineering.* McLean, VA: The Mitre Corporation.

International Civil Aviation Organization. (2000, November). English language proficiency training and testing in the United States. *Proficiency requirements in common English study group first meeting* (PRICESG-SN/05). Montreal, Canada.

Kachru, B. (1992). *The other tongue: English across cultures*. Urbana: University of Illinois Press.

Mathews, E. (2001). Provisions for proficiency in common aviation language to be strengthened. *ICAO Journal*, *3*, 24-26 & 41.

Mathews, E. (2002, September) *The proposed ICAO language proficiency requirements*. Paper presented at the 8th International Aviation English Association Seminar, Warsaw, Poland.

Mell, J. (n.d.). *Emergency calls – messages out of the blue.* Available from International Civil Aviation Organization Website, http://www.icao.int/anb/sg/pricesg/background/OotB.htm

Mitsutomi, M. and O'Brien, K. (2001). Global aviation safety and English competence. *Proceedings of the 2nd Annual American Society of Behavioral Sciences*, Las Vegas, NV.

Mitsutomi, M. (1999). PEC (Pilot English Competency) before PIC (Pilot in Command). *Proceedings of the International Aviation Training Symposium*, 351-358, Oklahoma City, OK.

Nayar, P.B. (1994). Whose English is it? *TESL-E*, 1,1, F-1. Retrieved February 26, 2002, from http://www-writing.berkeley.edu

O'Hare, D. and Roscoe, S. (1994). *Flightdeck Performance*. Ames, IA: Iowa State University Press.

The Air traffic controllers' Handbook. (n.d.). 7110.65J. Air traffic control Services, Federal Aviation Administration.

Villaire, N. (1994). *Aviation safety: more than common sense*. Casper, WY: International Aviation Publishers, Inc.

International Journal of Applied Aviation Studies, 3,1

Analysis of Some Aviation Rules of Thumb

William R. Trippett
Central Washington University
400 E. 8th Avenue
Ellensburg, WA 98926-7515
E-mail: w7vp@premier1.net

Introduction

On his internet web site entitled "Statistics," McBride (1996) repeated the oft-quoted rubric of the military that applies well to aviation activities: "Measure it with a micrometer, mark it with a piece of chalk and cut it with an ax" (p. 1).

To this has often been added, "Beat it to fit and paint it to match."

McBride used this rubric to introduce his discussion of significant numbers in mathematics, stating, "… every time we perform a mathematical operation on a measured value we must maintain the integrity of the measurement's accuracy" (p. 1).

Nothing could be truer for aviators who often need quick results but do not need answers calculated to a high degree of mathematical precision.

In designing equipment and systems aviation engineers must be mindful of accumulation of tolerances. The pilot, on the other hand, most often does not have time or data that is sufficiently accurate or reliable for precise math to be meaningful. Thus the genesis of aviation "Rules of Thumb."

It has been my practice in certain classes to introduce student pilots to these rules of thumb, while at the same time expecting them to know how accurate they are. This article will discuss and evaluate some of these rules of thumb and their accuracy. It is not the purpose of this article to provide significant breakthroughs in aviation science or engineering. "Rules of Thumb" are, after all, just that – rules established by pilots for easier management of aviation

---

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

operations. It also is not the purpose of this article to expect pilots to use the more precise math discussed. Rather the article is intended to show how reliable these rules are and how comfortable pilots should be in using them.[1]

The results of the evaluation were nevertheless interesting. Analysis of the initial bank angle rules, for example, demonstrated the preference for the 15 percent rule over others commonly in use. Similarly, the analysis of the rule for wind correction in holding patterns showed the justification for the change in the Aeronautical Information Manual (U.S. Department of Transportation, 2002–2003) implemented by the FAA several years ago.

Sources of Rules

Two of my favorite sources are *Aviation Rules of Thumb* (n.d.) and the quizzes located under the title "Brainteasers" on the website entitled AVweb (2003)[2]. The analysis of this article is based on the questions contained in one of the quizzes in the latter reference.

Analysis of the rules

Arc Length

AVweb's (2003) first question asked, "You are intending to fly the 20 DME arc from XYZ VOR radial 184 to radial 214. Your groundspeed is 120 knots. Approximately how much time will be required?" The answer given was 5 minutes and explained as follows:

> "At a distance [of] 60 nautical miles, one degree of arc equals one nautical mile. From the 184 radial to the 214 radial is 30 degrees of arc. That would be a distance of 30 nautical miles on the 60 DME arc, but on the 20 DME arc it is one-third of that, or ten nautical miles. At a groundspeed of 120 knots it takes five minutes to fly ten nautical miles."

How accurate is this answer? First, how accurate is the statement that one degree of arc at 60 nautical miles is equal to one nautical mile?

This statement can be easily tested by examining the following formula $d = \frac{1}{360} \times 2\pi \times r$ where d is the arc distance and r is the radius of the arc. Thus 1-360th of the circumference of a circle with a radius of "r" is the distance along the arc. Therefore, $d = \frac{1}{360} \times 2\pi \times 60 = \frac{1}{6} \times 2\pi \cong \frac{6.28}{6} \cong 1.047 nm$

---

[1] This article is intended to examine the action of the kind of small aircraft used in training and not larger commercial aircraft that have the use of flight directors and Flight Management Systems. For that reason, the analysis is limited to aircraft operating at less than 150 KIAS and at constant rate turns of 3 degrees per second.

[2] This website contains many other interesting articles and resources.

The arc distance is off by a bit more than 4 per cent. The total distance over the 30 degrees in the question but at 20 nm DME would be $\frac{20}{60}\times 30\times 1.047=10.47nm$

At 2 nm per minute, the time to travel this distance would actually be $t=\frac{10.47}{2}=5.23$ minutes. The time, like the distance, is off by about 4.7 percent.

Headwind and Crosswind

AVweb's (2003) second question was, "You are approaching to land on a runway where the wind is blowing at 15 knots at a 40-degree angle to the runway heading. What is the approximate crosswind component?" The answer given by AVweb was, "At a 40-degree angle, the crosswind component is between 60 percent of the wind speed and two-thirds of the wind speed. Sixty percent of 15 knots is 9 knots. Two-thirds of 15 knots is 10 knots. We're giving credit for either answer".

The actual crosswind component given by the formula: $V_{xw}=V_w\sin\theta$
Where V_w is the velocity of the wind, V_{xw} is the velocity of the crosswind component, and θ is the angle of the wind to the nose. Thus at 40°, the crosswind component is: $V_{xv}=15\sin 40^\circ=15\times 0.643=9.64$ knots. The answer actually is closer to 10 than to nine.

In navigation classes, the author teaches that at 45° from the nose or the tail, both the crosswind and the headwind/tail wind are equal to 70% of the wind velocity. For each 15 degrees toward the head (tail), the headwind (tailwind) increases by 20 percent and the crosswind decreases by 20 percent. Similarly, for each 15 degrees toward the wing, the headwind (tailwind) decreases by 20 percent and the crosswind increases by 20 percent. Thus for a 10 knot wind 30 degrees off the nose, the headwind component is 70%+20%=90% of 10 knots or 9 knots. The cross wind is 70%-20% or 50% of 10 knots or 5 knots. Similarly, a wind at 10 knots that is 60 degrees off the tail will provide a 70%-20% or 50% of 10 knots of tailwind and 70%+20% or 90% of 10 knots or a 9 knot crosswind.

These are not exact values since the sine values are 30°, 45° and 60° are 0.500, 0.707 and 0.866 and the cosine values are 0.866, 0.707, and 0.500, respectively. However, the reliability of the wind figures is not accurate either; therefore, the estimates that are made based on this method are as close as anyone can reasonably predict.

Bank Angle

AVweb's (2003) next question was, "What approximate bank angle is required for a standard rate turn at a true airspeed of 130 knots?" The answer given is, "To calculate the approximate bank angle for a standard rate turn, divide the true airspeed by 10 and add seven. One hundred thirty divided by ten is 13. 13 plus 7 is 20."

Evaluation of this question is more complex and requires a basic understanding of the dynamics of the level turn. Figure 1 is a simplified diagram of the forces acting on an aircraft.

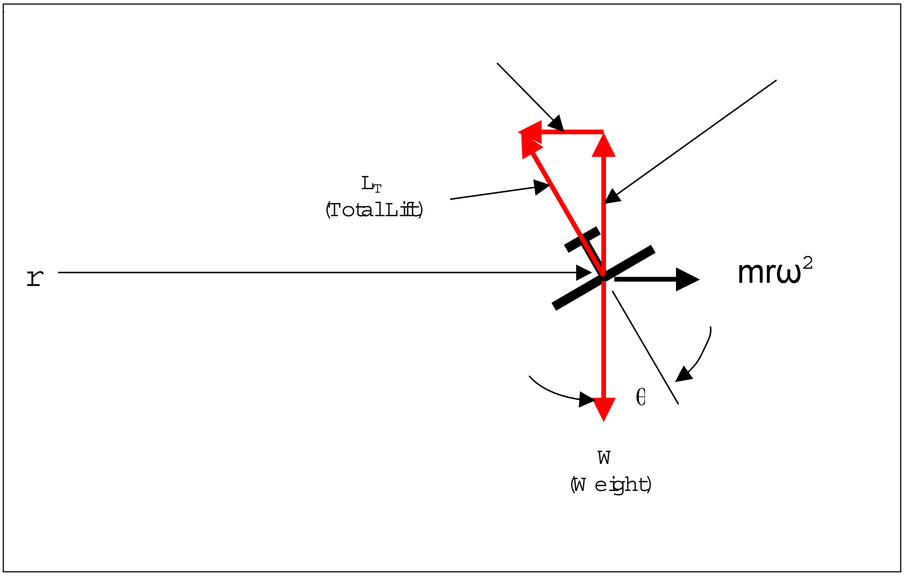


Figure 1. Forces acting on an aircraft in a level turn.

From the diagram, it can be seen that the outward "centrifugal force" is equal to the inward horizontal component of lift. Thus, $L_H = mr\omega$.

Where L_H is the horizontal component of lift, m is the mass of the aircraft, r is the radius of the turn, and ω is the angular velocity of the aircraft. However, since $L_H = L\sin\theta$ and $L = \frac{W}{\cos\theta}$, then $L_H = \frac{W\sin\theta}{\cos\theta} = W\tan\theta = mg\tan\theta$. Equating this to the centrifugal force, $mg\tan\theta = mr\omega^2$. The masses on each side of the equation cancel, leaving $\tan\theta = \frac{r\omega^2}{g}$. For aviation purposes ω is constant at 3 degrees per second for a standard rate turn. Also, in physics we learned that $V = r\omega$ and thus, $r = \frac{V}{\omega}$. Substituting this into the equation above, $\tan\theta = \frac{V\omega}{g}$. The equation is completed by including the conversion factors.

If V is in knots, it must be converted to feet per second. Since there are 6076 feet in a nautical mile and 3600 seconds in an hour, the conversion is $V_{f/s} = \frac{6076}{3600}V_{kts} = 1.6878V_{kts}$ feet per second. ω must also be converted from degrees per second to radians per second. There are 2π radians in a full circle so the conversion from 3° per second is $\omega_{rad} = \frac{3}{360}\times 2\pi = \frac{1}{60}\pi = 0.05236$ radians per second. g is the acceleration of gravity or 32.2 feet per second squared.

Substituting these values into the equation results in $\tan\theta = \frac{1.688\times 0.0524}{32.2}V_{kts} = 0.00274V_{kts}$. If, for example, the aircraft is traveling at 120 knots, the formula would predict an initial bank angle of 18.22°. Compare this to the predicted bank angle of 18° based on 15 percent of airspeed. In fact, the closeness of the mathematically accurate angle is very close to the 15% rule of thumb angle. The 10% plus 7 rule is less accurate. Figure 2 below shows how close these angles are to the mathematical solution for small aircraft.

