Federal Aviation
Administration

International Journal of Applied Aviation Studies

A Publication of the 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.

Individuals wishing to obtain a copy of the IJAAS on CD may contact Kay Chisholm by email at kay.chisholm@faa.gov, or by telephone at (405) 954-3264, or by writing to the following address:

International Journal of Applied Aviation Studies
Kay Chisholm
AMA-800
PO Box 25082
Oklahoma City, OK 73125

Go to Table of Contents

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.

David O'Hare   University of Otago, Dunedin, New Zealand
Esther Oprins   Air Traffic Control the Netherlands, Schiphol Airport, Netherlands
Jens Rolfsen   Det Norske Veritas, Høvik, Norway
Michael Siegrist   Institute for Environmental Decisions, Zurich, Switzerland
Steve Thatcher   University of South Australia, Mawson Lakes, Australia
Claudia van de Wal   VALK Foundation, Netherlands
Lucas van Gerwen   VALK Foundation, Netherlands
Leo Vermeulen   University of Pretoria, Pretoria, South Africa
Stephen Walsh   Interaction Training Associates, Netherlands
Damien J. Williams   University of Bristol, Bristol, UK
Robert Wolfger   Austrian Airlines, Strasshof, Austria
Rogier Woltjer   Linköping Institute of Technology, Linköping, Sweden
Beat Zimmermann   Air Navigation Institute, Thun, Switzerland

U.S.A.

Robert I. Aceves   St. Cloud State University, St. Cloud, MN
Vicki Ahlstrom   FAA, William J. Hughes Technical Center, Atlantic City, NJ
Amy Alexander   Aptima, Inc., Woburn, MA
Francis "Frank" Ayers   Embry-Riddle Aeronautical University, Daytona Beach, FL
Larry Bailey   FAA Academy, Oklahoma City, OK
R. Kurt Barnhart   Kansas State University, Salina, KS
Robert Baron   The Aviation Consulting Group, Myrtle Beach, SC
Ellen Bass   University of Virginia, Charlottesville, VA
J. Matthew Beaubien   American Institutes for Research, Washington, DC
Wendy Beckman   Parks College, St. Louis University, St. Louis, MO
Dennis B Beringer   Civil Aerospace Medical Institute, Oklahoma City, OK
Robert O. Besco   Professional Performance Improvement, Dallas, TX
Aleta Best   Department of Transportation, Washington, DC
Kate Bleckley   Civil Aerospace Medical Institute, Oklahoma City, OK
James P. Bliss   Old Dominion University, Norfolk, VA
Deborah Boehm-Davis   George Mason University, Fairfax, VA
Brent Bowen   University of Nebraska at Omaha, Omaha, NE
Jonathan Bricker   University of Washington, Seattle, WA
Dana Broach   Civil Aerospace Medical Institute, Oklahoma City, OK
Philip Budd   Graduate Studies in Counseling, Bethany, OK
Ernesto A. Bustamante   University of Idaho, Moscow, ID
James N Butcher   University of Minnesota, Minneapolis, MN
Evan Byrne   National Transportation Safety Board, Washington, DC
John Cain   Florida Technology Institute, Melbourne, FL
Raymond E. Cain, Jr.   Florida Memorial University, Miami Gardens, FL
John A. Caldwell   Wright-Patterson Air Force Base, Dayton, OH
Kim Cardosi   Volpe National Transportation Systems Center, Cambridge, MA
Stephen M. Casner   NASA Ames Research Center, Moffett Field, CA
Thomas R. Chidester   Civil Aerospace Medical Institute, Oklahoma City, OK
Robert H. Chisholm   Support Systems Associates, Melbourne, FL
David Conway   Southeastern Oklahoma State University, Durant, OK

Jerry Crutchfield	Civil Aerospace Medical Institute, Oklahoma City, OK
Michael DeGrave	Kenyon & Kenyon LLP, New York, NY
Carmen Delgado-Morales	FAA Academy, Oklahoma City, OK
Charles DeJohn	Civil Aerospace Medical Institute, Oklahoma City, OK
James DeVoll	Office of Aerospace Medicine, Washington, DC
Archie Dillard	FAA National Headquarters, Washington, DC
Brian G. Dillman	Purdue University, West Lafayette, IN
Key Dismukes	NASA Ames Research Center, Moffett Field, CA
Nancy Dorighi	Ames Research Center, Moffett Field, CA
Wayne Dornan	Middle Tennessee State University, Murfreesboro, TN
Graham Elliott	FAA Academy, Oklahoma City, OK
Boyd Falconer	Russell Reynolds Associates, Chicago, IL
Richard Fanjoy	Purdue University, West Lafayette, IN
Douglas R. Farrow	Federal Aviation Administration, Washington, DC
Wai-Tat Fu	University of Illinois, Savoy, IL
Kenneth H. Funk II	Oregon State University, Corvallis, OR
Peter C. Gardiner	SCSI, Torrance, CA
Kenneth Gardner	FAA, Flight Standards, New Cumberland, PA
Patricia Garrambone	FAA, Boston TRACON, Boston, MA
A. F. Grandt	Purdue University, West Lafayette, IN
Mavis Green	Harvard University, Cambridge, MA
Carla A. Hackworth	Civil Aerospace Medical Institute, Oklahoma City, OK
Chris Hallman	Great Circle Consulting, Inc., Newman, GA
Steven Hampton	Embry-Riddle Aeronautical University, Daytona Beach, FL
Jeffrey T. Hansberger	George Mason University, Fairfax, VA
Frederick D. Hansen	Oklahoma State University - Tulsa, Tulsa, OK
Samuel B. Hendrix III	FAA Academy, Oklahoma City, OK
Kevin High	Western Michigan University, Battle Creek, MI
Jerry Higley	FAA Academy, Oklahoma City, OK
Alan Hobbs	NASA Ames Research Center, Moffett Field, CA
Judy Holcomb	FAA Academy, Oklahoma City, OK
Willem Homan	Western Michigan University, Battle Creek, MI
Amy Hoover	Central Washington University, Ellensburg, WA
Todd P. Hubbard	Oklahoma State University, Stillwater, OK
David R. Hunter	Artis, LLC, Reston, VA
Peter Hwoschinsky	FAA National Headquarters, Washington, DC
Sig Illing	FAA Academy, Oklahoma City, OK
Kathi Ishimaru	FAA, Northwest Mountain Regional Office, Renton, WA
Florian Jentsch	University of Central Florida, Orlando, FL
Jeffrey Johnson	St. Cloud State University, St. Cloud, MN
Merrill R. Karp	Arizona State University, Mesa, AZ
Deann King	FAA Academy, Oklahoma City, OK
Raymond E. King	Civil Aerospace Medical Institute, Oklahoma City, OK
William Knecht	Civil Aerospace Medical Institute, Oklahoma City, OK
Jefferson M. Koonce	Technology Systems, Inc. Wiscasset, ME
Bruce Kuhlmann	FAA Academy, Oklahoma City, OK
Mary N. Kutz	Oklahoma State University, Stillwater, OK
Florian Jentsch	University of Central Florida, Orlando, FL
Victor LaSaxon	The Boeing Company, Midwest City, OK
Chien-tsung Lu	Central Missouri State University, Warrensburg, MO

Rebecca Lutte	University of Nebraska at Omaha, Omaha, NE
Richard L. Mangrum	Kent State University, Kent, OH
Carol Manning	Civil Aerospace Medical Institute, Oklahoma City, 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
Raymon M. McAdaragh	FAA, Hampton, VA
H. C. McClure	MAC Air Consulting, Queen Creek, AZ
Kathleen L. McFadden	Northern Illinois University, DeKalb, IL
Jeremy Mell	French Embassy, Washington, DC
Harry Minniear	Indiana State University, Terre Haute, IN
Marjo Mitsutomi	University of Redlands, Redlands, CA
Richard H. Mogford	NASA Ames Research Center, Moffett Field, CA
John Morey	Dynamics Research Corporation, Andover, MA
Kathleen Mosier	San Francisco State University, San Francisco, CA
Vahid Motevalli	The George Washington University, Ashburn, VA
Martin Mumenthaler	Stanford University, Redwood City, CA
Tom Nesthus	Civil Aerospace Medical Institute, Oklahoma City, OK
Michael S. Nolan	Purdue University, West Lafayette, IN
Gary J. Northam	Parks College, St. Louis University, St. Louis, MO
Dale B. Oderman	Purdue University, West Lafayette, IN
Allen J. Parmet	Midwest Occupational Medicine, Kansas City, MO
Manoj Patankar	Parks College, St. Louis University, St. Louis, MO
Donald A. Petrin	Purdue University, West Lafayette, IN
James J. Picano	Mind Quest, LLC, Suisun, CA
Jean Potvin	Saint Louis University, St. Louis, MO
Julia Pounds	Civil Aerospace Medical Institute, Oklahoma City, OK
Roni Prinzo	Civil Aerospace Medical Institute, Oklahoma City, OK
Edward Pugacz	FAA, William J. Hughes Technical Center, Atlantic City, NJ
Yandong Qiang	Johns Hopkins University, Baltimore, MD
Stephen M. Quilty	Bowling Green State University, Bowling Green, OH
William B. Rankin II	Florida Memorial College, Miami, FL
Esa M. Rantanen	Rochester Institute of Technology, Rochester, NY
Robert Ripley	Auburn University, Auburn, AL
Vladimir Risukhin	Western Michigan University, Kalamazoo, MI
Charles L. Robertson	University of North Dakota, Grand Forks, ND
Vladimir Risukhin	Western Michigan University, Kalamazoo, MI
Steven M. Samuels	US Air Force Academy, Colorado Springs, CO
Nadine Sarter	The Ohio State University, Columbus, OH
David Schroeder	Civil Aerospace Medical Institute, Oklahoma City, OK
Gregory L. Schwab	Indiana State University, Terre Haute, IN
Scott Shappell	Clemson University, Clemson, SC
J. Anthony Sharp	Ohio University Airport, Albany, OH
John W. Sheremeta, Jr.	American Airlines, Lake Ronkonkoma, NY
Mark A. Sherman	Ohio University, Athens, OH
Kenneth Sperry	The Boeing Company, Midwest City, OK
John E. Stewart II	U.S. Army Research Institute, Fort Rucker, AL

Alan J. Stolzer	Parks College, St. Louis University, St. Louis, MO
Judith B. Strother	Florida Institute of Technology, Melbourne, FL
Craig Gerald Stroup	American West Airlines, Phoenix, AZ
Terry von Thaden	University of Illinois, Savoy, IL
Judith Foss Van Zante	NASA Glenn Research Center, Cleveland, OH
Ron Ward	FAA Academy, Oklahoma City, OK
Thomas Weitzel	Embry-Riddle Aeronautical University, Daytona Beach, FL
Michael Wetmore	Central Missouri State University, Great Falls, MT
Michael Wiggins	Embry-Riddle Aeronautical University, Daytona Beach, FL
Kevin Williams	Civil Aerospace Medical Institute, Oklahoma City, OK
Donna Forsyth Wilt	Florida Institute of Technology, Melbourne, FL
Peter Wolfe	Professional Aviation Board of Certification, Washington, DC
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.

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

EDITOR'S NOTES

Papers

Our lead article by *Amy L. Hoover* investigates the *Long Term Effect of Concurrent Task Management Training on Pilot Task Prioritization Performance*. Aviation students who had previously participated in an experiment to investigate short-term effects of concurrent task management training were retested eight months later. These students showed negligible change in task prioritization performance after eight months. Students who previously exhibited a positive short-term training effect had task prioritization performance similar to the control group after eight months.

Jonathan B. Bricker's study examines hypothesized demographic differences in the three components of the Air Travel Stress Scale in *Who Has Air Travel Stress? Demographic Differences in Air Travel Stress*. Results showed significant indicators among passengers related to gender, income, flights flown, whether leisure or business, and international flight experience. Implications for understanding and intervening on air travel stress are discussed.