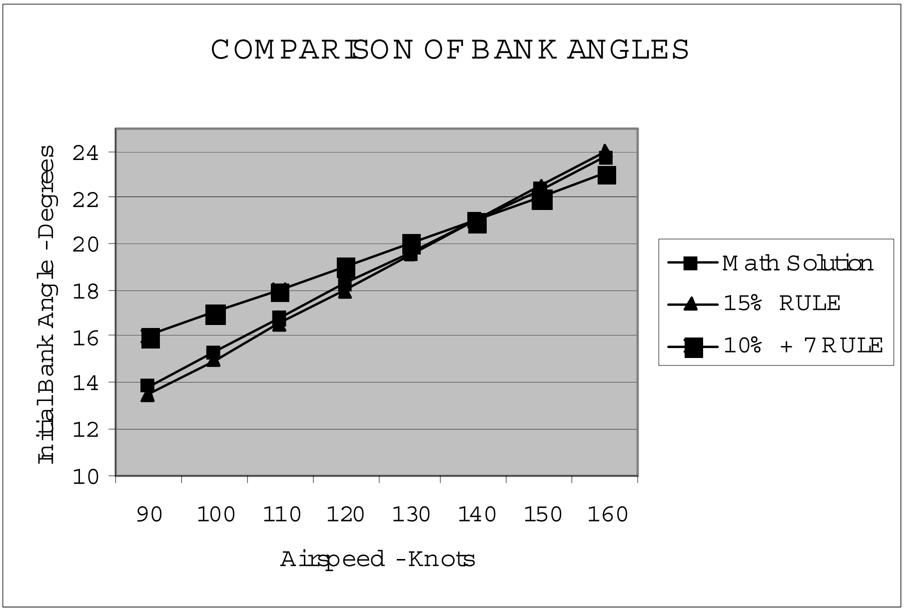


Figure 2. Comparison of computed bank angles for small aircraft.

Correcting for Crosswinds in a Holding Pattern

Probably the least likely candidate for scientific and mathematical analysis is the holding pattern. Wind direction and speed predictions are notoriously inaccurate and the ability to replicate the pattern exactly each time is nearly impossible, even with an autopilot or a flight management system. Nevertheless, Air Traffic Control (ATC) must set aside protected airspace for holds based on a number of conditions.

Under FAA Order 7130.3A (U.S. Department of Transportation, Federal Aviation Administration, 1998), ATC will choose from a list of 31 templates to be used for establishing the protected area of a particular hold. The controller must take into account the following criteria for holds in each instance:

a. Indicated airspeed of the aircraft.
b. Navigation aid and navigation system reliability.
c. Wind effect.
d. Longest distance to any of the navaids being used to establish the holding fix.
e. Aircraft altitude.

Figure 3 below is a typical example of a holding area template as depicted in FAA Order 7130.3A (1998). Typically, instructors refer to the "protected area" when discussing holding procedures with students. The reader should understand that the so-called "protected area" may be different when considering the use by ATC for separation purposes and when used by airspace designers for terrain avoidance purposes. The templates referred to in FAA Order 7130.3A are primarily intended to be used for separation purposes.

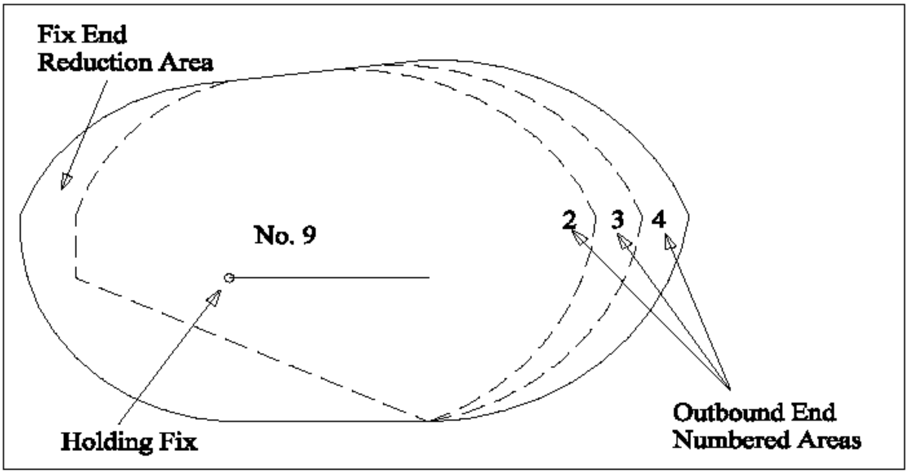


Figure 3. Typical Holding Pattern Template.[3]

Many instructors are familiar with the rules of thumb and the practices recommended by the Aeronautical Information Manual (AIM) (U.S. Department of Transportation, 2002–2003). For example, the AIM provides the following with respect to drift correction:

> "Compensate for wind effect primarily by drift correction on the inbound and outbound legs. When outbound, triple the inbound drift correction to avoid major turning adjustments; e.g., if correcting left by 8 degrees

---

[3] The reduction areas are not discussed in this article but can be used by air traffic controllers under certain circumstances. The reader is referred to FAA Order 7130.3A for further discussion of the use of reduction areas.

> when inbound, correct by 24 degrees when outbound" (Section 5-3-7, I, 6c).[4]

This rule is primarily designed for pilots of small aircraft who are not using flight directors or other computerized means of performing holding patterns. Moreover, the use of the standard 3° bank angle suggests aircraft operating at fairly low airspeeds.[5]

In a non-scientific poll of a number of experienced instrument instructors, the author found that there was no universal acceptance of this rule. Some instructors reported that they prefer to use the rule of twice the inbound correction on the outbound leg. Analysis of the procedure will show why this may be the case.

In order to analyze this rule, consider the following:

In Figure 4(a), the holding pattern is shown as holding north on the 360 degree radial of an imaginary VOR to a fix arbitrarily called FIXXX. Assume that the navaid defining the fix provides accurate positional definition on the inbound course and that the hold calls for right hand turns. Assume also, for the purpose of this analysis, that the aircraft is cruising at an airspeed of 105 knots and that the wind direction is from 045° at 15 knots.

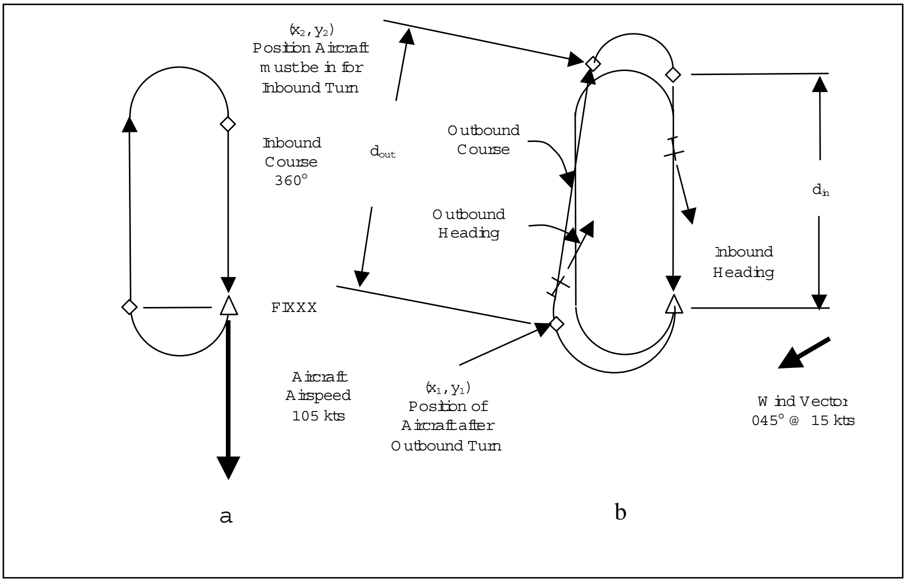


Figure 4. Effect of Wind Correction on Holding Pattern (Not to Scale).

---

[4] This has not always been the rule. In earlier versions of the AIM the recommended ratio was 2:1. Figure 12-6 of the Instrument Flying Handbook, FAA AC61-27C (Revised 1980) is an example of the use of the ratio 2:1.

[5] See the discussion above regarding the initial bank angle that would be required for high speed aircraft to use 3° per second as the rate of turn in a hold.

Figure 4(a) represents the standard holding pattern. The inbound course is located on the navaid radial, in this case the 360° radial, and the inbound distance is arranged so that at the ground speed of the aircraft the inbound time to the fix is one minute.

Figure 4(b) shows the effect of wind. Since the aircraft can be considered to be turning in a moving air mass, the position of the aircraft making the turn can be summed with the movement of the air mass.[6] Thus the position of the aircraft, when it completes the outbound turn, is, in the example, moved south and west of the position it would have been in had there been no wind. The distance the aircraft has moved laterally due to the turn will depend on the radius of the turn. As shown above under the discussion of Bank Angle, the radius is related to the airspeed and rate of turn by the formula: $r = \frac{V}{\omega}$, where V is the true airspeed of the aircraft (not the groundspeed) and ù is the angular velocity, or rate of turn, in radians per second.

In order to keep the radius in nautical miles, the velocity must be converted from knots (nautical miles per hour) to nautical miles per second and the rate of turn must be converted from degrees per second to radians per second. The resulting equation is: $r = 0.005236\text{V}$.

If the aircraft were to complete a heading change of 180°, then, the aircraft would have moved one diameter (or two radii) perpendicular to its movement through the air mass. The ending position of the turn could then be summed with the movement of the air during the one minute it takes to make the turn. This assumes that the aircraft has turned exactly 180°, which the diagram in Figure 3(b) shows not to be true for either the outbound or the inbound turn. Evaluation of the error that this introduces must be considered in deciding the actual relationship between outbound heading and inbound heading.

The position of the aircraft as it completes the turn inbound must be moved to account for the y-component of the wind vector as it effects the groundspeed of the aircraft so that the time from the end of the turn to the fix is still one minute. As a result, the position of the aircraft as it enters the inbound turn can be determined from the position of the aircraft at the end of the turn. The outbound course is determined by connecting the position of the aircraft at the end of the outbound turn to its position at the beginning of the inbound turn. The distance of this leg is determined by the formula $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$, where the (x_1, y_1) is the position of the aircraft at the end of the outbound turn and (x_2, y_2) is the position of the aircraft at the beginning of the inbound turn. Similarly, the angle to the 360° course line is determined by the formula $\tan^{-1}\frac{(x_2 - x_1)}{(y_2 - y_1)}$, where

---

[6] This assumption is based on the principle of linear superposition which, while not proved in this discussion, nevertheless applies.

the positions are the same as in the distance determination.

The actual headings to be used for the inbound course and the outbound course can be computed using standard methods (Trippett, 2002)[7]. In the example the inbound heading, after correction for the fact that the aircraft turns more than 180° outbound and less than 180° inbound will be 174.2°, or a wind correction angle of 5.8°. The outbound course can be determined to be 8.84°. Correcting for the wind yields an outbound heading of 13.6°. The ratio of the outbound correction relative to the reciprocal of the inbound course is thus 13.6/5.8 = 2.36, which is significantly less than the factor of three recommended by the AIM.

Correcting the error introduced by the actual number of degrees the aircraft turns inbound and outbound is accomplished by using a process of successive approximations. The first approximation assumes that the aircraft turns 180 degrees outbound and 180 degrees inbound and that the position of the aircraft at the end of the outbound turn and at the beginning of the inbound turn are based only on the movement caused by the wind. The outbound course in computed from this data and the actual value of the outbound and inbound turns recalculated. The process is repeated until the successive recalculations between one iteration and the next are small enough to be ignored. In the example, the differences are so small after the third recalculation as to be regarded as insignificant.[8]

Figure 5 shows the relationship between wind direction, airspeed/wind speed ratio and outbound to inbound wind correction ratio.[9] The chart confirms what most pilots would immediately realize – that the ratio is affected by how fast the aircraft is traveling in comparison to the wind velocity.

---

[7] The formulas (Trippett, 2002) used by the author are as follows:

Wind Correction Angle:

$$\alpha = -\sin^{-1}[\frac{V_w}{V_{TAS}} \cdot \sin(\phi + \beta - (\theta + 180))]$$

Where

α = Wind correction angle

ϕ = Wind direction (true)

β = Magnetic variation

θ = Course (magnetic)

V_w = Wind velocity

V_{TAS} = True Airspeed

Ground Speed:

$$V_{gs} = V_W \cdot \cos(\phi + \beta - (\theta + 180)) + V_{TAS} \cdot \cos[\sin^{-1}(\frac{V_W}{V_{TAS}} \cdot \sin(\phi + \beta - (\theta + 180)))]$$

[8] After the third recalculation the maximum percentage difference between it and the previous calculation is less than 0.1%.

[9] The data for this chart can be obtained from the author. Tests using the predicted data were run on using the On Top™ flight Simulator. The predicted values agree closely with these tests.

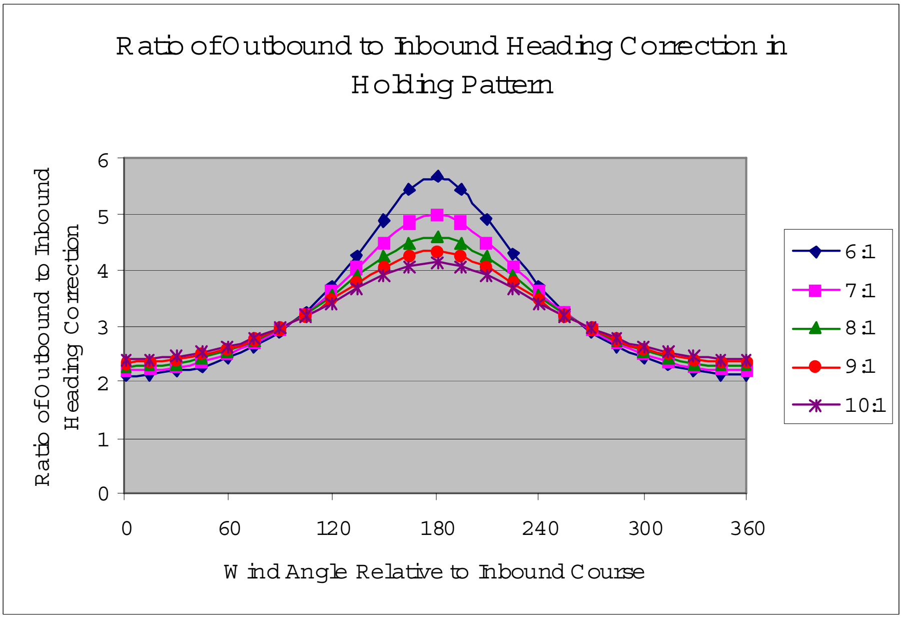


Figure 5. Comparison of Inbound to Outbound Wind Corrections in Standard Holding Patterns.[10]

The variables involved in holding patterns are so numerous that applying them cannot be accomplished by the pilot with the mathematical precision described in this article. Nevertheless, the choice of the 3:1 rule for outbound to inbound wind corrections can easily be seen from Figure 5 as an approximation of the potential values that will be encountered depending on wind direction and the ratio of airspeed to wind speed. The choice of a ratio of 3:1 is within the middle ground (and close to the average) of the expected range of values.

But is it the best choice?

Figure 5 shows the correlation between the ratio of the outbound wind correction and the inbound wind correction as compared to the rule-of-thumb ratio of 3:1. From this figure it can be seen that for winds that tend to have a tailwind character, the correlation is better than for winds that tend to have a headwind character, depending on the ratio of the airspeed to the wind velocity.

These data suggested that where the wind tends to be from the tail, the pilot should use a ratio on outbound to inbound heading correction that is 3:1 or less, tending toward 2:1. However, where the wind tends to be from the nose of the

---

[10] Note that while the ratio of drift correction at 180° and 360° is meaningless since it would require the division of zero by zero, it is assumed without proof that the functions represented by Figure 4 could be shown as continuous on the intervals $0° < \phi < 180°$ and $180° < \phi < 360°$ and that Figure 4 approximates the limit as the $\Delta\phi$ approaches zero at 180° and 360°.

aircraft, the ratio should be more than 3:1, tending toward a ratio of 4:1 or greater, depending on the comparative value of the wind velocity to the aircraft airspeed and the actual angle of the wind to the inbound course.

Conclusion

There are literally dozens of rules of thumb used by aviators to approximate solutions to problems that would otherwise be overwhelming in the cockpit if the mathematical niceties described herein were required for flight. The use of those rules is well justified, at least for the rules described in this article.

In succeeding articles, the author proposes to discuss other rules commonly used by aviators with the similar purpose of determining how correct they are.

References

Brainteasers (2003). *The Internet Aviation Magazine and News Service. http://www.avweb.com/brain/.* Aviation Publishing Group.

McBride, J. R. (1996, September 2). *Statistics.* http://www.geocities.com/CapeCanaveral/Lab/3975/Chem105/105lec1.htm.

Rules of Thumb (n.d.). http://www.geocities.com/acornil/rules_thumb.htm.

Trippet, W. R. (2002). Mathematical solutions to wind correction and groundspeed problems. *International Journal of Applied Aviation Studies (2)*1. Oklahoma City, OK: FAA Academy.

U.S. Department of Transportation, Federal Aviation Administration (1998, March 18). *FAA Order 7130.3A, Holding Pattern Criteria*.

U.S. Department of Transportation, Federal Aviation Administration (2002, February 21 – Changed 2003, February 20). *Aeronautical Information Manual, Official Guide to Basic Flight Information and ATC Procedures*.

International Journal of Applied Aviation Studies, 3,1

A Bright Day Dawns in International Air Carrier Liability: The New Montreal Convention

Robert F. Ripley
Auburn University
Department of Aviation Management
Lowder Business Building, Suite 446
Auburn, AL 36849-5248
Phone: (334) 844-6845
FAX: (334) 844-4927
Email: ripleyf@business.auburn.edu

Abstract

This paper presents the genesis of international aviation law and discusses the provisions of the 1999 Montreal Convention, which amended earlier international air carrier agreements.

In the early days of international air travel many far-sighted individuals recognized that commercial aviation had enormous potential; but unless the liability issues concerning the loss of passengers and cargo were resolved, this potential would never be fully realized. As a result of these concerns, the Warsaw Convention was drafted in 1929 and ratified in 1933. The Warsaw Convention protected the rights of passengers by requiring the payment of damages while at the same time capping liability limits for the airlines that caused their loss. In 1966 a meeting was held in Montreal that became known as the Montreal Convention and a bilateral agreement was consummated that effectively increased the damages awarded to passengers who suffered losses as the

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

result of an accident. The 1999 Montreal Convention further amended the damage awards to persons who had suffered losses as the result of an accident. The provisions of the 1999 Montreal Convention are much fairer to the consumer and thus will result in less litigation for the loss of life, baggage and cargo.

Background

When international air travel first began in the early part of the twentieth century, the need for a universal standard of liability was essential. The politicians and leaders of the aviation industry knew this and it lead them to create a doctrine that would establish a standardized set of regulations governing liability in the event of an accident or loss of cargo. The act was entitled *Convention for the Unification of Certain Rules Relating to International Carriage by Air* (ICAO Doc. 7838) and was later referred to as the Warsaw Convention. This treaty was drafted in 1929 and ratified in 1933. The following year the United States became a signatory (Kreindler, 1999). It was designed to create a uniform standard of liability for international air travel in the event of an airline crash and to encourage the availability of liability insurance for international airlines. The main stipulations of the treaty were designed to limit the monetary compensation paid by the airlines to the parties involved and to set rules on jurisdiction for the litigation following the accident. The Warsaw Convention protected the rights of passengers by requiring the airlines to pay money for damages caused by an accident or loss of cargo, and the convention's limits on liability ensured that the airlines would not suffer a catastrophic loss that would result in the bankruptcy of the airline after an accident. The Warsaw Convention described when, why and how an air carrier would be held liable for an accident or loss of cargo.

The foundation for liability under the Warsaw Convention was stipulated in Article 17 (Kreindler, 1999). This article held the airlines liable for accidents resulting in injury or death, either on the aircraft or in the course of embarking or disembarking the aircraft. This liability was limited under Article 22, and in the original doctrine it was limited to approximately $8,300 U.S. per passenger in cases involving injury or death. Article 25 allowed this limitation to be lifted if the plaintiff can prove willful misconduct. An airline's actions constitutes willful misconduct when an act is committed with an intentional or reckless disregard for the safety of others or with an intentional disregard of duty necessary to the safety of another's property. This charge was very difficult to prove in court, and the liability limit was almost never lifted. Registered cargo and baggage liability was limited to approximately $16.50 U.S. per kilogram for loss or damage. These articles provided the monetary limits and the restrictions on jurisdiction were addressed separately.

The location in which a plaintiff could claim jurisdiction for his impending litigation was limited under Article 28 (ICAO Doc. 7838). There were four specific forums established where a victim could file suit against a carrier. These venues included the following:

a. The country in which the airline is incorporated
b. The country in which the airline is headquartered
c. The country where the ticket was purchased
d. The passenger's ultimate destination

In the past seventy-five years there have been numerous attempts to reform this document to keep up with the changes in the world's economy.

Since the drafting of the Warsaw Convention, several additional countries have been added as members. These nations contributed to the formation of a governing body over civil international aviation called the International Civil Aviation Organization (ICAO). Liability limits also have increased through various protocols and amendments to the original doctrine.

The first meeting to update the Warsaw Convention was held in the Hague in 1955 and hence has been referred to as the Hague Protocol. The limits governing death or injury were raised to approximately $16,000 U.S., but the baggage and cargo limits were unchanged. The wording regarding exoneration from limited liability was altered slightly. However, these limits were low even for the 1950s, and it was still almost impossible to prove that any conduct by the carrier or one of its agents was considered intentional and enough to justify lifting the restrictions. Consequently, the need to revise this amendment was evident.

The United States was adamant about restructuring these liability limits. They threatened to pull out of the Warsaw Convention entirely and protested by not ratifying the Hague Protocol. In 1966 ICAO member nations met in Montreal and held the first of its conferences to amend the Warsaw Convention. The settlement reached there was not a protocol, but a bilateral agreement between air carriers operating in the U.S. and the United States Government. The limit for death or injury of a passenger who was a United States citizen was raised to $75,000 U.S. Even though these limits were enforced, most nations felt the Warsaw Convention needed to be updated further to meet the demands of the current economy.

The United States certainly was not the only country upset with the low monetary liability limits. Therefore, in 1971 in Guatemala, the ICAO tried to increase limits to approximately $100,000 U.S. for a passenger's injury or death and approximately $1,000 U.S. for loss of cargo, but member nations showed their disapproval with the agreement by not giving the support it needed to be ratified. The document also met opposition because it attempted to change the wording of Article 17. The delegates at the convention replaced the word "accident" with "event" in the article's description of what constitutes an accident. This virtually would have abolished the need for the accident to be aviation-related. The specific monetary limits were not the only parts of the Warsaw Convention that were out of date. The protocols pertaining to and including the original Warsaw Convention had used a French currency called the Poincare

Franc, which had not been used by the French since the late 1930s.

The global unit of exchange for currencies was introduced at the Bretton Woods Conference in 1944 (Batra, 2000). However, it was not used in the context of the Warsaw Convention until after the four protocols drafted in Montreal in 1975. The meetings produced an agreement that incorporated a Special Drawing Right (SDR) into monetary limits of the Warsaw Convention. Derived from the International Monetary Fund, a Special Drawing Right acts as a unit of exchange for all currencies including the Poincare Franc.

Another successfully passed protocol dealt with passengers attempting to gain standing under local laws.

In March of 1999, the protocol drafted at the 1975 Montreal Convention pertaining to a passenger's right to claim damages under local law came into effect in the United States. This protocol related to Article 24, and it prevented any passenger from claiming damages that resulted in death, bodily injury, or loss of baggage after an accident under local law, as it pertained to Article 17 of the Warsaw Convention. This has been enforced retroactively in some cases, and the 1999 Montreal Convention used the same wording. The issues surrounding the limits on liability also were addressed in these protocols.

One of the protocols from the 1975 Montreal Convention raised the monetary limits for the airlines in the event of an accident or loss of cargo. At this conference the International Civil Aviation Organization attempted to raise the liability limits for the airlines to about $130,000 U.S., but the protocol did not allow for the limits to be lifted in situations where the accident was fault based. Therefore, these restrictions put unfair burdens on airplane manufactures that could be held accountable for additional recoveries. This protocol was not passed by several nations, including the United States, because of its weak stance on any form of unlimited liability against the airlines.

There always have been complaints from in and outside the aviation industry concerning the low limits established by these protocols. Japan and Italy are examples of two countries that took action to increase liability for their airlines and, consequently, they took measures to make their airlines fully liable for both international and domestic travel. These actions, the inability to pass some additional protocols, and the wide variations in cost of living conditions motivated the ICAO to conduct a study to analyze the problems with the current limited liability standards.

This study began in 1994 with the assistance of the International Air Transport Association. The main initiative of the study was to determine the adequacy of the current liability limits and of proposed limits, as well as the costs associated with providing higher limits would have for all carriers. The study group consisted of private international air law experts from around the world, but the vast majority of these specialists represented developed nations. They distributed questionnaires that inquired about each member nation's views toward the adequacy of current liability limits for passenger, cargo, and baggage. Seventy-two percent of the International Civil Aviation Organization States that responded were dissatisfied with the current regulations (Batra, 2000). The study determined that the increase in ticket prices would be less than two dollars round trip to cover the added insurance expense. The group found that the current liability limits on international air travel were not acceptable worldwide because of the diversity of socioeconomic circumstances and variance in the cost of living. Therefore, they issued steps on how they believed the situation should be solved, and many of their ideas represented the foundation for the present agreement.