In *Using Probability and Set Theory to Examine Illustrations of Situation Awareness, Todd P. Hubbard* notes why there are misunderstandings of the construct of situation awareness. This article analyzes popular models of situation awareness and transforms them into five rules of probability and the general propositions of set theory. Suggestions are presented for the classroom as well as insights on how to observe and study situation awareness during simulator training.

Raymond E. King, David J. Schroeder, Carol A. Manning, Paul D. Retzlaff, and Clara Williams assess the viability of using of the Minnesota Multiphasic Personality Inventory-2 in lieu of the 16 Personality Factor in *Screening Air Traffic Control Specialist (ATCS) Applicants for Psychopathology*. A sample of 1,014 ATCSs completed the MMPI-2 as part of the research program, after being cleared with the 16 PF. The gathered data are used to estimate the number of future candidates that would be referred for follow-up psychological evaluations, given varying MMPI-2 scale cut scores

Analysis of the Drift Cues from a Tactile Belt to Augment Standard Helicopter Instruments showed significant improvement of drift control during takeoff and reduced drift error during hover. The *Ian P. Curry, Arthur Estrada, Catherine M. Grandizio, and Bradley S. Erickson* study also found that fatigued pilots reported a significant reduction in visual and physical workload with the belt. Results indicate that the belt significantly improved pilot perception of drift and situation awareness and reduced mental stress.

Stephen M. Casner's study of *General Aviation Pilots' Attitudes toward Advanced Cockpit Systems* surveyed 134 general aviation pilots. The results showed that general aviation pilots have usually positive attitudes about advanced cockpit systems and exhibit a strong preference for using them. Pilots recognized potential pitfalls associated with advanced cockpit systems but were more likely to ascribe the problems to other pilots than they were to themselves.

Emergencies and off-nominal situations will challenge the safe and efficient operation of NextGen. *Barbara K. Burian's* article *Perturbing the System: Emergency and Off-Nominal Situations under NextGen* focuses on three issues: 1) defining "emergency" and "off-nominal," 2) identifying the full-range of emergency and off-nominal situations and their effects on human operators, technologies, procedures, and NextGen operations, and 3) determining performance capabilities, limitations and external pressures affecting human response to these situations.

Does strategic team training at an air traffic control task increase long-term performance? In *Performance Assessment of Strategic Team Training in Simulated Air Traffic Control, Christopher P. Barlett, Christopher L. Vowels, John D. Raacke, and James Shanteau* compared teams receiving strategic team training with teams receiving factual training. Results suggest that the strategic training method produced positive long-term effects on performance in the dynamic decision environment.

The Significance of Demographic Characteristics in Airport Driver Training Programs is the second in a series by *William B. Rankin*. This article examines the problem of runway incursions at OEP-35 airports. The study examined if demographic characteristics are a significant factor in the airport driver training that employees receive at OEP-35 airports. The data suggested that demographic characteristics are significant factors in airport driver training and vary by geographic region.

Arlynn McMahon's research studied *Pilot Perceptions on Using a Ballistic Parachute System*. Results from a survey of 1,003 respondents showed pilots felt that an aircraft equipped with a parachute was safer than one without; however, flight experience affected pilot opinions. The study also revealed that pilots made decisions differently when considering flights in a parachute equipped aircraft during four scenarios.

Book Reviews

Critical Incident Stress Management in Aviation edited by Jörg Leonhard and Joachim Vogt.   Review by *Todd P. Hubbard*

Aviation English Course Books   Review by *Graham Elliott and Theresa White*

Aviation English for ICAO Compliance by Henry Emery and Andy Roberts with Ruth Goodman and Louis Harrison.

Cleared for Takeoff: English for Pilots, Books 1 & 2 by Liz Mariner

English for Aviation by Sue Ellis and Terence Gerighty

KC

International Journal of Applied Aviation Studies

Long Term Effect of Concurrent Task Management Training on Pilot Task Prioritization Performance

Amy L. Hoover

Department of Aviation
Central Washington University
400 E University Way
Ellensburg, WA 98926
hoovera@cwu.edu

Abstract

Seventeen university aviation students who had previously participated in an experiment to investigate short-term effects of concurrent task management training were retested eight months later. Pilots flew simulated flights on a Frasca 141 FTD. Nineteen task prioritization challenges were embedded within each flight scenario. Pilots from the previous study's control group showed negligible change in task prioritization performance after eight months. Pilots from the previous study's experimental group who previously exhibited a positive short-term training effect had task prioritization performance similar to the control group after eight months. These results indicate the short-term effect of training did not persist. Total flight time logged during the eight-month interval between trials was moderately correlated with reduction in task prioritization errors.

A pilot's ability to prioritize tasks for attention effectively is an important flying skill and is a primary component of concurrent task management (CTM) as defined by Funk (1991) and Funk et al. (2003). CTM is the process by which pilots selectively manage concurrent tasks by assessing and prioritizing them, allocating resources in order of priority, and continuously updating their prioritization scheme to complete the flight mission safely and effectively (Funk et al., 2003). A task prioritization error occurs when a pilot gives preferential attention to a lower priority task rather than to a task that should take higher priority with regards to flight safety (e.g., it is more critical, more urgent, or not being performed satisfactorily) (Funk, 1991; Funk et al., 2003). Although pilots generally practice effective task management, there are many instances in which failure to prioritize tasks properly has led to a potentially dangerous incident or even a fatal accident (Chou,

---

Madhaven, & Funk, 1996; Damos, 1997; Dismukes, Loukopoulos, & Jobe, 2001; Latorella, 1996; Raby & Wickens, 1994; Rogers, 1996; Schutte & Trujillo, 1996). Funk et al. (2003) suggested that CTM performance might be improved through training.

Background and Objective

Previous experimental analysis (Hoover and Russ-Eft, 2005) showed that pilots who participated in a short term training course (experimental group) had a 54% decrease in task management errors over a two week period of time compared to pilots who did not participate in the training (control group). However, long-term effects of that training were not analyzed.

The objective of this study was to determine the long-term training effect on single pilot task prioritization performance in simulated flight.

Method

Pilots who had participated in the previous study (Hoover and Russ-Eft, 2005) comprised the group for this study. Of the twenty-seven original participants, seventeen were available and were retested eight months later. Pilots flew a one hour simulated instrument flight on a Frasca 141 flight-training device (FTD). Task prioritization challenges and associated errors were defined based on CTM theory developed by Funk et al. (2003) and on procedures established by Hoover and Russ-Eft, (2005).

Participants

Hoover and Russ-Eft (2005) tested twenty-seven pilots from the Central Washington University (CWU) Aviation Department enrolled in the intermediate or higher stage of instrument training and who comprised a relatively homogeneous group with respect to their task prioritization performance. Eight months later, seventeen of those pilots were retested for this study. The seventeen pilots were not selected based on any specific criteria; rather they were the pilots still available for testing after eight months. Of those pilots retested, eight were from the previous study's control group and nine from the experimental group. All pilots had logged previous instrument time on the Frasca 141 FTD used in the experiment. Each pilot reported his or her flight experience, training, and FTD time during the eight-month interval since the previous experiment.

Flight Training Device

Simulated flights were flown on a Frasca 141 FTD configured as a normally aspirated single-engine fixed-gear aircraft using the same performance parameters that participants were accustomed to in airplanes and FTDs during their normal training. The avionics package included a Bendix/King stack with dual KY196 Communication radios, dual KN53 navigation radios, KDI 572 DME, KR 87 ADF, KT76A transponder, KMA 24 audio panel with marker beacons, and GNS430 IFR enroute and approach certified GPS. The FTD recorded all primary flight data including aircraft heading, altitude, airspeed, power settings, and position.

Procedure

Pilots flew simulated flights as per the CWU Standard Operating Procedures manual; all checklists, flow checks, and callouts were the same used in their normal flight training. Simulated flights were conducted in a line oriented flight training (LOFT) format. Because comparison of pretest and posttest error data from the previous study (Hoover and Russ-Eft, 2005) indicated no practice effect due to repeating the same LOFT scenario over a two-week period, this study used the same LOFT scenario eight months later. The scenario placed pilots in a high workload environment in Seattle Class B airspace and included radar vectors as well as pilot navigation, two precision instrument approaches, a multistage missed approach, and a holding procedure. For each flight, a Certified Flight Instrument Instructor (CFII) operated the FTD and acted as Air Traffic Control (ATC). The LOFT scenario was scripted with respect to ATC communications and procedures. Flights were observed and coded by a live observer who then verified error scores with the FTDs flight data records.

Prioritization Scheme

The task prioritization scheme defined by Hoover and Russ-Eft (2005) was used:

- Aviate task: Included all primary aircraft control inputs (pitch, power, yaw, and roll), operation of lift and drag devices (flaps), and operation of other primary aircraft systems that affected airspeed, altitude, climb or descent rate, and changes in lift, thrust, and drag.
- Navigate task: Included items related to the current and future position of the aircraft, including vectors, course intercepting and tracking, identification of intersections and waypoints, and programming and operating the GPS and other navigation radios.
- Communicate task: Included communications with ATC.

Definition of Task Prioritization Errors

Using the same procedure as the previous experiment, opportunities for 19 potential task prioritization errors were embedded at 13 challenge points throughout the one hour simulated flights. At each challenge point, the pilot was given an opportunity to divert his/her attention from a more important or more urgent task to a less urgent or less important task and had to decide which task was most critical to perform first. Types of prioritization errors included ignoring an aviate (flight control) task in order to navigate, classified as an aviate/navigate (A/N) error (n=7). If a pilot ignored an aviate task to perform a less important communication task it was classified as an aviate/communicate (A/C) error (n=7). The third type of error involved putting a less critical communicate task ahead of a more important navigate task, which defined a navigate/communicate (N/C) error (n=5).

Several prioritization challenges were embedded at a point in the flight where a pilot might make a task prioritization error, for example fixating on one task to the exclusion of another, and thus did not require any intervention. Others required

the CFII to act as ATC and call the pilot with information or instructions just before the pilot was leveling off or about to intercept course, or to cause a failure to a navigational facility or an aircraft system.

Performance criteria for determining if an error occurred was based on FAA-S-8081-4C Instrument Rating Practical Test Standards with respect to altitude, air-speed, heading, intercepting and tracking course, use of checklists, procedures, and ATC communications.

Results

Task prioritization error data were recorded as a frequency distribution of raw scores and converted to a ratio score (number of errors: total number possible). Pilots from the previous study's control group showed negligible change in their task prioritization performance after eight months. Pilots from the previous study's experimental group, although having shown a reduction in errors after participating in the short term training course, had task prioritization performance similar to the control group after eight months (Figure 1).
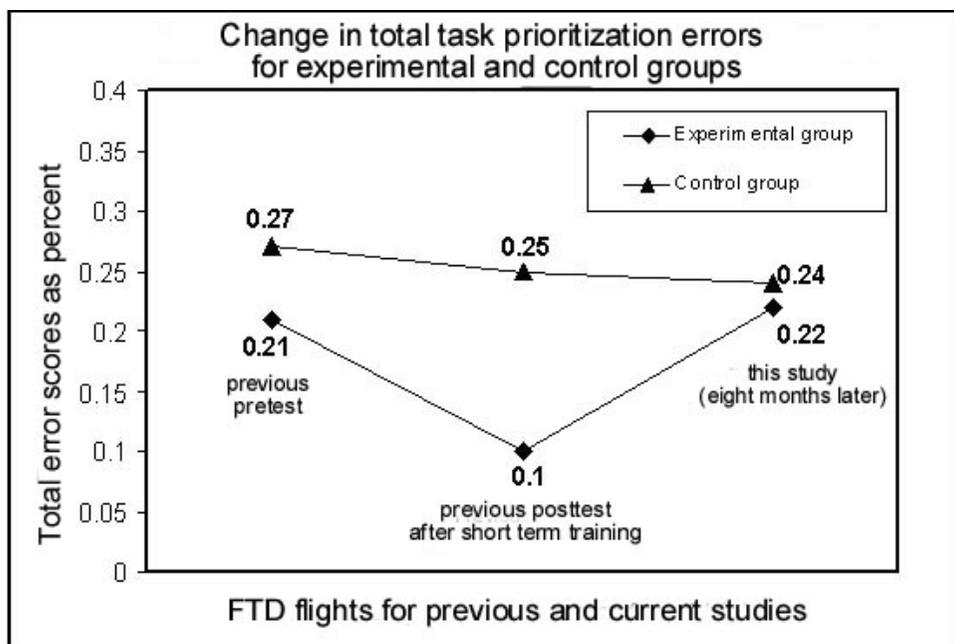
Figure 1. Change in total task prioritization error scores for experimental and control group pilots expressed as a percent of total possible errors. Pretest and posttest data from the previous study reflects scores before and after a two week time period in which experimental group pilots received training and control group pilots did not.