The ICAO council supported these recommendations and forwarded them to a legal committee. The council later organized a Special Group for the modernization and consolidation of the Warsaw System to supplement the work of the legal committee. It was in early 1996 when the ICAO council gave its initial endorsement of the study group's proposals concerning the need to draft a new agreement and recommendations on what types of changes needed to be made. The convention in Montreal to officially draft the treaty was held from May 10-28, 1999 and was referred to as the Montreal Convention of 1999.

1999 Montreal Convention

The meeting in Montreal included representatives from the majority of the ICAO member nations, and all countries in attendance signed the document, which was entitled the *Convention for the Unification of Certain Rules for International Carriage by Air* (ICAO Doc. 9740). However, only twenty-five nations (as of November 20, 2002) have ratified it, and according to Article 53, the agreement will not come into effect until sixty days after the thirtieth acceptance, approval, accession or ratification (Hamilton, 2001; Weber, 2000).

This document encompassed the work done by everyone involved in the process from the formation of the study group to the legal committee and all other parties mentioned earlier. The end product was to be the stepping-stone for the future of international air carrier liability.

The *Convention for the Unification of Certain Rules for International Carriage by Air* contained several sections and articles. The ones discussed in this report were some of the most influential, and they included a retention of the structure of the Warsaw Convention, a two-tier liability structure, limits of liability for delay, limits for loss of baggage and cargo, exoneration of liability, review of limits and advanced payments, an optional clause, jurisdiction, reimbursement for legal fees, recovery of non-compensatory damages, ticket and flight information, proof of insurance, and the statute of limitations (Kreindler, 1999; ICAO Doc. 9740; Weber, 2000). These categories are outlined as follows:

Retention of Structure

The preamble of the new policy acknowledged the importance of maintaining the structure of the original assembly held in Warsaw in 1929 and ensuing protocols. Therefore, the preamble recognized the need to modernize the Warsaw Convention, and to protect the consumer's best interest through equitable compensation. However,Article 55 stated that this convention superceded the Warsaw Convention and all of its protocols, and all specific intercarrier agreements. That article established the 1999 Montreal Convention as a completely new agreement.

Two-Tier Liability System

Article 21 stipulated that, for the death or injury of a passenger, compensation will be strictly payable by the airliners to passengers not exceeding 100,000 Special Drawing Rights (approx. $135,000 U.S.). The airline could be held accountable for unlimited liability in cases where the passenger can prove that the airline or one of its agents acted with negligence. This condition forced an airline to prove that it took all necessary procedures to avoid damage, or that those measures were not feasible. This defense is almost impossible to prove in court, and the airlines failure to do so would result in the airlines being held liable without limitation for all provable damages.

Liability Limits Due to Delay, Baggage and Cargo

Also included in Article 21 were the restrictions that limit payment by the air carriers to 4,150 SDR (approx. $5,600 U.S.) for damaged caused by delay. For baggage or cargo loss, delay or damage the liability was limited to 1,000 SDR (approx. $1,350 U.S.) per person. However, a passenger can make a special declaration at the time the baggage is checked. The passenger then will be quoted a fee, and if paid, the carrier becomes liable for the declared sum, unless the carrier can prove that the sum declared is greater than the passengers' real interest in the delivery.

Exoneration of Liability

According to Article 20, the carrier has the right to attempt to prove that the person claiming compensation was the cause of the accident or a direct contributor through negligence. By accomplishing this, the airline can be relieved, wholly or partly, from its liability.

Review of Limits and Advanced Payments

Built-in to require that the liability limits be reviewed every five years to adjust to inflation and current economic conditions, the escalation clause was found in Article 28. These renewal meetings will begin five years after the agreement has been put into force. The rate of inflation factor will be weighted with the average of the annual rate of change in the Consumer Price Indices of the currencies that make up the SDR. The possibility that the airline may have to make advanced payments for liability after an accident according to national law was included in Article 28, and if such payments are made it will not constitute a recognition of liability, and may be offset against payments made by the carrier for future damages.

Optional Clause

The air carrier may at its discretion set its own liability limits to be higher or unlimited.

Jurisdiction

In Article 33, a fifth forum was added to give passengers more control over location when they are involved with litigation against the airline. The fifth forum allowed for the passenger to seek action resulting from death or injury in the territory or state in which at the time of the accident the passenger has his/her principal or permanent residence or from which the carrier operates service for the carriage by air, either on its own aircraft or another carrier's aircraft pursuant to a commercial agreement. The four original forums remain in place.

Passenger Contributory Fault

Article 20 established that, if the airline can prove that passengers claiming compensation contributed to or caused the damage by their negligence or other wrongful act or omission, the carrier can be fully or partly exonerated from their liability. However, this is seldom appropriate in any case because most passengers do not contribute to major airline crashes. In the 1929 Warsaw Convention this was to be exclusively handled according to forum law, but the new agreement does not explicitly turn the matter over to forum law. Although, in the event that this article becomes relevant, it is presupposed that forum law would apply in an applicable situation.

Reimbursement for Legal Fees

The court may award the plaintiff compensation for legal expenses and court costs if permitted under forum law, and only if it pertains to a case that involved the plaintiff recovering an amount that was more than the carrier's written offer that was made within six weeks of the accident or an offer made before the commencement of the suit. This rule occurred in Article 22.

Recovery of Non-Compensatory Damages

The recovery of any exemplary, punitive damages, or any type of non-compensatory damages were strictly not attainable, under Article 29. This issue was not addressed in the 1929 Warsaw Convention, and the United States Federal Court system had previously ruled that Article 17 only provided for compensatory damages.

Ticket and Flight Information

Article 3 no longer invoked any sanctions for failure to deliver a ticket or preserve the flight information. This delivery of a ticket was an important factor for the limited liability according to the 1929 Warsaw Convention.

Proof of Insurance

In Article 50, a mandatory insurance clause stipulated that any carrier must surrender proof that they have adequate insurance to any state party for any country in which they operate.

Statute of Limitations

The two-year statute of limitation was retained in the 1999 Montreal Convention.

Table 1
Changes in Liability Limits

Conference	Warsaw	Hague	Montreal	Guatemala	Montreal	Montreal	Montreal
	1929	1955	1966	1971*	1975	1975*	1999*
Proposed liability limits for death or injury in U.S. dollars**	$8,300	$16,000	$75,000	$100,000	SDRs Applied	$130,000	$135,000
Proposed liability limits for loss of cargo in U.S. dollars**	$16.50/ kg	$16.50/ kg	No Change	$1,000/ Person	SDRs Applied	No Change	$1,350/ Person

* Not Ratified as of 5/01/01
** Approx. U.S. dollar values

Contracting Nations

As of November 20, 2002, twenty-five nations deposited an instrument of ratification, acceptance, approval or accession as required by the terms of the Convention (http://www.icao.int/icao/en/leb/mt199.htm). These nations and the date of the deposit of the required instrument are as follows:

Belize 08/24/99
CzechRepublic11/16/00
Greece 07/22/02
Kenya 01/07/02
Kuwait 06/11/02
Mexico 11/20/00
Namibia 09/27/01
Japan 06/20/00
Panama 09/13/02
Slovakia 10/11/00
Slovenia 03/27/02
Peru 04/11/02
Romania 03/20/01
Paraguay 03/29/01
Barbados 01/02/02
Nigeria 05/10/02
United Arab Emirates 07/07/00
Jordan 04/12/02
Bahrain 02/02/01
Botswana 03/28/01
New Zealand 11/18/02
Canada 11/19/02
Cyprus 11/20/02
Syrian Arab Republic 07/18/02
The former Yugoslav and Republic of Macedonia 05/15/00

Conclusion

The drafting of a new Warsaw Convention was long overdue and the conditions stated in the 1999 agreement were merely preliminary figures established to lead the international aviation industry into the next century. The limits declared in Montreal were low for today's economy and there was plenty of opposition from the nations that represent the International Civil Aviation Organization. However, any limits set by the ICAO will be met by resistance from parties that feel their interests were not addressed properly, yet it is impossible to go from a figure below $100,000 to a limit of over a $1,000,000 overnight. Therefore, if this agreement and future arrangements are going to work, the states that are affected by it must be able to accept the terms agreed upon in Montreal in 1999 for what they are – a stepping stone for the future. If these conditions are accepted as a starting point then the Warsaw Convention modernization plan will be able to grow every five years in accordance with the stipulations stated in the escalation clause.

It is important to note that when the 1999 Montreal Convention provisions go into effect, the consumer will benefit greatly. Specifically, there will be little, if any, litigation by passengers to recover full compensatory damages for injury or death; few, if any, attempts by airlines to avoid paying compensatory damages above 100,000 SDRs, unless they can clearly show that they were not the proximate cause of the accident; and, no litigation over the destruction, damage or loss of cargo.

References

Batra, J. C. (2000, Summer). Modernization of the Warsaw system–Montreal, 1999. *Journal of air law and commerce 65* (429).

Blacks law dictionary, 6th ed. (1991). West Publishing Company,

Convention for the unification of certain rules relating to international carriage by air, Oct. 12, 1929, 137 L.N.T.S. 11, 49 Stat. 3000, TS No. 876, ICAO Doc.7838 [Warsaw Convention].

Convention for the unification of certain rules relating to international carriage by air (1999, May 28). ICAO Doc. 9740 [Montreal Convention of 1999].

Convention for the unification of certain rules for international carriage by air done at Montreal (1999, May 28). *Current list of countries that have accepted, approved, accessioned to or ratified the1999 convention for the unification of certain rules for international carriage by air*. Retrieved March 2003 from http://www.icao.int/icao/en/leb/mt199.htm.

Hamilton, J. S. (2001). *Practical aviation law*.Ames: Iowa State University Press.

Kreindler, L. (1999, December 30). Millennium revolution. *New York law journal*.

Rodriguez, B. I. (2000, Winter). Recent developments in aviation liability law. *Journal of air law and commerce* 66(21).

Weber, L. J. (2000, February 17 & 18). Airline liability under the new Montreal convention of 1999. *Aviation law and insurance symposium*, Daytona Beach, Florida.

Weigand, T.A. (2001). Accident, exclusivity, and passenger disturbances under the Warsaw convention. *American University international law review, 16* (891).

Author Information:

Robert F. Ripley (JD, Jones School of Law) is an Associate Professor of Aviation Management in the Department of Aviation Management and Logistics, College of Business, Auburn University. He has published in various academic and practitioner publication outlets and is active in several professional organizations. His teaching and research focus is in the areas of Aviation Law and Policy and Aviation Human Factors.

International Journal of Applied Aviation Studies, 3,1

Leadership Values: Are They Industry-Specific? Can They Be Learned and Unlearned?

A Comparison of the Importance of Leadership Values in Generic Leadership Roles Versus Aviation Leadership

David B. Carmichael
1136 Fox Lake Lane
Edmond, Oklahoma 73034-7309
Phone: 405-348-8050
E-Mail Address: dcarmich@okcu.edu

Mary N. Kutz
6108 Winfield Drive
Oklahoma City, OK 73162
Phone: 405-720-9091
E-Mail Address: drkutz@cox.net

and

Dovie M. Brown
16108 Silverado Drive
Edmond, Oklahoma 73013
Phone: 405-348-1388
Email: deebrown1@msn.com

Abstract

Prompted by research conducted in 1998 pertaining to the Characteristics of Successful Aviation Leaders in Oklahoma, *this article examines the literature for basic answers to some fundamental questions pertaining to values in leadership. The Oklahoma study*

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

involved interviews with successful aviation leadership for advice pertaining to education of future aviation leaders. One of the issues that surfaced from the leaders in that study was the importance of teaching values to tomorrow's aviation leaders. This notion raised a number of questions pertaining to values that must be answered by future research.

Are leadership values such as honesty simply inherent to such industries as aviation where life and death decisions often depend upon an industry-wide commitment to honesty? Or are values critical to survival regardless of the industry? At what point can values be unlearned or so stressed by the realities of profits or other industry pressures that they take a back seat to those pressures? Are values teachable and should they be taught either as a part of industry-specific educational curricula or by organizational training modes to achieve a best fit with the organizational culture?

These and other questions and issues explored in this article pertaining to values have far-reaching implications for the education of future leaders, not only in the field of aviation, but also other organizations of our day. Perhaps an understanding of the role of values in a variety of environments could provide valuable insights into the leadership pressures that create leadership breakdowns and failure of not only some of the major air carriers, but also the Enrons, Global Crossings, Arthur Andersens and other contemporary examples of organizational leadership gone awry.

Introduction

A 1998 study of the *Characteristics of Successful Aviation Leaders in Oklahoma* (Kutz, 1998) raised some basic issues pertaining to the importance of values to aviation leadership. The Oklahoma study involved interviews with successful aviation leaders throughout the state for advice pertaining to education of future aviation leaders. One of the issues that surfaced from the leaders in that study was the importance of teaching values to tomorrow's aviation leaders. This notion raised a number of questions pertaining to values that must be answered by future research.

Are leadership values such as honesty simply inherent to such industries as aviation where life and death decisions often depend upon an industry-wide commitment to honesty? Or are values critical to leadership survival regardless of the industry? Are values teachable and should they be taught either as a part of industry-specific educational curricula or by organizational training modes to achieve a best fit with the organizational culture? At what point can values be unlearned or so stressed by the realities of profits or other industry pressures that they take a back seat to those pressures?

A partial answer is found in a similar study entitled, *High Achievers in Fed-*

eral Service in Oklahoma (Carmichael, 1994). In this study, the researcher found that values were a central source of guidance to the subjects. For these individuals, such issues as public service, religious teaching, family, honest communication, work ethic, leadership responsibility, and sensitivity to others were central to their being and to their professional conduct. Only two of the four agencies represented in the studies were aviation related. Integrity was particularly important to the high achievers in the Internal Revenue Service. The value of service to patients was particularly important to personnel from the Veteran's Administration Medical Center.

As a part of their research of generic leadership roles, the authors examined the accounting profession as well as the aviation industry. Accounting, particularly in the area of audit, is especially sensitive to integrity and damage to the reputation of their firms. Even in professions very sensitive to reputation, i.e., accounting, the literature reviewed by the researchers indicates that the profession has been very concerned about audit firms and their clients becoming too intertwined in their interests for audits to be objective.

The field of aviation has the same sensitivity to reputation and integrity with the additional critical issue of human life at stake. Sadly, however, historical evidence indicates that aviation is subject to the same human frailty as the audit profession. The profit motive and the pressures inherent in the environment may well cause serious breaches of ethics, even with human life at stake.

Purpose

Questions pertaining to values have far-reaching implications for the education of future leaders not only in the field of aviation but also other industries of our day. Perhaps an understanding of the role of values in a variety of environments could provide valuable insights into the leadership pressures that create leadership breakdowns and failure of not only some of the major air carriers such as Pan Am, Trans World Airlines, and Value Jet, but also the Enrons, Global Crossings, Arthur Andersens and other contemporary examples of organizational leadership gone awry. This article is only the beginning of a long and involved process of exploring and understanding leadership values and their impact on organizations as well as their implications for the education of future leaders. The purpose of this article is to explore what we know today about the role of values in leadership success and offer insights and recommendations for future exploration.

Values Defined

In order to establish a common understanding of values it is important to define the term as it relates to other terms such as ethics, standards, and principles, all of which describe an aspect of leadership that shapes the organizational culture.

In a contemporary textbook, *Organizational Behavior*, (Robbins, 2001) in defining values, quoted Rokeach as follows:

> "Values represent basic convictions that a specific mode of conduct or end-state of existence is personally or socially preferable to an opposite or converse mode of conduct or end-state of existence."

Robbins continued this thought by pointing out that in Rokeach's definition, values contain a judgmental element that carry an individual's ideas as to what is right, good, or desirable, and that values have both content and intensity variables.

The Department of Defense Directive (DODD) 5500.7 *Standards of Conduct: Joint Ethics Regulation* defined values as "core beliefs such as duty, honor and integrity that motivate attitudes and actions" (p. 37). They further defined ethics as "standards by which one should act based on values or core beliefs" (p. 37). For purposes of this writing, we define values similarly to the Air Force definition as "core beliefs that motivate attitudes and actions." We define ethics as "the standards by which one should act in keeping with core values."

Values in Generic Leadership Roles

Regardless of the industry, the setting of organizational standards and the immersion in corporate values must be communicated or people will draw their own conclusions from their own perceptions of what they see and that can spell disaster for a company. Organizational leadership has the responsibility of modeling values and involving their people in the process of creating, communicating, and modeling shared values.

Organizations and individuals face ethical issues and value judgments in almost every aspect of doing business even in the communications process. Individuals face value judgments on what they should reveal on an employment application to a prospective employer. Many resumes contain untruths and many companies ask discriminatory interview questions, both of which raise ethical issues and require value judgments. Companies make value judgments and ethical decisions about advising consumers of flawed products. In the aviation business as in other businesses, the pressures of a rapidly evolving business environment place ethical demands on employees and management to cut corners to meet unrealistic deadlines that affect productivity.

Many business leaders believe companies with high ethical standards are also the strongest competitors. Internal ethics programs have been implemented in many of these companies as a means of evaluating the moral integrity of business communication practices.

Kouzes and Posner have done some of the most extensive research on leadership and values from the perspective of both leaders and constituents. In their book, The Leadership Challenge (1995), they described their research base of 60,000 leaders and constituents in both public and private organizations worldwide. They conducted their research in a variety of ways beginning with the question "What values (personal traits or characteristics) do you look for and admire in your superiors?" (p. 20) from which they gleaned approximately 225 listed values. Subsequent research involved interviews and actual case studies and stories. Honesty was the value most identified in every survey. They described honesty as the "single most important ingredients in the leader-constituent relationship" (p. 22), and the leader's behavior provided the evidence. If leaders practice the values they talk, constituents are willing to entrust them with their careers and even their lives. They further described the relationship between honesty and values and ethics in that people refuse to follow leaders who lack confidence in their own beliefs and who will not express their values to their constituents (Kouzes & Posner, 1995, p. 20-25).

Values in Aerospace Leadership

Because comments from the interviews with Oklahoma aviation leadership in the 1998 study sparked interest in further research of value-oriented leadership issues, we begin with a look at values in the aviation/aerospace industry specifically and utilize the audit industry as a basis for comparison purposes.

Aviation is the leading industry in the state of Oklahoma. The Oklahoma study included interviews with top-ranked leaders in a variety of specialties within the aviation field ranging from a military organization of over 20,000 employees to leaders of air carrier organizations, airport managers, aviation professional organizations, and government organizations as well as astronauts, national and state political and community leaders in positions of influence in the aviation community. The leaders interviewed were asked to describe their guiding values and principles. Consistent with the findings of Kouzes and Posner (1995) in their study of leadership values, the Kutz study found that honesty was the most frequently mentioned value. Some of the leaders expressed the notion that honesty is inherent in the aviation community in that your life and the lives of others depend upon it. The pilot must depend on the Air Traffic Controller who assigns airspace and the mechanic who signs off on the aircraft to be honest in the performance of their job. Nevertheless, some of the leaders expressed concerns that values should be taught and retaught throughout the educational process.

Ash (2001) described the importance of ethics and values to today's aero space leadership efforts in the United States Air Force.

> Technology changes; operations and tactics change; and we speak of revolutions in war as well as generational differences like baby boomers and Generation Xers. But truth and honesty are timeless, and they are also as fundamental to discipline and military effectiveness as anything else. Herein lies the leadership challenge. Ask any academy commandant if maintain-

> ing the honor code and getting cadets to live according to sound ethics are not among the greatest challenges in producing tomorrow's leaders (p. 33).

Murphy (2000) described the process used by the military to inculcate its values to new recruits:

> In the Air Force, new recruits never see an F-15. Instead, they have the common values and expectations of the Air Force drilled into them until this fundamental understanding is complete. Afterward, all the training, all the evolving standards of performance, and all the awards and successes are built on these first planks in the organizational platform (p. 33).

He later described integrity as one of the core values of the Air Force for promoting ethical conduct. This is consistent with the Kutz (1998) study in which honesty and integrity were identified as core values critical to successful aviation leadership.

As aviation technology of the 20th Century gives way to the complex aerospace technology of the 21st Century, the responsibility for strong values and ethics permeates even the aviation communications process. Accurate use of the language in communicating in the aviation industry becomes more complex and more important. In the October 23, 2000 issue of *Aviation Week & Space Technology*, an editorial by Murphy addressed the ethical responsibility of aviation to help news media accurately report aircraft accidents suggesting that we in aviation do not do enough to help the media determine the facts and put them in context. Rather than have a spokesman with little or no technical expertise attempt to brief the media and leave a lot unsaid so that misinformation fills the vacuum, Murphy suggested establishing "a dialogue with reporters and editors to help them understand the accident investigation process and identify reliable sources of information or comment" (p. 102). Rather than utilizing press leaks in an effort to slant one side of the story so that perceptions of air safety become a by-product of miscommunication, the aviation and aerospace industry can take the high road even in their communications to the public.

Aviation is a proud industry responsible for much of the technological and economic growth as well as societal development in this country. It has a proud history based on the highest standards grounded in deep values of honesty, respect for human life, and the overall good of mankind. In an era when standards and values are increasingly compromised, the temptation for aviation to lower those standards and reexamine deeply imbedded values may erode not only the public image but also safety in flight. The impact can be more significant than in other industries. When timeliness, economic concerns, and other pressures of the 21st Century tempt us to reexamine those high standards, safety frequently comprises the pivotal point that defines the line that must not be crossed.

The weight of establishing and maintaining high values and standards in aviation organizations falls directly and indirectly on the leader or executive who sets the tone and the emphasis on values by communication to employees, but more importantly by example. Kern (1998) pointed out that 95 percent of all aviators fly as part of some organization and the decision makers of those organizations decide many of the trade-offs between safety and efficiency, new equipment and old, selection of personnel and types and condition of equipment. Most studies affecting flight safety have focused on human and machine in the areas of piloting, maintenance, dispatch, and air traffic control He postulated that we need to look at what management does, and why. It is the "key choices that management makes that signals its priorities and aviators are quick to take note" (p. 101). He cited examples of one major carrier that came under intense scrutiny after five major accidents over a two-year period because of the elimination of two preflight checks in an attempt to achieve more on-time departures. Some of those accidents were the result of leaving with too little fuel to complete the trip. Bottom line, economic pressure affected decision-making practices and basic values were compromised with grave results.

A Comparison of Aviation Values with those in the Audit Profession

While the aviation field generally has very strong values, it is not alone in this orientation. Other institutions are similarly driven by powerful ethical values that have shown themselves to be vulnerable to human frailty.

The field of financial audit is particularly value-laden. Because of the dependence of financial markets on audit reports of the economic health of the firm, objective analysis of such data by an independent auditor is critical. This is why the recent Enron scandal has focused so much attention on one of the largest audit firms, Arthur Andersen.

Ringle (2002) discussed the ethical philosophy of Arthur Andersen himself who, in addition to establishing the firm by the same name, was the first salaried president of the New York Stock Exchange. This article quoted Andersen in a 1941 commencement speech to St. Olaf's College, "Man must be moved by high moral and ethical concepts in all of his relationships. Without this anchorage, he is…lost."

Moving forward to the current day,Ahrens (2002) argued that Arthur Andersen's attempts to recover its public image damaged significantly by the Enron scandal were discussed. The article pointed to newspaper advertisements taken out by Arthur Andersen as a method of "damage control" in the wake of the Enron financial collapse. The same article pointed to a February 6 advertisement in which Andersen chief executive Joseph F. Beradino mentioned eight times that "changes" or "fundamental changes" were needed in the firm's practices.

The Ahrens article continued stating that the Andersen advertisements are the work of Chiopak, Leonard, Schecter, and Associates, a Washington communications firm known for its work with companies having difficulty with public image. Andersen hired this firm in response to overwhelming media requests for company response. The result was a series of newspaper advertisements around the country at a cost of up to $80,000 per full-page offering. This expense shows how anxious Anderson is to shore up its flagging image in the face of the scandal.

Literature on the audit profession is replete with articles on the need for audit independence. One such article (Barlas, 2000) reported that Jerry Sullivan, executive director of the Public Oversight Board (POB), indicated a strong interest in examining Big 5 accounting firms to assure that auditor independence is maintained. POB was interested in also looking at whether audit partners have possibly compromised their independence by maintaining relationships with audit clients. Barlas cited the fact that the Securities and Exchange Commission (SEC) enlisted the POB in the wake of a January 2000 report by Jess Fardella, an independent SEC consultant, who examined independence violations at PricewaterhouseCoopers (PwC). PwC agreed to spend $2.5 million to set up an internal education program as part of a settlement of these charges.