Figure 2 shows that the distribution of errors for pilots in the control group was within 5% of their scores from the previous study. Pilots in this group showed a slight decrease in A/N errors, which ranged from 19% to 22% in the previous study to 17% in this study. A/C errors showed a 5% decrease in this study from the pre-

vious study, and N/C errors showed a 2% to 5% increase from the previous study. Figure 3 shows that A/N errors for the experimental group pilots varied only 3% from their scores in the previous study. Although these pilots had a significant decrease (47%) in A/C errors after participating in the previous study's two week training course, eight months later they had A/C error scores within 1% of their original (pretest) scores. Error rates for N/C errors, which showed a 78% decrease after training in the previous study, were within 3% of their original (pretest) scores after eight months.
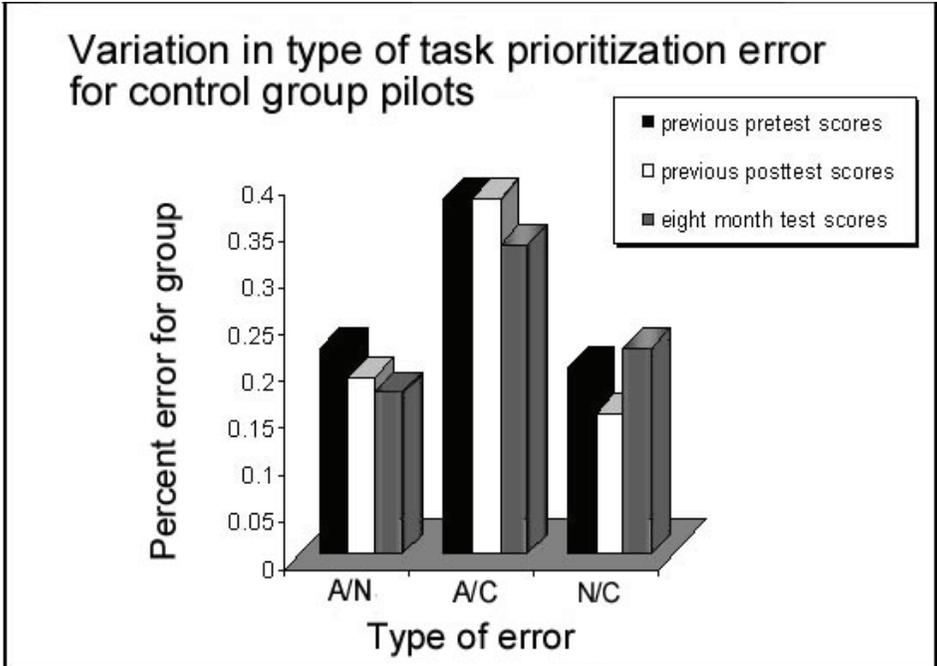
Figure 2. Variation in each type of task prioritization error for pilots in the previous study's control group. Pretest and posttest data from the previous study reflects scores before and after a two-week period in which experimental group pilots received training and control group pilots did not.
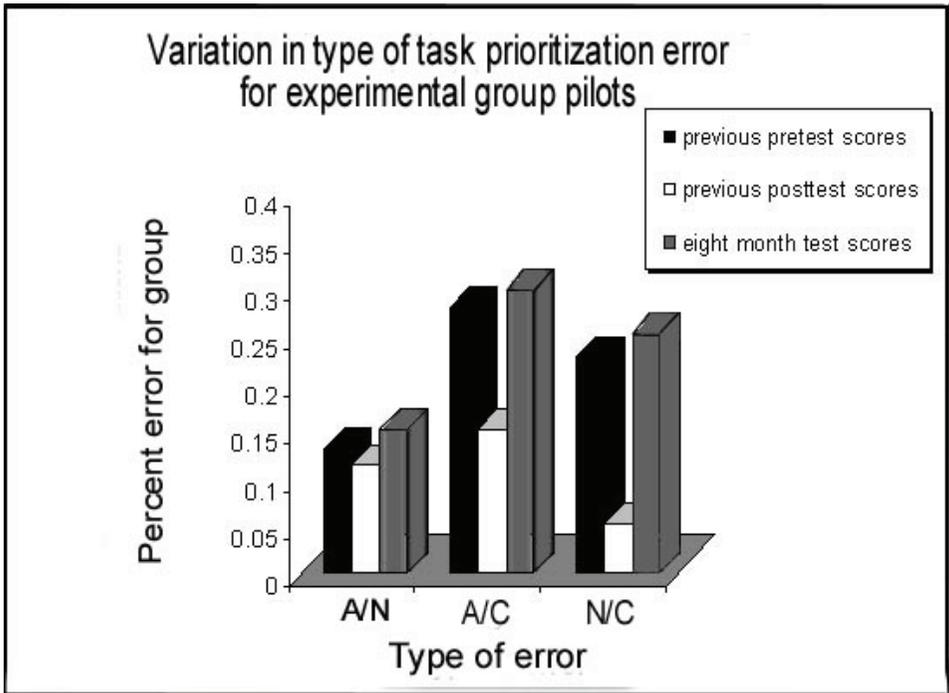
Figure 3. Variation in each type of task prioritization error for pilots in the previous study's experimental group. Pretest and posttest data from the previous study reflects scores before and after a two week period in which experimental group pilots received training and control group pilots did not.

The bivariate coefficient of determination interpretation of Newton and Rudestam (1999) was used to compare each pilot's total error score with the amount of total flight time, instrument time, time in all FTDs, and time in the Frasca 141 FTD they had logged during the eight month interval between the previous study and this study (Table 1). There was little to no correlation between the amount of instrument time or Frasca 141 FTD time logged during the eight month interval and pilots' error scores. Correlation between total time in all FTDs and error scores was low for the same period. Pilots' total time in flight logged during the eight month interval between studies showed a moderate correlation ($R^2 = 0.424$; Table 1).

Table 1
The coefficients of determination showed moderate correlation for total flight time logged indicating accrued flight time may be related to task prioritization performance.

Coefficient of determination for time logged during 8 month interval and task prioritization error scores			
	Total Flight Tme	Instrument Time	Total FTD Time	Frasca 141 Time
Pilot Error Scores	$R^2 = 0.424$	$R^2 = 0.084$	$R^2 = 0.197$	$R^2 = 0.010$

Discussion

In the previous study (Hoover and Russ-Eft, 2005) pilots in the experimental group, who received task prioritization training, showed a large decrease in total task prioritization errors between pretest and posttest flights compared to the control group, who did not receive training. Over the eight-month period, control group pilots' error scores were fairly constant with a mean ranging from 24-27%. After eight months, experimental group pilots' scores regressed toward a similar mean of 21-22%. In the previous study, these experimental group pilots showed a reduction in both A/C and N/C errors after short-term training, but after eight months, their error scores were similar to their original pretest scores. Thus, any training effect that previously occurred did not persist over the eight months between trials.

Learning is defined as a persisting change in behavior resulting from experience (Schunck, 2004) and entails codifying concepts from the short-term, or working, memory into long-term memory. In order to transfer information to long-term memory an individual must relate incoming information to concepts and ideas already in memory. The loss of the short-term training effect over time indicated by this study suggests that transfer did not occur for these pilots. Thus, the positive training effect demonstrated by Hoover and Russ-Eft's (2005) experiments could be attributed to a sensitization effect; experimental group pilots were focused on reducing their task prioritization errors during the short (two week) training session and did not actually retain the concepts in their long term memory.

Limitations

There was no control for extraneous variables over the eight-month period between the previous experiments and this study; variability in pilot training was not evaluated. Future studies are needed to ascertain the effect of variability in flight training on task prioritization performance.

Participants were selected for the previous study because they comprised a relatively homogeneous group with respect to their training and experience, and those tested in this study were a result of attrition of the previous group. Therefore, it is unknown whether results from this study can be generalized to a more variable group of general aviation pilots.

Recommendations

Several recommendations for future research are suggested based on the limitations just discussed. A time-series experimental design using one or more pilots could be used to evaluate longer term training effects. Experiments could be conducted with a different group of pilots and with a training course taught over a longer period. External validity issues could be addressed by comparing results between pilots who have different training backgrounds. A less homogeneous group of pilots might also be used to investigate whether pilots with varying levels of experience showed different training effects and if there were some specific level of experience at which pilots showed the greatest effect.

Conclusions

The positive training effect previously exhibited for these pilots did not persist over the longer term. Instrument time and time in FTDs during the eight-month interval did not seem to have an effect on any pilot's performance, but total flight time logged between trials was positively correlated with a reduction in task prioritization errors. These results beg the question as to whether or not an effective training course can be designed to facilitate learning over a longer period. More follow-up studies are needed to design and test longer term training courses and to evaluate the effects of that training.

Acknowledgements

References

Chou, C.D., Madhavan, D., & Funk, K.H. (1996). Studies of cockpit task management errors, *International Journal of Aviation Psychology, 6*(4), 307-320.

Damos, D. L. (1997). Using interruptions to identify task prioritization in part 121 air carrier operations. Proceedings of the *Ninth International Symposium on Aviation Psychology*(4), Columbus, Ohio State University.

Dismukes, R. K., Loukopoulos, L. D., & Jobe, K. K. (2001). The challenges of managing concurrent and deferred tasks. In R. Jensen (Ed.), *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus: Ohio State University.

Funk, K. H. (1991). Cockpit task management: Preliminary definitions, normative theory, error taxonomy, and design recommendations. *The International Journal of Aviation Psychology. 1*(4), 271-285.

Funk, K. H., Colvin, K, Bishara, S., Nicolalde, J., Shakeri, S., Chen, J. Y., & Braune, R. (2003). *Training pilots to prioritize tasks: Theoretical foundations and preliminary experiments.* NASA Grant NAG 2-1287 Final Report. Corvallis, Oregon.

Hoover, A.L. & Russ-Eft, D.F. (2005). Effect of Concurrent Task Management Training on Single Pilot Task Prioritization Performance. *International Journal of Applied Aviation Studies*, *5*(2), 233-252. Oklahoma City; FAA.

Latorella, K. A. (1996). Investigating interruptions: An example from the flightdeck. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting.* 249-253.

Newton, R. R., & Rudestam, K. E. (1999). Y*our statistical consultant: Answers to your data analysis questions.* Thousand Oaks, CA: Sage.

Raby, M., & Wickens, C. D. (1994). Strategic workload management and decision biases in aviation. *The International Journal of Aviation Psychology. 4*(3), 211-240.

Rogers, W. H. (1996). Flight deck task management: A cognitive engineering analysis. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*, 239-243.

Schutte, P. C., & Trujillo, A C. (1996). Flight crew task management in non-normal situations. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting.* 244-238.

Who Has Air Travel Stress?
Demographic Differences in Air Travel Stress

Jonathan B. Bricker

University of Washington

Abstract

To learn who experiences air travel stress, this study examined hypothesized demographic differences in the three components of the Air Travel Stress Scale: Air Travel Anxiety, Air Travel Anger, and Airline/Airport Trust. In two samples (N= 925; N = 2382) of air travelers, results showed that statistically significant indicators of higher Air Travel Anxiety were being a woman, being under age 54, having a gross annual household income of less than $60,000 US dollars, being a leisure traveler, and having eight or fewer domestic roundtrip flights in the past year. Indicators of higher Air Travel Anger were being a man, having a gross annual household income of at least $60,000, and being an international traveler. Finally, indicators of lower Airline/Airport Trust were being a man, being age 54 or younger, an income of at least $60,000, being a business traveler, and being an international traveler. Implications for understanding and intervening on air travel stress are discussed.