Levinsohn (2000) reported on a 4-0 vote by the SEC that proposed new rules that would ostensibly remove any potential conflict of interest that would impair the absolute independence of corporate audits or the credibility of financial reporting. More specifically, the proposal would bar accounting firms from providing audit clients non-audit services such as bookkeeping; financial information systems design and implementation; appraisal or valuation services; actuarial services; internal audit outsourcing; management functions; human resources, legal, and expert services; and broker-dealer, investment advisory, and investment banking services. Interestingly, members of Congress criticized the Chair of the SEC for failing to provide the House Commerce Committee or the Senate Banking Committee with empirical evidence showing that auditing plus consulting services reduces audit independence.

Levinsohn's (2000) article concluded by quoting the SEC Chair, Arthur Levitt, who replied, (to the Congressional criticism):

> An accountant is not independent when the accountant has a mutual or conflicting interest with the audit client, audits his or her own work, functions as management or an employee of the audit client, or acts as an advocate for the audit client. (p. 74)

The audit profession with the financial disaster of Enron as a possible casualty has resisted this direction.

Individual Values and Organizational Culture

Organizational culture is defined in a recent textbook on Organizational Behavior as, "the informal set of values and norms that controls the way people and groups in an organization interact with each other and with people outside the organization" (George & Jones, 2002). Organizational culture within individual firms and within industries is clearly operative in addition to, and perhaps superimposed over, individual values in organizational settings.

Again, in examining ethics in the accounting field, the issues of ethics within the profession versus individual values were studied in a recent scholarly article (Douglas, Davidson, & Schwartz, 2001). In this study, individual values were found to be important, but were overshadowed by the values of an organizational culture. More specifically, in this study researchers found that ethical culture affects individual idealistic values and idealism affects judgment. Moreover, the moderating effect of culture on judgment will be more pronounced if the culture explicitly supports distinct values.

In the same article, the researchers assert that:

> "Socialization theory leads us to expect an eventual convergence of personal values with those of the organization. Firms mold their members to fit the organizational environment, or select and promote individuals who already fit into the prevailing culture and cause those that do not to leave" (Douglas, Davidson, & Schwartz, 2001, p. 101).

It is easy to extend the findings in the Douglas, Davidson, and Schwartz article to the aviation industry. As found by Kutz (1998), ethics are quite strong in aviation. From the Douglas et al. study, one may logically infer that the culture in aviation is the primary driver of such ethics. If this inference is correct, it follows logically from the conclusions of Douglas and her colleagues that more ethical judgments might be achievable through more instruction in ethics and in a professional code of conduct. In other words, such standards can and should be taught.

Kouzes and Posner (1995) found that people in organizations, regardless of the industry, sometimes drift when they are unsure or confused about their values. They tend to know what is expected of them and can better handle conflicting demands when they have clarity concerning not only their own values but those of their leaders and the organization they represent (p. 214).

Are Values Teachable?

Bennett (1993) postulated that virtues must be learned; people are not born with them. There is some evidence to support the notion that they also can be unlearned. Environment and circumstances can erode even the most committed. Values can be weakened over time and gradually modified due to pressures of productivity, timeliness, and profits in organizations.

Ash (2001) described the urgency of educating Air Force troops on values. In his words,

> "…the educational process must never let up but continually reinforce ethical fitness. But soldiers, sailors, and airmen do not have time to read Aristotle in the heat of battle. They must have already probed the difficult and morally ambiguous issues, and they must have the benefit of a familiar code to carry them through challenging times" (p. 39).

The very survival of an organization could potentially be affected by a failure to continually educate and maintain a sense of shared values in an organization. Failure to develop through education and teachable moments a congruence of values throughout the organization or a failure of leadership to model those values, which sustain the organization through difficult days, can impact the survival of a company. For example, labor issues could bring a company to its knees if executives spend money in other ways while cutting salaries and espousing to employees and labor their concerns for survival of the company. If employees perceive that the actions of leadership are inconsistent in that they spend money foolishly (at least from the employee's perspective) on other things while cutting costs on labor, that failure to model conservative spending erodes trust and leaves a perception that there really is no danger of survival which ultimately brings a company down. It is no accident that a company like Southwest Airlines survived 9/11 with no layoffs and was among the first to begin the recovery process. Southwest is known for consistently modeling and continually educating their constituency on such values as the importance of low costs, the importance of every employee being a leader, the importance of every job (as demonstrated by the Chief Executive Officer showing up to help line employees), and a philosophy of "love" that means going out of your way to be there for each other and for customers (Freiberg & Freiberg, 1996).

Regardless of the industry, creating a cooperative, proud industry is the responsibility of its leadership. Kouzes and Posner (1993) referenced recent studies that document the key role played by "community in the relationship between leaders and their constituents" (p. 130) and that sense of community is a direct result of the leader's effort to build community through a common purpose and shared values. They contrasted competitive and independent leaders who are seen as obstructive and ineffective.

The challenge for leaders is how to develop that sense of community and shared values regardless of the type of organization being led.

Tichy (1997) described the success of the leadership of Ameritech in changing the values of more than 60,000 of its employees. Tichy ascribed the survival and performance of Ameritech as a tribute to the determination and success of its leaders in overcoming resistance and radically changing the values of the company. Tichy said, "changing people's values is even harder than changing their ideas but in the long run it is more important….For change to be effective, leaders must rewrite the software—the values that guide people's actions" (p. 127).

A concerted effort to teach values is an important part of organizational development regardless of the industry and can frequently affect the survival of the organization. That effort can be accomplished in a variety of ways including such informal methods as leadership modeling, using teachable moments, informal leader/employees discussions. All of those informal methods should supplement a sustained effort to develop a sense of common purpose and shared values through formal group discussions, organizational communiqués (i.e. newsletters, bulletins, etc.), and classroom instruction both in the academic and organizational training environment.

The role of academia in values education can be broad in terms of values and ethics classes, which establish the role standards play in determining the fate of organizations as well as coursework specific with examples and illustrations of the role standards play in individual, organizational and industry success. For example, academic instruction pertaining to regulatory requirements in aviation could be supplemented with discussion of consequences of violations and weakening of standards. Aviation naturally has a sense of pride in its accomplishments over the last Century and students of aviation history can learn the responsibility for maintaining that sense of pride through maintaining the high standards that have brought that industry to the forefront as the safest transportation mode. Similarly, other industries such as accounting with a proud history of high standards can benefit from leadership training in the role and importance of values to success.

The U.S. Air Force, in its new Developing Aviation Leaders (DAL) initiative, recommended integration of learning across institutions by developing joint courses and programs, shared leadership opportunities, faculty exchanges, and others to ensure that there is consistency in the educational process (Drohan & Murray, 2001, p. 21). Improved cooperation between academia and industry in developing leaderships skills aimed at organizational value development and renewed emphasis on leadership ethics may provide some much needed leadership skill development regardless of the industry.

Instruction in industry-specific values and ethics may be the catalyst that ties today's leadership education in technical, management, and professional skills to another level of success involving practical application of principles and values in the organizational environment.

Further research is needed in a variety of aspects of values education to avoid the pitfalls of moving too far in one direction or another. Tentatively, however, findings of contemporary research seem to favor the notion that values and value development are important regardless of the industry.

Summary

Even though loss of human life may be an outcome of lax standards and weakening values in aviation thus contributing to their perception that honesty is innate to the industry, aviation leaders do not have proprietary exclusiveness on high standards and values. Other industries such as the audit industry have a proud tradition of high standards that is important in retaining the respect granted that industry in the past. The very term "audit" implies examination and review for compliance with high standards. Not only an aviation organization but other organizations such as those in the audit industry, also known for high standards, can be brought to their knees by weakening standards that erode public trust. This calls attention to the urgency of continual and consistent values education processes regardless of the industry. The importance of continual education on values and ethics that permeates every aspect of every day with every client may indeed be the secret to survival and success in leading most organizations.

Although there are potential pitfalls in values education and much research is needed to avoid those pitfalls, the potential pitfalls of failure to develop leadership skills that address development of shared values in organizations are equally high. One possible pitfall in understanding values is the "one size fits all concept." Kouzes and Posner (1995) cautioned that "successful companies may have very different values and that specific set of values that serves one company may hurt another" (p. 215). On the other hand, values such as honesty are not industry-specific; although critical to the aviation industry in terms of consequences, research indicates that honesty is just as essential to success in audit and other industries with different consequences.

References

Ahrens, F. (2002). Firms on the spot roll out the spots making 'crisis ads' a fine art. *Washington Post*, February 9, 2002, Edition: F, Section: F, E1.

Ash, E. (2001). Purple virtues, a leadership cure for unhealthy rivalry. *Leadership Special Edition of Aerospace Power Journal. XV(2)*.

Barlas, S. (2000, March). Auditor independence rules. *Strategic Finance*, *81* (9), 25-26.

Bennett, W. J. (1993). *A treasury of great moral stories.*

Carmichael, D. B. (1994). *High achievers in Federal service in Oklahoma* (Unpublished doctoral dissertation, University of Oklahoma).

Drohan, T. & Murray, D. (2001). Responding to the "Developing Aerospace Leaders" initiative, a master attack plan for reforming undergraduate professional development. *Leadership Special Edition of Aerospace Power Journal.* XV(2).

Douglas, P. C., Davidson, R. A., Schwartz, B. N. (2001, November). The effect of organizational culture and ethical orientation on accountants' ethical judgements. *Journal of Business Ethics, 34* (2).

Freiberg, K. & Freiberg, J. (1996*). Nuts! Southwest Airlines crazy receipt for* *business and personal success.* New York: Broadway Books.

George, J. M., Jones, G. R. *Organizational behavior,* (*3rd* ed.). Upper Saddle River, NJ: Prentice-Hall.

Kern, T. (1998). *Flight discipline.* New York: McGraw-Hill.

Kouzes, J. M. & Posner, B. Z. (1995). *The leadership challenge.* San Francisco, CA: Jossey-Bass Publishers.

Kouzes M. M. & Posner, B. Z. (1993). *Credibility, how leaders gain and lose it, why people demand it.* San Francisco, CA: Jossey-Bass Publishers.

Kutz, M. N. (1998). Developing future aviation leaders: advice from today's leaders! *The Journal of Aviation/Aerospace Education and Research.* Spring 2000, 9(3), 21-27.

Kutz, M. N. (1998). *Characteristics of successful aviation leaders of Oklahoma* (Unpublished doctoral dissertation, Oklahoma State University).

Levinsohn, A. (2000, August). SEC wants audit integrity disclosed. *Strategic Finance, 82* (2), 74.

Murphy, J. D. (2000). *Business is combat, a fighter pilot's guide to winning in modern business warfare.* New York: Harper-Collins.

Robbins, S. P. (2001). *Organizational behavior* (9th ed.), Upper Saddle River, NJ: Prentice Hall.

Tichy, N. M. (1997). *The leadership engine, how winning companies build leaders at every level.* New York: Harper Business.

International Journal of Applied Aviation Studies, 3,1

Interview

Concepts for Air Traffic Course and Lesson Design for New Equipment Systems Training

On January 16th of 2003, the staff of the International Journal of Applied Aviation Studies (IJAAS) had an opportunity to talk with Dr. Bob Welp about some of the design concepts he has found to be useful in the creation of new courses used for skill development on new equipment. Dr. Welp is an instructional systems specialist for the Federal Aviation Administration. His Oklahoma office is at the Mike Monroney Aeronautical Center in Oklahoma City.

Purpose of the Interview

This interview presents concepts for the design of new equipment systems courses for Air Traffic Control personnel. The focus is on overall course structure and individual lesson structure for classroom and computer-based instruction (CBI) lessons. The concepts covered in this interview are design ideas from a variety of sources that have been implemented in the past five years and include a number of suggested conventions for the design of effective computer-based instruction. The content for courses that support new equipment systems will necessarily vary in length and complexity depending on the nature of the system and the job skills to be taught. Experience has demonstrated that regardless of this variability, the basic course and lesson structures and CBI conventions presented here can be used effectively.

IJAAS: First of all, how should I address you? Do you prefer to be called Dr. Welp?

Bob Welp: Bob is fine.

IJAAS: Bob, I know you work for the FAA in Oklahoma City, but isn't the division you work for based in Washington, D.C.?

BW: It is. This office supports our work with the Academy and other training functions.

IJAAS: The focus of this interview is on designing courses for new equipment. Is it more difficult to create a course for new equipment? the new equipment or simulation

.

Requests for reprints should be sent to Beverly Laughead, FAA Academy, AMA-502, P.O. Box 25082, Oklahoma City, OK 73125.