There were over 2.1 billion passengers boarding the world's scheduled airlines in 2006 (IATA, 2007). Despite the large and growing number of air travelers around the globe, many of whom fly out of necessity, little research has been conducted on the stress associated with taking a flight. For a number of years, and increasingly since the September 11, 2001 terrorist hijacking of US airliners, hundreds of media reports worldwide have anecdotally described the stresses of air travel, ranging from the hassles of long airport security lines to threats of airline hijackings or bombings. Moreover, media reports have suggested that air travel stressors such as airport crowds, flight delays, and cancellations have important consequences: they may lead some people to experience work-related stress or avoid flying altogether (see, for example, "Crowded skies," 2004; Rayner, 1998; Sharkey, 2000; Trucco, 2003; Zoglin & Donnelly, 2002).

The unique nature of air travel suggests that an empirically-tested measure of stress specific to air travel is needed. Recently, Bricker (2005) reported on the first known measure of air travel stress, the Air Travel Stress Scale (ATSS). He showed that the ATSS measures three components: (a) Air Travel Anxiety: anxious reactions to adverse air travel events, (b) Air Travel Anger: angry reactions to other passengers, and an antecedent of air travel stress, and (c) Airline/Airport Trust: the lack of trust that the airlines/airports will ensure one's comfort and safety. Each component showed good internal reliability and test-retest reliability over a 6 to 7 week interval. In addition, each component showed good evidence for discriminant and convergent validity when each was correlated with measures of stress, coping, and personality traits. The current study follows up on the Bricker (2005) study by examining demographic differences in scores on the three dimensions of the ATSS.

The Value of Examining Demographic Differences in Air Travel Stress
Examining demographic differences in air travel stress is important for a number of scientific reasons: first, such study could lead to a better understanding of the kinds of individuals who are most likely to experience this stress; second, findings from this inquiry can help generate testable hypotheses about why certain groups of people experience air travel stress more than others; third, such an investigation could also provide further evidence for the construct validity of the ATSS by showing that specific dimensions of air travel stress are related to specific demographic factors in conceptually meaningful ways.

Studying demographic differences in air travel stress is also relevant for practicing psychologists. For example, psychologists who specialize in treating individuals with work stress (Portello & Long, 2001), flying phobias (Bor, Parker, & Papadopoulos, 2001; Bor & van Gerwen, 2003), driving anger (Deffenbacher et al., 2000), and a variety of other life stresses are all pertinent psychologists who could develop a specialty in assessing and intervening on air travel stress since these problem areas are conceptually and empirically related to air travel stress (Bricker, 2005). Information on demographic differences could help these psychologists in their interventions. For example, in a clinical setting, psychologists could communicate this information to clients who belong to high-scoring demographic groups in order to help normalize their air travel stress and stimulate the identification of specific strategies for coping with this stress. In a worksite setting, psychologists could use this information to readily-identify traveling employees who may be susceptible to developing air travel stress in the future. An intervention to prevent air travel stress might then be specifically tailored to the needs of these subgroups.

Hypothesized Demographic Differences in Air Travel Stress
Gender. A number of demographics factors may be associated with air travel stress in theoretically meaningful ways. One demographic factor may be gender. Feingold (1994) in a meta-analysis found that males scored higher than females on anger whereas females scored notably higher than males on anxiety and trust. Based on this prior empirical evidence (Feingold, 1994), it is hypothesized that women would report higher levels of Air Travel Anxiety and Airline/Airport Trust whereas men would report higher levels of Air Travel Anger.

Age. There is empirical evidence that younger people score higher on anxiety measures than older people (Twenge, 2000). It is therefore hypothesized that younger travelers are more likely to experience Air Travel Anxiety than older travelers. In contrast, little attention has been given to age differences in the expression of anger, with one study suggesting that older people express less anger than younger people do (Thomas, 2002). Consistent with the hypothesis that older people have less intense anger or learn to manage it more constructively with age and experience (Thomas, 2002), it is expected that older travelers will express less Air Travel Anger than younger travelers. Finally, little attention has been given to age differences in trust, with one study showing that older people show more trust than younger people (Ho, 2005). Older people are hypothesized to have more Airline/Airport Trust than younger people.

Income. Considerable evidence has shown that lower socioeconomic status (SES), assessed with measures of educational attainment, income, and occupational status, tends to be associated with a higher prevalence of psychiatric disorders among children, adolescents, and adults (e.g., Kessler, Foster, Saunders, & Stang, 1995). Although air travel stress is not a psychiatric illness, this association is relevant because air travel stress is conceptualized as a psychologically distressing experience. In the context of air travel stress, it is likely that people with higher levels of income will report lower levels of Air Travel Anxiety and Air Travel Anger. Finally, those higher incomes may have greater expectations of the airlines and airports to ensure their comfort and safety. They may be more easily disappointed when airlines and airports do not meet their expectations.

Business vs. leisure traveler. Several demographic factors describing an individual's air travel patterns may be related to air travel stress. Because no prior studies have explored the relationship between these factors and *any* kind of stress, much less air travel stress, several speculations can be made. One of these travel pattern factors includes whether a person is primarily a business or a leisure traveler. First, a business traveler has more at stake when he/she flies. A cancelled flight may mean missing an important meeting and perhaps a lost business opportunity. Therefore, business travelers probably learn more quickly than leisure travelers about ways to anticipate and solve these kinds of problems. Business travelers are probably less likely to experience Air Travel Anxiety. For instance, business travelers may be more likely than other travelers to cope with flight delays by re-booking themselves on another flight. On the other hand, business travelers would probably have more difficulty trying to change other passengers' behavior. Thus, if business travelers have a tendency to try to fix even those situations they cannot fix (e.g., other passengers' rude behavior) then they may be more likely to experience Air Travel Anger. Regarding Airport/Airline Trust, a business traveler probably would expect more from the airports and airlines. A business traveler depends on them for his/her livelihood. Thus, minor failures by the airlines and airports (e.g., news of baggage screeners allowing potentially harmful objects to pass through security) would probably quickly erode trust in their ability to ensure service and safety. Therefore, being a business traveler is probably a marker of low Airport/Airline Trust.

Domestic travel. More frequent domestic air travel in the past year should probably be a marker of lower Air Travel Anxiety, Air Travel Anger, and Airline/Airport Trust. By virtue of their more frequent exposure to air travel environments, these travelers probably have learned ways to manage their anxious reactions to adverse air travel events and their annoyance with behavior of other passengers.

International travel. Data show that airports and airlines in Western Europe and parts of Asia are perceived by travelers to have higher security and service standards than airports and airlines in the United States (IATA, 2007). Thus, by comparing usual experience with airlines and airports in the US with those in these countries, the US-based international traveler would probably have less trust in the US airlines/airports' abilities to ensure his/her comfort and safety.

In summary, this study hypothesizes the following about demographic differences in ATSS components:

Gender: Women will report higher levels of Air Travel Anxiety and Airline/Airport Trust whereas men will report higher levels of Air Travel Anger.

Age: Compared to younger travelers, older travelers will report less Air Travel Anxiety, Air Travel Anger, and more Airline/Airport Trust.

Income: Travelers with higher income will report lower levels of Air Travel Anxiety, Air Travel Anger, and Airline/Airport Trust.

Business vs. leisure travel: Compared to those who fly mainly for leisure trips, those who fly mainly for business will report lower levels of Air Travel Anxiety and Airline/Airport Trust but higher levels of Air Travel Anger.

Domestic air travel frequency: More frequent domestic air travelers will report lower levels of Air Travel Anxiety, Air Travel Anger, and Airline/Airport Trust.

International air travel: International travelers will report lower levels of Airline/Airport Trust, Air Travel Anxiety, and Air Travel Anger.

All of these hypotheses will be tested in this study with two independent samples of air travelers.

Method

Participants

Participants for the first sample were 925 (72.3% recruitment rate; 925/1280) Seattle-Tacoma International Airport travelers (herein referred to as "Sample A") who were surveyed in January/February 2002 via an anonymous paper and pencil self-report survey packet. The demographic characteristics of this sample were 54.3% male, average age of 41, 56.0% had a gross annual household income of at least $60,000, 50.8% flew for business at least half time, their median number of domestic roundtrips within the past 12 months was three, and 46.7% had taken at least one international roundtrip in the past 12 months.

Because of the potential that results from one sample may not replicate in another sample, the analyses were repeated with a separate sample of air travelers. Participants for this second analysis were a nationwide sample of 2382 (recruitment rate incalculable) self-identified air travelers who completed demographic measures at baseline (April 2002) and the ATSS six to seven weeks later. For simplicity, this sample will be herein referred to as "Sample B." The demographic characteristics of this sample were 38.2% male, average age of 39, 59.1% had a gross annual household income of at least $60,000, 47.0% flew for business at least half time, their median number domestic roundtrips within the past 12 months was four, and 39.4% had taken at least one international roundtrip in the past 12 months. Compared to the Sample A, Sample B included: (a) 16.1% fewer male participants, (b) participants who were 2 years younger on average, (c) 3.1% more participants with an income of at least $60,000, (d) 3.8% fewer business travelers, (e) participants with one less median number of domestic roundtrips, and (f) 7.3% fewer percent of participants who had taken an international roundtrip within the past 12 months.

Measures

Demographics. Self-report survey of participants' gender, age, gross household yearly income, usual purpose of flying within the past 12 months (i.e., business versus pleasure), number of domestic (within the US) roundtrips within the past 12 months, and number of international roundtrips within the past 12 months. The average of the Pearson correlations between each of these demographic variables was .16 in Sample A and .21 in Sample B.

Air Travel Stress Scale. The directions for the survey are: "A number of statements which people have used to describe their experiences of air travel are given below. Please select the rating that BEST indicates the extent to which you agree with the following statements. There are no right or wrong responses." Respondents rate the extent to which each item is true for him/her on a response scale ranging from 0 ("completely disagree") to 5 ("completely agree"). Sample items from each of the three scales are as follows: "I fear that I will miss a connecting flight" (Air Travel Anxiety), "I would feel resentful if I had to sit near loud/talkative passengers" (Air Travel Anger), and "I trust the airlines" (Airline/Airport Trust). The scale was labeled "Air Travel Experiences Scale" in order to help prevent revealing the intent of the survey and thereby reduce response bias. For Sample A, the means, standard deviations, and alpha reliabilities, respectively, were 1.89, .88, and .79 for Air Travel Anxiety, 2.10, .90, and .71 for Air Travel Anger, and 3.20, .80, and .82 for Airline/Airport Trust. For Sample B, the means, standard deviations, and alpha reliabilities, respectively, were 2.13, .93, and .83 for Air Travel Anxiety, 2.48, .87, and .73 for Air Travel Anger, and 2.57, .96, and .88 for Airline/Airport Trust. The correlation between Air Travel Anxiety and Air Travel Anger was .49 (p < .001) in Sample A and .45 (p < .001) in Sample B, Air Travel Anxiety and Airline/Airport Trust was -.27 (p < .001) in Sample A and -.32 (p < .001) in Sample B, and the correlation between Air Travel Anger and Airline/Airport Trust was -.21 (p < .001) in Sample A and -.30 (p < .001) in Sample B.

Procedure

For Sample A, data were collected by trained data collectors that solicited air travelers' participation by verbal request. Potential participants were approached sequentially (one-by-one) where they were seated at a random sample of gate departure lounges, baggage claim areas, and airport concessions (e.g., restaurants and massage service). Although these methods of approaching potential participants were intended to result in a representative sample of those who were in the airport on the occasions of data collection, there is a reasonable potential for selection bias.

When a traveler agreed to participate in the 15-20 minute survey while at the airport, the data collectors gave the participant an informed consent form and the survey packet. When a participant completed the survey packet, a data collector took the packet and thanked the participant for being in the study. Participants were not compensated. The 105 individuals who elected to complete the survey later were given a self–addressed–stamped envelope to mail in their survey. Forty-three percent of these individuals actually mailed in their survey. There was no evidence that these participants significantly differed ($p > .05$) from the rest of the sample on demographic characteristics.

For Sample B, data were collected via a website. In response to a University of Washington press release describing the data collection website, the following media sources wrote articles about the study that invited potential participants: *Atlanta Journal Constitution, Frequent Flyer Magazine*, and *MSNBC*. Data collection occurred over a period of two weeks.

The website was secured by 128-bit encryption. Participants provided their e-mail addresses so that they could participate in a brief follow-up survey for the test-retest reliability analysis. The e-mail address was not linked to their survey data. Anyone entering the website address was allowed to participate in the study. After agreeing to an online Information Statement at the beginning of the website, participants completed the demographics survey. All participants were entered in a drawing for one of ten $50 cash prizes. Participants who completed the survey multiple times were detected by the recording of the IP address of the computer that submitted the responses. All of the survey responses provided by these 28 participants were deleted. Participants for this sample provided their e-mail address. Because one purpose of this sample was to determine the test-retest reliability of the ATSS (reported in Bricker, 2005), six weeks after the completion of their initial data collection, a follow-up recruitment message was sent to the e-mail address the participants provided. To maximize the follow-up rate (72.0%), this e-mail was sent two times, seven days apart. Thus, the interval of time between completing the demographic survey at baseline and the ATSS at follow-up was six to seven weeks.

When participants logged into the website, they entered their e-mail address and the personalized password they created during the baseline survey. After they entered their password, they completed the ATSS. Several demographic questions were also asked to double-check the matching of their baseline and follow-up survey. (Follow-up surveys from nine participants whose surveys could not be matched were deleted.) Participants completing this follow-up survey were entered in a separate drawing for one of ten $100 cash prizes. All the data collection pro-

cedures were approved by the University of Washington Human Subjects Review Committee.

Statistical analysis. In both samples, six Multivariate Analyses of Variance (MANOVA) were calculated to test differences in the three ATSS scales' scores for each of the six demographic measures. For the first analysis, the comparison variable was gender. For the second analysis, the comparison variable was age. Specifically, four identical 15-year interval age groups (10-24, 25-39, 40-54, 55-69) were created in order to make the analyses of both samples readily comparable. For the third analysis, the comparison was whether the participants' gross household incomes were above the sample median of $60,000. For the fourth analysis, the comparison was whether the participants flew at least half time for business within the past 12 months. For the fifth analysis, the comparison variable was the participants' frequency of roundtrip domestic flights within the past 12 months. The domestic flight frequencies were divided into four quartiles (0-2, 3-4, 5-8, 9+ roundtrips). In order to make the analyses readily comparable, these four quartiles were the same in both samples. For the sixth analysis, the comparison variable was whether the participants had flown at least one round-trip international flight within the past 12 months.

The outcome variables were the three components of the ATSS: Air Travel Anxiety, Air Travel Anger, and Airline/Airport Trust. Effect sizes were calculated with the η^2, which is interpreted as the percentage of the outcome variable's variance that is explained by the comparison variable (Pedhazur, 1997). Post hoc comparisons were conducted using the Bonferonni procedure (Edwards, 1985), in which the alpha level of each pairwise comparison is divided by the total number of comparisons. These post hoc comparisons were conducted for the age group analysis and domestic flight frequency analysis because they included more than two groups. Consistent with Perneger (1998), all Bonferonni comparisons conducted with the Sample A were then confirmed in Sample B.

Results

Table 1 revealed a significant multivariate effect for gender in both Sample A ($F(4, 920) = 19.24$, partial $\eta^2 = .06$, $p< .001$) and Sample B ($F(4, 2377) = 51.78$, partial $\eta^2 = .10$, $p<.001$). In both samples, women reported higher levels of Air Travel Anxiety, lower levels of Air Travel Anger, and higher levels of Airline/Airport Trust. Note that for all the results, the scale scores ranged from 0 to 5, with higher mean scores reflecting higher levels of the given construct.

Table 1
ATSS Components by Gender

Measure	Men		Women		
	M	SD	M	SD	F
Air Travel Anxiety					
Sample A	1.78	.87	2.02	.87	15.75***
Sample B	1.87	.89	2.30	.90	101.28***
Airline Travel Anger					
Sample A	2.32	.93	2.13	.94	8.04**
Sample B	2.62	.99	2.43	1.01	14.90***
Airline/Airport Trust					
Sample A	3.23	.89	3.38	.78	6.30*
Sample B	2.50	.97	2.62	.95	6.08*

*p < .05, **p < .01, ***p < .001.

Table 2 revealed a significant multivariate effect for the four age groups in both Sample A ($F(4, 920) = 10.62$, partial $\eta^2 = .04$, $p<.001$) and Sample B ($F(4, 2377) = 12.44$, partial $\eta^2 = .03$, $p<.001$). In both samples, Bonferonni post hoc tests revealed that participants aged 55-69 reported significantly less Air Travel Anxiety than all of the younger age groups ($p<.01$). In both samples there were no significant differences in Air Travel Anger scores among the four age groups. Also in both samples, 55-69 year-olds reported higher levels of Airline/Airport Trust than all of the younger age groups ($p<.05$).

Table 2
ATSS Components by Age

Measure	Age 10-24		Age 25-39		Age 40-54		Age 55-69		F
	M	SD	M	SD	M	SD	M	SD	
Air Travel Anxiety									
Sample A	2.08	.89	1.93	.80	1.92	.92	1.65	.87	5.88*
Sample B	2.17	.93	2.21	.94	2.10	.90	1.89	.89	7.64*
Airline Travel Anger									
Sample A	2.21	.95	2.25	.92	2.26	.92	2.19	.92	.27
Sample B	2.39	1.06	2.48	1.03	2.54	.98	2.54	.97	1.54

Airline/Airport Trust								
Sample A	3.28	.78	3.17	.85	3.29	.84	3.56	.78	7.22*
Sample B	2.51	.94	2.47	.98	2.54	.92	2.74	.95	11.52*

*$p < .001$.

Table 3 revealed a significant multivariate effect for the two income groups in both Sample A ($F(4, 920) = 6.68$, partial $\eta^2 = .02$, $p<.001$) and Sample B ($F(4, 2377) = 21.91$, partial $\eta^2 = .05$, $p<.001$). In both samples, participants reporting gross annual household incomes of at least $60,000 reported significantly lower levels of Air Travel Anxiety and Airline/Airport Trust as well as significantly higher levels of Air Travel Anger than participants reporting household incomes of less than $60,000.

Table 3
ATSS Components by Income

Measure	Under 60K		Over 60K		F
	M	SD	*M*	SD	
Air Travel Anxiety					
Sample A	1.93	.89	1.83	.86	3.85*
Sample B	2.21	.93	2.06	.91	12.22***
Airline Travel Anger					
Sample A	2.15	.92	2.31	.94	6.11**
Sample B	2.37		2.58	.99	18.02***
Airline/Airport Trust					
Sample A	3.33	.81	3.25	.86	1.99*
Sample B	2.73	.95	2.48	.95	30.20***

*$p < .05$, **$p < .01$, ***$p < .001$.

Table 4 revealed a significant multivariate effect for whether some flew for business or pleasure in both Sample A ($F(4, 917) = 3.67$, partial $\eta^2 = .02$, $p<.05$) and Sample B ($F(4, 2370) = 24.00$, partial $\eta^2 = .06$, $p<.001$). In both samples, business travelers scored significantly lower on Air Travel Anxiety and Airline/ Airport Trust. Business travelers scored significantly higher on Air Travel Anger in Sample B whereas this same comparison was not significant in Sample A.

Table 4
ATSS Components by Type of Travel (Pleasure vs. Business)

Measure	Pleasure		Business		F
	M	SD	*M*	SD	
Air Travel Anxiety					
Sample A	1.93	.89	1.80	.88	3.84*
Sample B	2.22	.93	1.97	.92	23.41**
Airline Travel Anger					
Sample A	2.24	.93	2.23	.93	.249
Sample B	2.42	1.01	2.60	.99	10.61**
Airline/Airport Trust					
Sample A	3.35	.83	3.21	.88	4.27*
Sample B	2.68	.97	2.45	.97	19.88**

*$p < .05$, **$p < .001$.

Table 5 revealed a significant multivariate effect for the four domestic flight frequency groups in both Sample A ($F_{(4, 918)} = 7.67$, partial $\eta^2 = .03$, $p<.001$) and Sample B ($F_{(4, 2372)} = 43.99$, partial $\eta^2 = .09$, $p<.001$). In both samples, more frequent domestic travelers reported significantly lower Air Travel Anxiety. Bonferonni tests revealed that, in both samples, participants who took at least 9 roundtrips in the past 12 months scored significantly lower on Air Travel Anxiety than the three other, less frequent domestic traveler groups ($p<.01$). Only in Sample B did more frequent travelers score significantly higher on Air Travel Anger and significantly lower on Airline/Airport Trust. For Sample B, participants who took at least 9 roundtrips in the past 12 months scored significantly higher on both Air Travel Anger and Airline/Airport Trust than the three other, less frequent domestic traveler groups ($p<.01$).

Table 5
ATSS Components by Number of Domestic Roundtrips in the Past Year

Measure	0-2		3-4		5-8		9+		F
	M	SD	M	SD	M	SD	M	SD	
Air Travel Anxiety									
Sample A	1.99	.88	1.86	.90	1.88	.85	1.71	.88	4.27*
Sample B	2.33	.88	2.21	.91	2.10	.94	1.88	.90	20.54**
Airline Travel Anger									
Sample A	2.22	.93	2.21	.91	2.34	.93	2.22	.94	.57

Sample B	2.42	1.02	2.45	.98	2.47	1.01	2.67	1.00	6.22**
Airline/Airport Trust									
Sample A	3.35	.83	3.34	.77	3.35	.81	3.18	.95	1.46
Sample B	2.62	.99	2.64	.94	2.63	.97	2.40	.94	6.45**

*p < .01, **p < .001.

Table 6 revealed a significant multivariate effect for whether someone flew internationally in the past 12 months in both Sample A ($F(4, 920) = 3.65$, partial $\eta^2 = .01$, $p<.05$) and Sample B ($F(4, 2373) = 19.96$, partial $\eta^2 = .04$, $p<.001$). In both samples, international travelers scored significantly higher on Air Travel Anger and significantly lower on Airline/Airport Trust. However, international travelers scored significantly lower on Air Travel Anxiety in Sample A whereas this same comparison was not statistically significant in Sample B.

Table 6
ATSS Components by International Roundtrip Travel in the Past Year

Measure	No Roundtrips		1 or More Roundtrips		F
	M	SD	M	SD	
Air Travel Anxiety					
Sample A	1.86	.89	1.92	.86	.97
Sample B	2.22	.94	1.99	.89	27.84***
Airline Travel Anger					
Sample A	2.19	.96	2.29	.90	3.82*
Sample B	2.44	1.00	2.59	1.02	9.49**
Airline/Airport Trust					
Sample A	3.38	.83	3.20	.86	9.97**
Sample B	2.62	.95	2.50	.98	6.88*

*p < .05, **p < .01, ***p < .001.

Discussion

Summary of Results

This study showed a number of significant demographic differences in air travel stress that were consistent across both samples. Specifically, demographic indicators of higher Air Travel Anxiety were being a women, being age 54 or younger, having a gross annual household income of less than $60,000 USD, being a primarily a leisure traveler, and having eight or fewer domestic roundtrip flights in the past year. Consistent demographic indicators of higher Air Travel Anger were being a man, having a gross annual household income of at least $60,000, and being an international traveler. Finally, demographic indicators of lower Airline/Airport Trust were being a man, being age 54 or younger, income of at least $60,000, being a business traveler, and being an international traveler.

Evaluation of Study Hypotheses

As hypothesized, the results from both samples showed that *women* reported higher levels of Air Travel Anxiety, lower levels of Air Travel Anger, and higher levels of Airline/Airport Trust. These results were consistent with the sociocultural model of gender, which suggests that appropriate roles for women include more free expression of anxiety and trust; whereas, a more appropriate role for men is to express anger more freely (Eagly, 1987; Eagly & Wood, 1991). The results are also consistent with the possibility, articulated by Feingold (1994), that it is more socially desirable for women to express more anxiety and trust and less anger.