BW: There are a couple of unique challenges. First, you must have the training ready to go when the equipment is fielded—you don't have the luxury of relaxed schedules. Second, the system is often evolving and changing as you are developing the training, so you are constantly learning, checking, and recycling through the materials to ensure the training is complete and accurate. Third, you are usually working with national workgroups who have little or no training development experience or expertise. It's a steep learning curve for many of these people.

IJAAS: What would the course structure look like for a new equipment course?

BW: We start by designing a course to teach outcomes. By outcome, I mean to be able to use the new equipment and perform all tasks correctly. We then design individual lessons to teach terminal objectives to support the outcomes. For example, a terminal objective might be to perform a single task correctly. In my opinion a course should be designed around outcomes and objectives and be taught in the following sequence:

First, teach the basic knowledge about the system. Describe the purpose, functions, and the architecture of the system, whether using CBI or classroom instruction. Second, organize subsequent lessons around terminal objectives (i.e. job tasks) or similar functions so that each lesson has a meaningful theme, can be practiced, and is easily related to job functions. Each lesson should first teach the knowledge required to learn the task(s) or functions taught in the lesson. For lengthy courses, we've found it useful to intersperse practice lessons that require students to determine when a task should be performed and to require them to perform multiple tasks. This method helps the learner translate the training to the job. Third, toward the end of the course students need to be able to practice (when appropriate and possible) the skills and tasks in a simulated environment such as a simulator or computer-based instruction (CBI) or on the actual equipment. Fourth, at the end of the course include a knowledge test and a performance check to ensure all tasks can be performed correctly. The conditions for the performance check should be as realistic as possible. Lengthier courses may require periodic reviews and testing to avoid a long end-of-course test.

Examples of basic course structures for En Route, Terminal Radar Approach Control (TRACON) and Tower training environments are illustrated in Figures 1 and 2. Both structures begin with *classroom lecture, proceed to CBI to teach equipment operations, return to classroom lecture for national procedures and end with practice and performance of air traffic control tasks using the new equipment or simulation.*

Classroom Lecture	**CBI**	**Classroom Lecture**	**(DYSIM or other Simulation**
• Introductions • Outcomes • Standards • Course Organization	• CBI lessons alternate with classroom reviews and simulation practice; periodic knowledge and performance checks	• National Procedures	• Practice procedures & tasks • Performance Check

Figure 1. Generic Course Structure for En Route / TRACON Training.

Classroom Lecture	**CBI**	**Classroom Lecture**	**Hands-on practice or on-the-job training**
• Introductions • Outcomes • Standards • Course Organization	• CBI lessons with a knowledge and performance check at the end	• National Procedures	• Practice tasks and procedures (if possible) • Performance Check

Figure 2. Generic Course Structure for Tower Training.

IJAAS: You've brought some visual examples of two structures for course development. We plan to print those examples in the journal. Would you like to comment on them?

BW: The main difference between the two structures is in the way the CBI is presented. In the En Route and TRACON environments, CBI should be alternated with classroom discussion and simulation practice (a more detailed depiction of this cycle is shown in figure 3). For example, in the En Route environment, where class size is typically eight students, it is possible for all eight students to simultaneously take CBI, return to class for a brief discussion and review, and then go to Dynamic Simulation (DYSIM) to apply what they have learned before repeating the cycle. This is an effective approach to learning because it breaks the training down into smaller "chunks" and allows students to learn and apply one or two tasks at a time.

IJAAS: So there are two structures and what you just described was for

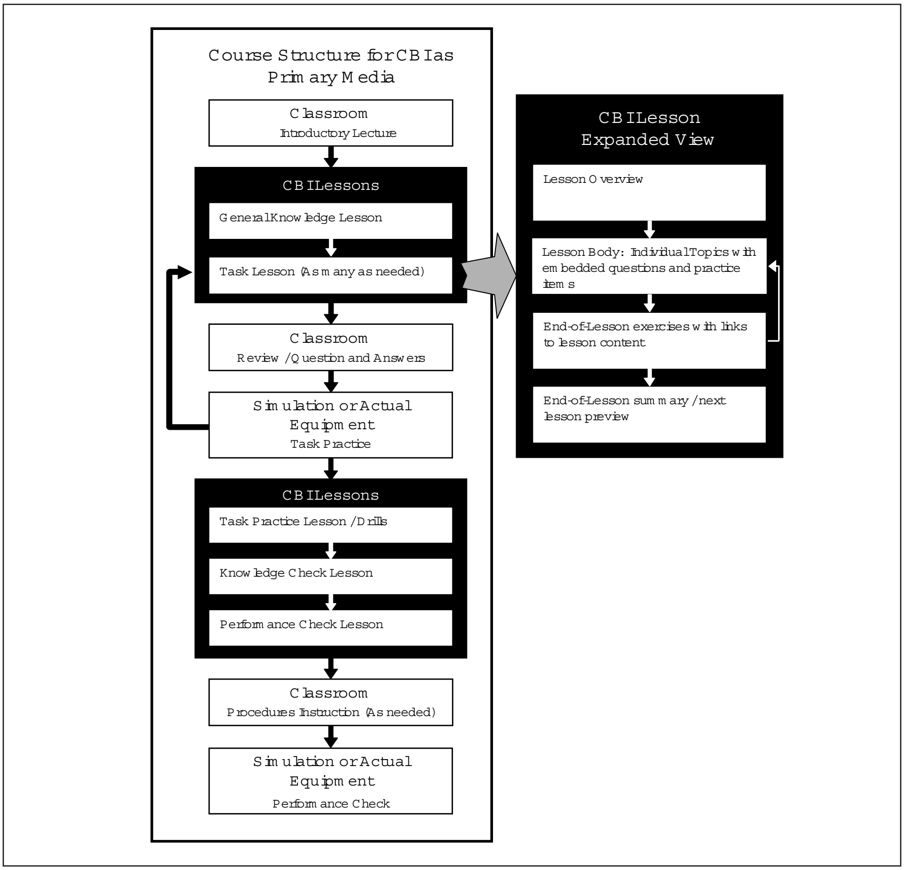


Figure 3. Computer Based Instruction Course and Lesson Structure.

the En Route and TRACON environment. What does the classroom and CBI mix look like for the Tower environment?

BW: In the Tower environment, where class size is typically four students, there are usually not enough CBI workstations to permit all students to take the CBI together, so this cycle will not easily work. However, the amount of information and skills that need to be learned to use a Tower air traffic system are also usually less, which reduces the need to "chunk" the instruction. So we have found that in the Tower environment it is often practical and effective to have students complete the CBI on their own prior to attending classroom instruction and practicing on the actual equipment. We include a performance check on the CBI to ensure students are ready to advance to the more realistic and complex environment of using actual equipment or simulation.

IJAAS: Are there other differences between the En Route/TRACON and Tower environments?

BW: Yes, another difference between the En Route/TRACON and Tower environments is the capability of providing realistic practice and performance checking The En Route and TRACON training departments can replicate realistic, dynamic air traffic scenarios through the use of simulation. This affords the air traffic controllers an opportunity to practice individual and multiple tasks. The Tower environment does not have such a capability. They employ site-specific systems and the visual aspects of the Tower cab view cannot be easily simulated. This often limits the realism possible for practice and performance checking.

IJAAS: I am familiar with the differences between learning to gain knowledge and learning to gain a skill. It would appear that your learning outcome is to achieve mastery of a set of tasks. Does this mean that each student must perform perfectly?

BW: I'm glad you made the distinction between two of the types of learning outcomes.

In our new equipment lessons we do aim toward the mastery level of competence. Mastery level competence is measured on end-of-course knowledge and performance checks. Considering the safety-critical nature of air traffic control and the fact that controllers must be ready to use the equipment right after they finish training, the standard for knowledge and performance checks should be 100%, but the performance check doesn't have to be a painful experience. A couple of key design features have helped us design knowledge and performance checks that result in high achievement and are well-received by students. First, the tests are carefully designed to include only important knowledge and skills. Second, we try to avoid the classroom setting, where the logistics of passing out, taking and scoring tests can be cumbersome. If the knowledge in the lesson requires testing, we recommend the test be on CBI rather than using a written test because the test can be self-paced and missed items don't bruise egos the way they might in a classroom setting. If the tasks taught in the lesson require testing, the test should be conducted first on CBI, then on the actual equipment or a simulator, if possible. Third, the tests are efficient in that students get a list of the items they missed and they can then retry the item or receive supplemental instruction. They only need to retake the items they missed. When asked, most students have told us they would rather be tested on important information and skills and be required to get them all correct rather than be tested on everything and only be required to achieve a 70 or 80 percent correct score. Also, for knowledge check items, when a student misses a question, we present a different question or an alternate form of the question so the student doesn't just focus on the answer. Instead, they review all of the content related to the objective because they know they will get a different question.

IJAAS: Does the goal or purpose of the training influence the design?

BW: The goal or purpose of training should be captured in the course outcomes, and these outcomes should match the job as closely as possible. The goal or purpose should also include all tasks that are performed on the new equipment. In my opinion, the best training has course outcomes that include all tasks and require mastery under job-like conditions. Then the course design and structure decisions are made to support this goal.

IJAAS: Do all courses start in the classroom?

BW: Generally, the first lesson of a course is presented in a classroom lecture format and includes the course length, structure, and the schedule of learning events.

Table 1
Course Design for Tower Data Link Services Air Traffic Control Specialist Course.

	Training Task #	Length (min.)	Instructional Strategy	Media	Testing
Overview of TDLS	1, 2	20	Tutorial	CBI	Knowledge, Performance
Basic TDLS Operation	8, 9	15	Tutorial Demo Practice	CBI	Knowledge, Performance
Flight Data Input/ Output Application	3	5	Tutorial	CBI	Knowledge
Pre-Departure Clearance Application	4, 5, 12, 13	30	Tutorial Demo Practice	CBI	Knowledge, Performance
Digital Automated Terminal Information System (D-ATIS) Operation, Part 1	6, 7	15	Tutorial Demo Practice	CBI	Knowledge, Performance
Advanced D-ATIS Operation, Part 2	6, 7, 9, 10, 11	15	Tutorial Demo Practice	CBI	Knowledge, Performance

Review and Practice	All	15	Test	CBI	Knowledge, Performance
CBI Knowledge Test	All	10	Test	CBI	Knowledge test (90% to pass)[1]
CBI Performance Test	All	10	Test	CBI	Performance test (100% to pass)
Facility Instruction on Local Procedures and Configurations	All	TBD	Tutorial	Lecture	N/A
On-the-job training and performance check		TBD	Demo Hands-On Practice	OJT	Performance test (100% to pass)

The first lesson may also present general system knowledge if not presented elsewhere.

IJAAS: Let's talk about course design in a little more detail. We've briefly covered some aspects of course design, but could you give us an example of a specific course?

BW: This graphic (see Table 1) provides an example of how the guidelines were applied to the Tower Data Link Services (TDLS) ATCS Course as derived from the TDLS Training Development Plan. General system information is taught in the first lesson – Overview of TDLS. Notice that each lesson teaches at least one task (column 2) and all tasks are taught. Also notice that on-the-job training was used for the performance check. This was because neither simulators nor training devices were available. To help ensure that the course would be effective under such circumstances, the CBI was designed to closely replicate TDLS functions and included practice and a performance check for all tasks. Given the comprehensive nature of the CBI, it was determined that a classroom lecture was unnecessary. The information normally required in the first lesson (course length,

---

[1] In this course, the knowledge test standard was raised from 70% to 90% as a trial to determine the feasibility of using higher standards. Results of this trial and a second trial using a 100% standard demonstrated that standards for knowledge tests could be set at 100% if two design features were incorporated: (1) limit the test items to knowledge that is truly essential to learn to operate the system properly; and (2) provide access to review material when a question is missed and permit retries of the question.

etc.) was placed in a separate document so instructors could inform students before taking the CBI. The TDLS Course design illustrates how the basic course structure serves as a guide for course design and then is refined to adapt to the unique circumstances or requirements of training delivery in the field facility.

IJAAS: I would like to probe a little deeper into the mechanics of lesson plan development for the classroom. You've given us some general things to think about, but perhaps you can fill the details of this developmental process.

BW: Classroom lessons will generally be of three types: (1) instructional lessons for presenting the topics and tasks, (2) review lessons to go over student questions and review key points, and (3) practice lessons for conducting practice on actual equipment that may be located in the classroom. Instructional lessons should be organized around the four elements: the purpose, lesson objectives, lesson body, and lesson review.

IJAAS: Perhaps you can talk about these four elements in more detail.