The results from both samples also showed that participants aged 55-69 reported significantly less Air Travel Anxiety and higher levels of Airline/Airport Trust than all of the younger age groups. The stress of air travel appears to be less of a problem for older people than younger people perhaps because older people are more experienced and accepting of the challenges of air travel. Also consistent with Twenge's (2002) hypothesis that younger people experience more anxiety than older people, this study suggests that younger travelers experience more air travel anxiety than older travelers. Future research should test that hypothesis, perhaps with other forms of anger and contexts for expressing anger. Finally, the results support the hypothesis that older people have more Airline/Airport Trust than younger people. One interpretation of this result is that older people have the wisdom of experience that gives them more realistic expectations of the airlines and airports.

Regarding income, participants in both samples who reported gross house-hold incomes of at least $60,000 appeared to have lower levels of Air Travel Anxiety. Contrary to what was hypothesized, those with incomes of at least $60,000 reported higher levels of Air Travel Anger than participants reporting household incomes of less than $60,000. Finally, there was support for the hypothesis that those with higher incomes would report lower Airline/Airport Trust. This result suggests that those with higher incomes also have greater expectations of the airlines and airport to ensure their comfort and safety.

Several demographic factors pertaining to individuals' air travel patterns were associated with air travel stress. In both samples, business travelers scored significantly lower on Air Travel Anxiety, a result that is consistent with the interpretation that business travelers may have more resources to respond to adverse air

travel events. Compared to leisure travelers, business travelers may know more effective ways to cope with the particulars of the air travel system (e.g., how to quickly re-book an airline ticket) or may be more able to afford access to services to help them travel more effectively (e.g., airline airport lounge club memberships). Business travelers also scored significantly lower on Airline/Airport Trust, a result that is consistent with the explanation that business travelers probably expect more from the airports and airlines because they depend on them for their livelihood. Minor failures of the airlines and airports to ensure a passenger's comfort and safety may erode business travelers' trust in the airlines and airports. Finally, there was no support for the hypothesized differences in business vs. leisure travelers' Air Travel Anger. Consistent with what was hypothesized, participants who took at least nine roundtrips in the past 12 months scored significantly lower on Air Travel Anxiety than the three other, less frequent domestic traveler groups. Perhaps because of their more recent and frequent exposure to air travel environments, these travelers have learned ways to deal with the situations that lead to Air Travel Anxiety. The results also suggest that eight or less yearly roundtrip domestic flights could be used as a rough marker of those will tend to experience higher levels of Air Travel Anxiety. In contrast, there was no support for the hypotheses about the role of frequent domestic travel in Air Travel Anger and Airline/Airport Trust.

A final demographic characteristic explored in this study was whether the participant flew internationally within the past 12 months. There was no consistent evidence to support the hypothesis that international travelers had less Air Travel Anxiety. Contrary to what was hypothesized, international travelers scored significantly higher on Air Travel Anger. As can be further explored in future research, international travelers may actually become more bothered by other passengers' rude behaviors. International travelers also scored significantly lower on Airline/Airport Trust. This finding is consistent with the data that airports and airlines in Western Europe and parts of Asia are perceived by travelers as having higher security and service standards than airports and airlines in the United States (IATA, 2007). Thus, by comparing usual experience with airlines and airports in the US with those in these countries, an international traveler may view the domestic airports and airlines as less capable of ensuring comfort and safety.

Finally, the results of Sample A and B were largely similar. The consistent results between the two samples provide for more robust conclusions because the methods of sample selection were different. In addition, the few instances in which the results were not consistent can be explained by the differences in sample selection.

In addition to the future research suggested already, other directions for future research include exploring (1) longitudinal changes in air travel stress using the ATSS, and (2) how these changes may be associated with longitudinal changes in the services provided by airlines/airports and their employees. Also valuable would be to explore the extent to which air travel stress may differ according the specific airlines or airports an individuals tends to use while traveling.

Limitations

The demographic differences in air travel stress had small to moderate effect sizes. These findings are nonetheless valuable because so little is known about demographic differences in air travel stress. A second limitation of this study is that all of the data were cross-sectional or had only a short-term (i.e., six to seven week) follow-up, thus preventing any causal interpretations from being drawn. For example, it is conceivable that air travel stress leads one to fly less or that flying less leads one to be more vulnerable to experience air travel stress. Third, because all of the data were self-report there was no external validation of the ATSS. Future research should consider further validating the ATSS by correlating it with physiological measures of stress (e.g., cortisol), similar to what has been recently done in the context of research on train commuting stress (Evans & Wener, 2006). Finally, people may misreport their anxiety, anger, and trust—perhaps to present themselves in a socially desirable way. This possibility was not too likely because the correlation between the ATSS and social desirability was low (Bricker, 2005).

Finally, the Internet-based Sample B may limit the generalizability of the findings. First, it was not possible to determine the response rate of these samples. Besides the self-selection that may have occurred in the samples for any number of reasons, participants in the Internet-based Sample B might have been highly self-selected. Internet access is greater among several demographic groups, especially younger people and those with higher incomes (Dillman, 2000). However, survey data indicates that 56% of the US population was using the Internet (Victory & Cooper, 2001)—a figure that is rising every year. More pertinent, Internet usage among air travelers is high: 85% of air travelers surveyed at the Seattle-Tacoma International Airport in Sample A reported that they used the World Wide Web within the past 30 days. To the extent that air travelers were able to participate in this study because they had access to the Internet, this paper's Internet-based sample may generalize to most air travelers.

Implications for Worksite-based Psychologists and Airport/Airline-Based Interventions

A psychologist in a worksite setting could use this information to help target specific components of air travel stress management skills interventions to the needs of certain demographic groups. For example, an intervention focused on managing Air Travel Anger could be tailored to men, those with higher incomes, and international travelers. In contrast, an intervention focused on Air Travel Anxiety could be tailored to women, those who are age 54 or younger, those with lower incomes, leisure air travelers, and infrequent domestic air travelers. Each targeted subgroup can be taught specific skills for how to cope with the particular component of air travel stress they experience.

The results of this research might also be valuable in *public health interventions*. For example, brief videos played at airport lounges or on commercial airplanes can be developed that illustrate the stresses of air travel and simple ways to cope with them better. These videos might better address the specific components of air travel stress by being tailored to the needs of demographic groups, which tended to score higher on those specific components.

Acknowledgements

References

Bor, R., Parker, J., Papadopoulos, L. (2001). Brief, solution-focused initial treatment sessions for clients with a fear of flying. Counselling *Psychology Review, 16*(4), 32-40.

Bor, R., & van Gerwen, L. (Eds.). (2003). *Psychological Perspectives on Fear of Flying*. London: Ashgate.

Bricker, J. B. (2005). Development and Evaluation of the Air Travel Stress Scale. *Journal of Counseling Psychology*, *52*(4), 615-628.

Crowded skies are here again. (2004, January 23). *New York Times*, A–22.

Deffenbacher, J. L., Huff, M. E., Lynch, R. S., Oetting, E. R., & Salvatore, N. F. (2000). Characteristics and treatment of high-anger drivers. *Journal of Counseling Psychology, 47*(1), 5-17.

Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method.* New York: Wiley.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Eagly, A. H., & Wood, W. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin, 17*(3), 306-315.

Edwards, A. L. (1985). *Experimental design in psychological research*, 5th ed. New York: Holt, Rinehart, and Winston.

Evans, G. W., & Wener, R. E. (2006). Rail commuting duration and passenger stress. *Health Psychology*, *25*(3), 408-412.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, *116*(3), 429-456.

Ho, G. (2005). Age differences in trust and reliance of a medication management system. *Interacting with Computers, 17*(6), 690-710.

International Air Transport Association (IATA). (2007). *Global Airport Monitor*.

Kessler, R. C., Foster, C. L., Saunders, W. B., & Stang, P.E. (1995). Social consequences of psychiatric disorders I: Educational attainment. *American Journal of Psychiatry, 152*(7), 1026-1032.

Kite, M. E., Deaux, K., & Miele, M. (1991). Stereotypes of young and old: Does age outweigh gender? *Psychology & Aging, 6*(1)*,* 19-27.

Pedhazur E. J. (1997). *Multiple regression in behavioral research*. Third Edition. Forth Worth: Harcourt Brace.

Perneger, T.V. (1998). What is wrong with Bonferroni adjustments. *British Medical Journal, 316*(7139), 1236-1238.

Portello, J.Y., & Long, B. C. (2001). Appraisals and coping with workplace interpersonal stress: A model for women managers. *Journal of Counseling Psychology, 48*(2), 144-156.

Rayner, R. (1998, March 8). Nowhere, U.S.A. *New York Times Magazine*, Section 6, 42–46.

Sharkey, J. (2000, January 12). Business travel: On the road; Try to relax. At least two psychologists are paying attention to stress from life on the road. *New York Times*, p. C-8.

Thomas, S. (2002). Age differences in anger frequency, intensity, and expression. *Journal of the American Psychiatric Nurses Association, 8*(2), 44-50.

Trucco, T. (2003, December 21). Travel Advisory: Survey finds rudeness a problem for travelers. *New York Times*, Section 5, p. 3.

Twenge, J. M. (2000). The age of anxiety? Birth cohort change in anxiety and neuroticism, 1952—1993, *Journal of Personality and Social Psychology, 79*(6), 1007-1021.

Victory, N. J., & Cooper, K. B. (2001). *A nation online: How Americans are expanding their use of the Internet*. Washington D.C.: U.S. Department of Commerce.

Zoglin, R., & Donnelly, S. B. (2002). The nation's best run airport—and why it's still not good enough. *Time Magazine, 160*, 46–52.

Using Probability and Set Theory to Examine Illustrations of Situation Awareness

Todd P. Hubbard

Oklahoma State University
318 Willard
Stillwater, OK 74078
todd.hubbard@okstate.edu

Abstract

To some degree many collegiate aviation students have been misinformed on the nature of situation awareness (Dekker, 2005). Situation awareness was a construct developed by cognitive psychologists to help further examine the interior workings of the processes at play. However, if one were to study situation awareness with Endsley (2000), or Pew (2000), or Dennehy and Deighton (1997), the student might come away with misunderstandings of how this construct works: not because these scientists are misinformed, but because their illustrations must be presented with warning labels. Not all illustrations are accurate. This is particularly true when a Venn diagram or set theory has been used to illustrate situation awareness. This paper examined several of the more popular models of situation awareness and transformed them into five rules of probability and the general propositions of set theory. Useful tips on how to properly present situation awareness in the classroom were added, as well as thoughts on how to observe and study situation awareness during simulator training.

Based on the interactionist model of situation awareness, where Person and Environment (P-E) exist together in operational space, the Person (pilot) is described in terms of the subjective (P_s) and the objective self (P_o), and the Environment (flight environment) is described as being both subjective (E_s) and objective (E_o) (Dennehy & Deighton, 1997). Combinations of Person and Environment reveal two views of actual events. One can view pilot performance and behavior in terms of what he or she actually did in the flight environment, or by what the pilot could have done, given his or her capacity to perform or behave differently in a flight environment. The interactionist approach is embraced by or at least

applauded by Pew (2000), Uhlarik and Comerford (2002), and Endsley and Rodgers (1997).

Given the theory's preoccupation with the subjective and objective self, Dekker (2005) cannot agree with the interactionist model when it is used to explain pilot behavior in actual events, such as aircraft accidents (Hollnagel, Kaarstad, & Lee, 1999). To suggest that the pilot's objective self—that which is derived from past flights, simulator sessions, tests, or general knowledge evaluations—is an indicator of a pilot's potential, might be relevant when deciding to upgrade the pilot to the Captain's chair, but potential ability becomes meaningless in actual events. Even superior pilots, those who have the most potential to perform flawlessly, can be caught off guard or miss vital cues that lead to poor decisions and an unfortunate end of the flight (Dismukes, Berman, & Loukopoulos, 2007).

Illustrations showing the interaction between capacity (objective self) and actual displays of skill and knowledge (subjective self) have provided some insight into how the subjective interacts with the objective in both Person and Environment. Those presenting the illustrations, such as Dennehy and Deighton (1997), Harris (1997), and Pew (2000), have strained to make the connection between the subjective and objective. These illustrations have encouraged many to conclude that the proper appraisal of causation is performed after the fact, when details about pilot capacity are made known. Perhaps the most damaging way these illustrations have been used, is in depicting the construct situation awareness (SA) (see Figure 2).

Pew (2000) shows how ideal, achievable, and actual situation awareness events can be construed as being overlapping, intersecting, or separate and distinct. Although conceptual modeling, like that illustrated by Pew, helps one understand the interaction of person with environment, there is something fundamentally not helpful about the illustration. And this fundamental flaw is not exposed so well as by set theory and the visually intuitive Venn diagram.

Purpose
The purpose of this paper is to use set theory and Venn diagrams to examine illustrations of SA and to shed light on better ways to explain the concept. Three illustrations of SA (Dennehy & Deighton, 1997; Pew, 2000), two based on an interactionist model and one based on information processing theory, were examined to see if they could be conformed to or transformed by the rules of set theory and statistical probability (see Appendices A & B). Figures, approximating these illustrations, are shown in this paper after the reference list, for two reasons: (1) the visual sequence helps the reader process the illustrations in a logical way, and (2) they do not litter the narrative and draw attention away from technical arguments that require more concentration. After explanations of some of the logical-mathematic expressions, the reader will find a practical example of how set theory and Venn diagrams can be used to examine NTSB findings as to proximate cause after an aircraft accident has been investigated. Toward the end of the paper the reader will find useful tips for classroom instructors, on how better to illustrate situation awareness, and how to incorporate the manifestations of SA in classroom and simulator sessions.

How Should Situation Awareness be Modeled?

Beneficial models of situation awareness do not have within them any reference to SA, because scientifically speaking SA does not exist. Endsley's (2000) three-level model (perception, comprehension, and projection) is such an example, where SA is not used to define itself. One can see in Endsley's dynamic model of SA interactions between the levels and among the cognitive sub-elements. These fairly represent how a pilot might perceive him or herself in space-time at the present and perceive him or herself at another location in space-time in the future (Durso et al., 1999).

Endsley and Rodgers (1994; 1997), like Dennehy and Deighton (1997), Pew (2000), and Harris (1997), were all influenced by Neisser's perceptual cycle, an idea made popular in the early 1970s. Neisser found evidence of interactions between Person and Environment, which according to him affected the Person's perception of his or her place in space and time (Uhlarik & Comerford, 2002).

Early work with U.S. military pilots (Bell & Lyon, 2000; Endsley & Rodgers, 1994) created a ground swell of interest in situation awareness, which led to widespread use of the term to describe all sorts of events, from getting lost to being "behind the aircraft," (a breakdown caused by a lack of anticipation). By the late 1990s, cognitive psychologists were successful in gaining control over how SA would be studied and described, but not everyone accepted the more accurate cognitive explanations. Military pilots continued misusing the term, even though new evidence was available. In the classroom, those professors with special abilities in cognitive psychology or those instructors having formal training in cognitive learning theory altered their lesson plans to accommodate fresh perspectives of situation awareness. No other aviation course was more influential on how pilots perceived situation awareness than crew resource management (CRM). In CRM courses an instructor has the benefit of presenting the theory behind SA and soon afterwards observing the gaining of or losing of SA in the simulator. Not all CRM instructors were aware of the multifaceted view of SA, and hundreds of students graduated without ever knowing the complexity in the concept.

Cognitive psychology is still relatively new, compared to other sciences, having gained a foothold in the 1970s and having unseated behavioral psychology as the dominant view, but even though cognitive psychologists are just getting started, the subject of situation awareness has taken center stage at many conferences and symposia. Despite the subject's popularity, little new has been discovered since the 1990s. There have been comments made on the subject, but actual studies that examine the sub-elements of SA are fewer in number. This makes the 22-item list of situation awareness characteristics by Dennehy and Deighton (1997) even more important (Table 1). Having achieved a 35% agreement on the 22 items or events that characterize situation awareness, one can be relatively sure that these items are indeed part of the larger construct of situation awareness. However, those conducting research in SA cannot declare victory just yet. Thirty-five percent is a good start.

Table 1

Sub-elements of SA (adapted from Dennehy & Deighton, p. 286)

1. Memory	12. Overall Awareness
2. Decision making ability	13. Perception
3. Awareness of the environment	14. Safe flying
4. Present-state knowledge	15. Knowledge of procedures
5. Knowledge of goals	16. Air-performance awareness
6. Knowledge of aircraft systems	17. Attention
7. Spatial awareness	18. Mental picture
8. Workload	19. Course training
9. Stress	20. Future-state knowledge
10. Time perception	21. Information processing
11. Cockpit resource management	22. Aircrew attitude

These items are indirectly confirmed by Endsley (2000) in her three-level model and they were valued by Uhlarik and Comerford (2002) in their critique of the various definitions and descriptions of situation awareness. Having the stamp of approval from Endsley, and Uhlarik and Comerford is important, because these researchers have been commissioned to perform studies for the Federal Aviation Administration (FAA) in conjunction with human factors researchers from the Civil Aerospace Medical Institute on the Mike Monroney Aeronautical Center campus in Oklahoma City, Oklahoma. This collaboration with the FAA has benefited the aviation community by exposing situation awareness theory and practice on the international stage.

As depicted in Figures 2-4, Pew's Venn diagram of situation awareness displays three views of SA: the ideal, the achievable, and the actual (2000). Pew confirmed that his depiction was an intersecting Venn diagram, and not just an illustration of a theoretical point. If one then assumes that the Venn diagram Pew used conforms to the logic of set theory and statistical probability, then by these rules one should be able to create intersections and unions of the ellipses shown. However, Pew's Venn illustration fails on several levels, even though the narrative before and after Pew's diagram (2000, p. 36) presents a logical argument for situation awareness.

Straight away, one can dismiss the diagram on the basis of falsifiability (more on falsifiability later). Each event in a Venn diagram must be falsifiable, and since situation awareness is a construct and is not falsifiable, then it does not fit within the rules of set theory nor can it be logically argued.

After a second look, the diagram reveals the interactionist relationship between the objective and subjective, with the terms ideal, achievable, and actual. Even

though Pew did not use the term interactionist when he labeled his Venn diagram, he did use terms that are quite interactionist, such as achievable and actual. Achievable SA is what the objective self could do under different circumstances. Actual SA is what the subjective self did. Beside these interactionist tags one also sees Ideal SA. As will be discussed, the ideal is unhelpful for hypothesis testing and for determining causal factors in aircraft accidents (Dekker, 2005; Dismukes et al., 2007).

Ideal SA could be of some benefit to scientists, if one could be sure that all the elements of SA were included in that ideal. Since scientists are limited in perspective by their own human ability and the combined abilities of all other humans, no one could be sure that the ideal was achieved. Scientific method does not require that we find facts, only indications that we are on the right path. This methodology has been in operation since the early 17th century, being expressed and bounded by Descartes, Pascal, and Newton (Kasser, 2006; Kreeft, 2004). Therefore, the ideal must remain just that, the ideal, and not connected to Achievable and Actual SA. Pew's depiction of the ideal is nothing more than theoretical conjecture, much like a statistician saying that for all sets in Space S there must be at least one null set. One must agree that the ideal is *out there* in the universe, but while studying situation awareness in pilots, the ideal remains hidden.

Before moving on, falsifiability must be better defined. As with all constructs, each resists being falsified. A hypothesis cannot be created with an indivisible construct of terms, the overall construct can be changed by any of its constituent parts. If a term can be sub-divided into its constituent parts, then one must examine how the sub-parts interact before declaring how the construct has changed. Given the list of 22 elements (Table 1) from Dennehy and Deighton (1997), to properly test a hypothesis on situation awareness, perhaps as many as 22 alternative hypotheses would have to listed; thus, realizing at the same time that changes to any one hypothesis could alter all the other outcomes.

Here are a few thoughts shared by Kasser and Hall on the topic of falsifiability. Kasser (2006), Professor of Philosophy at North Carolina State University, and Hall (2005a; 2005b), Professor Emeritus of Philosophy from the University of Richmond, both agree that the dividing line between science and pseudo-science is drawn on the boundary between empiricism and meta-physics. Both professors pointed to Popper's work on falsification in the 1950s as a means to understand the benefits of logical positivism, which for this paper is a theory made relevant and observable through set theory and logical syllogisms. One cannot test a hypothesis unless the hypothesis presents an element which can be observed and measured and ultimately declared to be true or false. The scientist, at the end of a period of observation and study, must be able to say whether the hypothesis is true or the hypothesis is not true. One cannot scientifically approach the subject of SA unless situation awareness can be observed, tested, and falsified. Situation awareness, in the form presented by Pew, cannot be observed and tested. In fact, what Pew presents are outcomes of the interaction of the constituent elements of SA, not the action of SA.

Set Theory and Probability as Criteria for an Examination

Statistics and Philosophy are in the same family of academic inquiry, particularly when it comes to logical arguments concerning probability theory (Hays, 1994). Both statisticians and logicians use Venn diagrams to visually represent the relationships between events and the space (*S*) within which these relationships operate (Figure 1). These technicians do this because visual illustrations of set theory can show at what point two events form an intersection, a union, or even if two elements operate within a third event, such as Pew's ill-fitting illustration (Figure 2) of SA (Pew, 2000).

If a Venn diagram were to be written out in mathematical notation (see Appendix A) in the absence of the visual illustration, the intuitive nature of relating elements and events might be lost to the viewer, unless abstract mathematical ideation suits you (see Appendix B). However, if both the illustration and the mathematical notation were presented together, the intuitive logic of the Venn diagram and the mathematical description would form a matching association which could be easily translated into instructional and testing strategies in the classroom (see Appendix B). A practical guide will be introduced later.

Identifying Events within Space S

Venn diagrams, named after J. Venn, have provided us with a visual picture of probability (Hays, 1994). You will need to refer to Figures 1-4 in order to understand what the following logical notation means. Space S, seen in Figures 2-4 is equal to 1. Space S cannot equal greater than 1, therefore, all that resides within Space S must be less than 1, unless, of course, all the sets within Space S are mutually exclusive and exhaustive, and there is no intersection of two or more sets or union of two or more sets (see Hays' probability rule 5 in Appendix B). If Space S can represent all that is knowable, then events A and B (A or B≠Ø), which are not mutually exclusive, must be part of what is knowable. In other words, events A and B do not exist unless they can be observed. Therefore, as depicted in Pew's example of SA (Figure 2), if one proposes, as Pew does, that ideal, achievable, or actual represents three observable events, then it must be possible to examine each event empirically. However, because Pew uses a construct (situation awareness) instead of an observable phenomenon, these events cannot be analyzed or measured in their present state, so any Venn reference to SA states cannot be made. Pew's illustration might make some sense if explained using another method, but as Dekker (2005) pointed out, any Venn illustration of SA that uses the term SA as the event name is really not a Venn diagram, because all the properties of set theory and probability are violated at the same time. If, however, one were to describe SA in terms that could be observed, say the 22 items provided by Dennehy and Deighton (1997), then it is very possible to create a constellation of events within Space S that either intersect, form a union, or are mutually exclusive.

To see how Dennehy and Deighton's 22 items (events) interact, at least one view of that interaction, observe how in Figure 5, the interacting items are superimposed around a central core or intersection in a Venn diagram. Following the protocols of set theory and probability, one can say that each event intersects with all the other events in the one Space S that all share. Thus Memory ∩ Perception,

Mental Picture ∩ Attention, Time perception ∩ Stress, Workload ∩ Spatial aware-
ness, and every other combination of events where one or more events intersect
(∩) with one or more other events shows just how complex SA can be and how
incomplete and wanting Pew's Venn depiction can be when trying to understand
relationships and interactions. If all 22 items were given a number (see Table 1)
to represent each, then one could say that 1∩2, 1∩2∩3, 1∩2∩3∩4, and
1∩2∩3∩5….∩22 are all possible combinations. If relationships were based on a
union between events, then 1∪2, 1∪2∪3, 1∪2∪3∪4, 1∪2∪3∪4∪5, 1∪2∪3∪4∪5….∪22
would express varied relationships of union.

$$\text{Space S (S = 1.00)}$$

Situation Awareness as Probability Statements. Referring now to Appendix B
and Table 1, it is possible to reinterpret Endsley's (2000) dynamic model as prob-
ability statements.