BW: When describing the purpose, we begin with a general description of what will be taught. For example, "This lesson will introduce the Tower Data Link Services and provide general system information." We follow the purpose with lesson objectives, in plain language describing the behavior to be learned. For example, "You will identify the purpose, functions, and characteristics of the Airport Movement Area Safety System." Or, "You will be able to open or close a runway, change the configuration of the airport, and change the direction of the leader lines on the Airport Movement Area Safety System." The objectives also serve as a guide to the main topics of the lesson. We then tell the students whether or not the objectives will be tested and how they will be tested. We recommend presenting the content of the lesson, one objective at a time, teaching the knowledge that is specific to each task before teaching the steps to perform the task. There are at least three ways you can sequence the content of a lesson. You can organize around visual information, such as displays or menus, in which case the sequence would be left to right, top to bottom of the display.

IJAAS: Does that describe a person's natural scan technique?

BW: Yes it does. When displaying visual information it is important to follow the normal scan pattern used by most learners. Learners start at the top left and scan to the right and down.

IJAAS: This would be particularly important to know if you were a CBI developer. What other techniques do developers use to display information?

BW: Although everything on a computer screen is visual, we can't forget that there is a cognitive, meaning-seeking element to processing information. One technique we use is to teach concepts, procedures, or operations from simple to complex. Procedures and operations are usually displayed in steps. This is a natural and job-related sequence that establishes a direction and meaning of the content. One last sequence I would like to mention is for rules. Although it's a little confusing, Air Traffic rules are called procedures and they tell you when you must take certain actions. These rules are best taught starting with general situations and then covering specific and unusual situations.

IJAAS: Lectures can be rather boring if the students don't have an opportunity to interact. Do you provide activities for the students?

BW: I'm glad you mentioned interaction. I think it's very important and it can take several forms depending on the nature of the content and the media you're using. Let me take a minute to describe some of the types of interaction you can use in the classroom. One is embedded questions. We use embedded questions to enhance recall of facts or concepts. Another is the "What if" application questions. A "What if" question prompts students to apply rules. When practical, we also include hands-on practice. Hands-on practice provides an opportunity for skill development, especially if students need to be able to perform quickly. Interaction improves learning when it is meaningful and it better prepares students for testing. I like to close my comments about classroom lesson plan development with a mention of the lesson review. The review should not be simply a list of the objectives/topics covered. It should be used to cover any teaching points that are key, complex or controversial, and every review should include a reminder to the instructor to ask students if they have any questions.

IJAAS: You've covered the key elements of the classroom lecture method, but we haven't heard much about CBI. Is CBI development different from traditional classroom lesson plan development?

BW: The instructional techniques used in CBI development are similar to those used in the traditional classroom. Remember, we are only talking about a change in medium, not a change in educational approach. The CBI medium offers us some opportunities that are not possible in lecture. The two most significant of these that come to mind are the capability to replicate equipment displays and entry devices and being able to allow students to work at their own pace. A generic course and lesson structure in which CBI is the primary medium used for the en route / TRACON environment is shown in this graphic (see Figure 4). Notice that while CBI provides much of the instruction, it is not the only medium used. Classroom instruction, simulation, and actual equipment, are generally necessary for a course to be effective.

IJAAS: So no one medium holds all the keys to successful completion of the course. Each one has its strengths and weaknesses.

BW: Right. So let's talk about CBI development. Effective CBI will generally have four types of lessons: (1) instructional lessons for presenting the topics and tasks, (2) practice lessons for task practice, (3) knowledge check lessons for testing knowledge, and (4) performance check lessons for testing task performance. Similar to classroom lecture, CBI lessons should be organized around the four elements of lesson overview, lesson body, end-of-lesson exercises, and end-of-lesson summary. The lesson overview informs the students of the topics they will learn about or the tasks they will learn how to perform. An effective CBI overview can be accomplished on a single screen display. The lesson body presents the individual topics of the lesson. For new equipment systems training, this should again follow the pattern of first presenting the knowledge relevant to the skills or tasks being taught in the lesson. Each skill and/or task would then be presented (e.g., steps presented for students to follow) or demonstrated followed by an opportunity for the student try the skill/task. This process would continue until all skills or tasks for the lesson have been taught. At the end of the lesson, present the student with questions about the topics and/or practical exercises to practice the tasks. The students should be required to complete each question or exercise correctly. Because we're still teaching at this point, incorrect responses should result in corrective feedback so the student can retry the question. Students should also be told what they missed and have the opportunity of returning to the point in the lesson where that content is presented (i.e., hyperlink). Finally, end the lesson with a list of the topics covered.

IJAAS: When you talked about classroom lecture development, you specifically mentioned what you would do in the first classroom lesson. What does the first CBI lesson look like when you develop a new equipment course?

BW: Unless taught elsewhere, the first CBI lesson should present general knowledge about the system or topic to be taught. For new equipment systems, this includes such topics as: system purpose, functions, features, components, functional architecture, capabilities, specifications, input and display devices, and input and navigation methods. In addition to general knowledge taught, the first lesson of CBI should also contain an overview of the CBI lessons, performance and knowledge test information, and navigation information. In the overview inform the students of the number of lessons and lesson topics. Describe the knowledge and performance check lessons and the standards for completing each. Finally, explain any special procedures required to navigate the CBI, such as the use of a trackball, special keyboard, or other input device. The last three lessons should always be task practice, knowledge check, and the performance check.

IJAAS: CBI development is still relatively new. It is much younger than other media. In the past few years, what have you learned about improving CBI? Do you have anything you would like to pass on to our readership?

BW: Sure. From 1998 to 2002, eight air traffic new equipment systems training programs using CBI were developed (or revised), evaluated, and fielded. These programs include: Airport Movement Area Safety System (initial system and Build 5), Tower Data Link Services, User Request Evaluation Tool (initial system and Build 2), Controller-to-Pilot Data Link (ACARS and ATN versions), and Airport Surface Detection Equipment Model X. Lessons learned for course development and design were compiled for each of these for use in improving subsequent training programs. Many of these lessons learned were directed at streamlining and enhancing CBI and have resulted in a number of CBI design conventions that have been implemented over this period. The list does not include basic principles of instructional design, and I don't consider it to be an exhaustive list of conventions for delivering instruction via CBI. The list does, however, represent the most significant design improvements we have made in the last few years and has clearly resulted in improved CBI (see Addendum). One convention is to design lessons to be about 20 minutes and no longer than 30 minutes. This results in several benefits. It gives students a greater sense of progress and accomplishment in completing the lessons, reducing the frustration of long lessons. It is a pace better suited to CBI than the traditional classroom schedule of 50 to 60 minutes because CBI usually covers about twice as much content than classroom lecture in the same amount of time. It invites the students to go take the CBI because students feel like they can make progress in the course and finish a lesson or two even when they have limited time to get to the CBI room. Another area we have been learning about is how to teach multiple methods of operating equipment. When multiple methods are possible (i.e., from different input devices such as keyboard, mouse or trackball or from different screens, menus, or hot keys), generally it is best to teach all methods (unless the experts indicate otherwise). But during End-of-Lesson Exercises and the Performance Check, allow the student to choose which method to perform the procedure.

IJAAS: What if teaching all these methods is impractical or interferes with learning. What would you do then?

BW: Sometimes there are just too many possible methods and it would either take too long to teach all of them or, there are so many that students have difficulty mastering even one. When that happens, we teach, practice, and test a single, expert-preferred method, mention that other methods are possible, and direct students to the operator manual for more information. Another alternative is to teach, practice, and test a single expert-preferred method first, then selectively teach one or more alternate methods, as necessary. The key is to select methods that are known to be useful, such as a short-cut method.

IJAAS: Well, we have taken a close look at the development process for new equipment courses. You've made suggestions for development in both the traditional classroom and for CBI. And you ended our discussion with a rather detailed list of conventions you use in developing CBI lessons. We appreciate your taking the time to meet with us. Do you have any parting comments?

BW: Yes I do. The progress we have seen the last four and one half years has been possible because of the way we are approaching our training programs. In the past, each program was a single, isolated process where the there was no mechanism or central repository for sharing ideas across programs or building on the best concepts. Each program seemed to start from ground zero. Now we are using the quality concept of "continuous improvement." We do this by using a common scale to measure results and we put together lessons learned for each program. We basically expect each training course to be complete, accurate and effective for all students. But in addition, we rate the courses on a variety of dimensions and gather comments from students and others involved with the training. When a new program starts, we examine the ratings and comments not only to improve the course we are working on, but also to carry forward the best ideas and tackle any remaining problem areas for future courses. This way, we are constantly improving the way we design, develop, and deliver the training. A side benefit of this kind of approach is that encourages everyone to get involved in generating ideas and solving problems, including students, and it creates a sense of "ownership" among the many people and organizations associated with the training. I'd say it has helped us create a common goal among training team members – "beat the last training program's score and achieve excellence."

IJAAS: Thanks for meeting with me today and providing such an insightful look into lesson design for new equipment training.

BW: It was my pleasure.

Acknowledgements

Several people have made significant contributions to the training pro-grams mentioned in this article. Les Jones, a specialist in ATX-100, really got the process going by requesting the Data Link training program be developed as a "model" program. Paul DeBennedittis, the lead for Data Link training took this request to heart by challenging the training team to develop excellent instruction and he was also willing to experiment in the design of skill drills and other innovative ideas. Gary Washburn, a subject matter expert for the team worked closely with our CBI contractor, PTI, Inc., to develop and refine the first Data Link drills. In our Tower Data Link Services training project, Nancy Hyde of PTI, Inc. developed a computer-based instruction program with outstanding, realistic exercises that received numerous "perfect" ratings from students and later won an International Society for Performance and Instruction award for excellence. Ken Kurdziel and the entire training team for the User Request Evaluation Team were instrumental in implementing a "chunk and apply" approach in which students learned a few skills on CBI and then applied these skills in Dynamic Simulation. I've just mentioned a few of the people and their contributions - there are a lot more who I haven't named, but I mainly want to emphasize that these concepts have come from a lot of dedicated people.

Addendum

General CBI Design Conventions

1. Enable students who have exited CBI prior to completion to return to the same location in the instruction where they exited.
2. Highlight the "next" button on the CBI navigation bar to cue students to advance to the next screen.
3. Use bulleted text phrases (i.e., 5-7 words) rather than complete sentences and maintain adequate "white space" on the screen.
4. Avoid acronyms and engineering terms and wherever possible, use commonly known air traffic terminology.
5. For instructions on the left of the screen, initially tell the students a couple of times (in the narration) to use the instructions to perform the function. Then, just say to do the function.
6. During exercises that include narration, allow students to attempt the exercise (i.e., stop audio and respond to a question or make an entry) rather than having to wait until the narration is complete.
7. Alternate narrators between lessons; preferably using a male and female voice.
8. Design lesson to be about 20 minutes and no longer than 30 minutes.
9. Keep narration segments relatively short (30 seconds maximum) to maintain interest and attention and to prevent students from having to replay lengthy segments when they are distracted.
10. When multiple methods for performing an equipment procedure are possible (i.e., from different input devices such as keyboard, mouse or trackball or from different screens, menus, or hot keys), generally it is best to teach all methods (unless the experts indicate otherwise). During the end-of-lesson exercises and the performance check, allow the student to choose which method to perform the procedure.

    When teaching all methods is impractical, interferes with learning, or is not necessary, consider alternate approaches such as:

    - Teach, practice, and test a single, expert-preferred method.
    - Teach, practice, and test a single, expert-preferred method, but mention that other methods are possible and say where to go to find out how to perform them.
    - Teach, practice, and test a single expert-preferred method first, then teach one or more alternate methods, as necessary.

11. Provide practice exercises for each skill or task. In the lesson body, follow every presentation of a skill or task with an exercise that requires the student to try that skill or task at least once.

12. When equipment system entries have to be made, show the entry the student made, even after incorrect responses (so the student can find and correct an error) and provide the actual system feedback.

13. In the end-of-lesson exercises do not provide feedback on correct questions or exercise items (other than displaying the word "Correct") unless there is a key point to make. For incorrect questions or exercises, provide corrective feedback and give the student another opportunity to complete the item. At the end of the exercises, display a list of the items and indicate which items the student completed correctly on the first try. Provide a "hot link" from the items on the list to the corresponding instruction in the lesson.

14. Generally, a CBI course should require students to proceed through all instruction in a pre-established sequence that has been designed to optimize learning (an exception might be when using pre-tests to target instruction). Once the instruction has been completed, the student should be able to return and navigate through the CBI with relative ease, essentially permitting random access to content.

15 The standard for knowledge and performance checks should be 100%.