Endsley's dynamic decision making model prominently features situation aware-
ness as three levels, named earlier in this paper. After examining the model and
looking for similarities between the 22-item list of Dennehy and Deighton and the
Endsley model, the following items appear in both places.

Table 2
*Compatible Sub-elements from Endsley (2000) and Dennehy and Deighton
(1997)*

Stress	Memory	Perception
Workload	Training	Attention
Information processing	Knowledge of aircraft systems	Knowledge of goals
Mental picture	Knowledge of goals	Awareness of environment
Decision making ability	Knowledge of procedures	

Starting with Probability Rule 1 (Appendix B, Hays, 1994), one can claim that
*p(*Memory∪~Memory) [read: probability of Memory "or" "not" Memory] can be
rationally explained as *p*(~Memory) = 1—*p*(Memory). Within Space S there must
exist both Memory and "not" (~) Memory. Simply put, one can differentiate between
Memory and any other item on the shorter list. Thus, Memory, Stress, Training,
Perception, and the other 10 items exist by themselves and are not dependent on
other events for their existence. This gives scientists the opportunity to study only
one sub-element of situation awareness, without studying all the others. How-
ever, one must bear in mind that studies focusing on one sub-element of SA do
not also insinuate that this one sub-element operates in isolation to all the other
sub-elements of SA. This understanding is of great importance when pilots review
an NTSB accident report.

Dekker (2005) used a variation of Probability Rule 1 in his explanation for
how human error is perceived among stakeholders in flight operations. As Equa-

tion 1 shows, human error is thought to be what is left when mechanical error has been eliminated as a cause for an incident or accident. This formula is not fully supported, and in some corners of psychology it is utterly useless (Hollnagel, Kaarstad, & Lee, 1999).

Human Error or p(~mechanical error)$= f$(1 - mechanical error)(1)

Probability Rule 2 says that for any event within Space S, only positive numbers are within probability range (Hays, 1994, p. 23). This means that all sets within Space S are positive. So, for example, p(Attention) is within Space S because its measure is stated in terms of positive numbers. Any sub-element can be inserted in the place of attention. Therefore, to be recognizable and probable, all situation awareness sub-elements must exist within Space S. Whereas Probability Rule 1 suggested that each sub-element could be studied apart from the others; Probability Rule 2 establishes the limits of any measure.

Probability Rule 3 states that $p(\emptyset) = 0$. If $p(S) + p(\sim S) = 1$, and $\sim S = \emptyset$, then $p(\emptyset) = 0$. Regarding situation awareness, the probability that SA is explained as a null set is zero. This finding suggests that some sub-element of situation awareness is always in operation, either in the Person or the Environment (Dennehy & Deighton, 1997). Therefore, no sub-element of SA can be put in null hypothesis format and be found to be true. For example, consider the following null hypothesis.

[$H_{\emptyset}$: Working memory has no affect on overall situation awareness]

The sub-elements on the short list (Table 2) must occur or the notion of situation awareness is rendered incomplete, vacant in regard to logic. Any of the sub-elements can replace working memory to complete a new null hypothesis.

Now, while it is true that the null hypothesis will be found to be false when examining the presence of any sub-element of SA, the degree to which any sub-element influences SA can be measured and hypotheses made regarding these influences. Airplane pilots are most interested in the interaction between sub-elements and their influence on SA. They like the idea so much that they construct simulator scenarios to examine and reexamine these influences; and, after analysis of many simulator sessions, they try to mitigate any negative influence that each sub-element might have on aircrew SA (see discussion of Line Oriented Flight Training [LOFT] in Helmreich & Foushee, 1993).

Probability Rule 4, in conjunction with Rule 1, states that any sub-element of SA can be depicted as being in the same Space S as other sub-elements, but that these sub-elements can be observed in isolation, rather than intersection (Hays, 1994). Therefore, if P=Perception, and A=Attention, then:

$$p(P \cup A) = p(P) + p(A) - (PA) \qquad (2)$$

A researcher could set up a study of perception and attention, but by choice not make any connections between the two. Instead, after manipulating the experimental variables, he or she could record changes in either sub-element. Although

in the actual dynamic scenario sub-elements interact naturally, in the laboratory it is sometimes useful to measure each sub-element separately. Once the sub-elements have been teased apart, and after each element has been examined to better appreciate its contribution to SA, any laboratory notes on the elements can be used in the classroom.

Using field notes to supplement lesson plans on SA should come with a warning label, and classroom instructors should pay attention to it. There is a difference between *affect* and *effect* in the written findings of research, and understanding the difference between these explanations is absolutely crucial. If an experimenter wishes to pull apart the greater situation awareness model, each sub-element extracted from the dynamic process has an affect and effect tag associated with the element. For example, based on the SA models created or commented on by Endsley and Rodgers (1994), Dennehy and Deighton (1997), Uhlarik and Comerford (2002), and Pew (2000), all processes start with sensory or cognitive stimulation or input. The brain organizes this input based on pattern recognition and schema associations from long-term memory. The result is called perception; and because perception involves a person's long-term memory, every person will have a slightly different perception of the event, even though the sensory or cognitive stimulation is the same. Perception has a dramatic influence on attention, workload management, and stress. Therefore, one could say that sensory input has an affect on perception, which in turn affects all the other elements in the situation awareness chain.

Viewed from a different perspective, one could also say that perception is the result or effect of sensory stimulation. This would also follow when examining attention. Attention is affected by perception, but the effect of perception on attention is measured in how attentive the pilot is. Therefore, if the effect of any preceding element on any other element, but particularly the next in line in Neisser's perceptual cycle can be measured, then both the researcher in the laboratory and the instructor in the classroom can examine effects of each element.

In the classroom, the instructor could create objectives for each of the sub-elements, and thoroughly examine each of them. In the lab, the same instructor could advise the students to look only at one element at a time; and then following close examination of each element in SA, draw the students' attention to how each element is an effect of the preceding element's affect. Probability Rule 4 shows how individual examination of an element can be performed apart from the dynamic process. Thus individual examination of elements is represented by p(P) + p(A), subtracting how P and A intersect (P∩A) in the dynamic event.

Probability Rule 5 states that when events within Space S are mutually exclusive, if they occupy all of Space S all partitions equal 1, or the entirety of Space S (Hays, 1994). Here again, one sees sub-elements of SA existing independently from other sub-elements, and acting separately. Each sub-element is mutually exclusive of any other sub-element (see Figure 6). Although mutual exclusivity in situation awareness can occur in theory, none of the scientists mentioned in this

paper would agree that all the sub-elements of SA have no intersections. Probability Rule 5 is a display of unions between elements and can be displayed like this.

$$p(A \cup ..... \cup L) = p(A) + ...p(L) = 1.00 \qquad (3)$$

Set Theory, Venn Diagrams, Situation Awareness

Notwithstanding what has already been shown to be possible through probability statements, in regard to the sub-elements of the construct situation awareness, set theory offers more opportunities for intersection and overlapping of sub-elements within situation awareness than the simple five rules of probability.

If one ignores the rules of probability and set theory, then Pew's (2000) Venn illustration of his three views of situation awareness would elicit a response free of rules and boundaries. Visual depictions can be deceiving if they are drawn to depict a concrete relationship. If a student is held harmless from accepting faulty logic, then any illustration would be acceptable. However, when Pew created his illustration, he disregarded the obvious connections of his illustration with the boundaries of set theory and probability. Indiscriminate students and some faculty might support Pew's illustration and some faculty might explain situation awareness as it appears in the illustration. However, Pew, the indiscriminate student and faculty member would be wrong, and perhaps only Pew would know the difference. One of the purposes of this paper is to expose faulty logic and unhelpful illustrations.

Thus, Ideal SA is in some way the all-encompassing depiction of situation awareness, with Achievable and Actual SA being sub-sets of the ideal. In set theory notation, let AchievableSA=AchSA, and ActualSA=ActSA and IdealSA =IdSA, so that (AchSA⊂IdealSA) and (ActSA⊂AchSA∩IdSA) are theoretically feasible, except for the recurring problem that situation awareness is not a measurable entity itself, but a made up term to explain how a number of events that operate together change a person's perception enough for the individual to make decisions that will influence his or her future. This point has been made earlier in the paper, so instead of using Pew's three levels, I shall use set theory to properly associate sets and their subsets from illustrations used by Dennehy and Deighton (1997) and Endsley (2000).

In Table 1, you will find the 22-item listing of sub-elements of situation awareness, and in Table 2 the list of compatible sub-elements of SA. Dennehy and Deighton (1997) use their items to illustrate the subjective and objective nature of the Person in the loop, somehow fitted within the subjective and objective Environment. In set theory, the subjective Person (P_s) is included in the objective Person (P_o) and both exist within the subjective Environment (E_s) and the objective Environment (E_o). Therefore there are overlapping sets and subsets. Thus, if P_o is the set and P_s is included in P_o as a subset, then $P_s \subseteq P_o$. However, based on analysis of how SA sub-elements interact during LOFT scenarios, and based on Probability Rules 4 and 5, there is evidence that each of the 22 sub-elements can interact as $P_s \subseteq P_o$ and $P_s \subset P_o$ within both E_s and E_o. The subset included in the set and the proper subset of P_s within P_o both exist together, and provide researchers with alternative ways to explain the interaction of the subjective and the objective Person or Environment.

For example, in $P_s \subseteq P_o$ a trainer can compare the overall ability of a student pilot on the basis of all training accomplished thus far. This qualitative view can be used to measure any single training event. Suppose a student pilot routinely performs well. However, on one flight, she appears to be distracted by something other than the flight environment. The student's overall performance (P_o) is the set, and performance on a single flight is the subject part of the equation (P_s), or subset. Since performance on one flight is different from other flights, then elements of her overall performance record (P_o) are not in the one flight (P_s). So, $P_s \subset P_o$ explains how the objective Person can contain elements of performance that do not appear in the subjective Person.

Unions Prevent Inclusivity.
Look again at Figure 5. The figure provides one view of the intersections of 22 events (sub-elements of SA). Is it just as likely that the figure could be adjusted to reflect Probability Rule 4? Theoretically yes, but in actual fact, situation awareness is always an integration of sub-elements. The studies in this paper reveal SA as a series of "and" statements, such as A∩B. Situation awareness is also sometimes A∪B, even if other sub-elements are C∩D, C∩E, C∩G, and so on.

If Stress, one of the 22 items of SA, is not observed in actual flight or the simulator, there is a possibility that Stress will not influence SA in the same way as a different observable sub-element would. Thus Stress (abbreviated as St) forms a union with the other sub-elements (St∪W, St∪P, St∪A, St∪M), but is not shown to be integrated with the other elements, such as workload (W), perception (P), attention (A), and memory (M) might be shown to be (W∩P∩A∩M) (Colle & Reid, 1997; Endsley & Garland, 2000). When even one element is shown to form a union rather than an intersection, then SA is not fully complete—it is not fully developed—because at some time in the future, the missing sub-element might become active again, thus changing the complexity of SA for that individual: the subjective Person.

Free Association among Relationships. There is nothing in set theory to prevent unions from becoming intersections and intersections from becoming unions. As the previous section on inclusivity demonstrated, when one sub-element forms a union rather than an intersection with other sub-elements, SA is incomplete, or not fully developed. However, as Endsley (2000) has shown in her dynamic model of SA, relationships among sub-elements are never static nor are they intended to be so even with the standardization of advanced pilot training (Endsley & Garland, 2000; Endsley & Rodgers, 1994; Endsley & Rodgers, 1997)

Again, selecting Stress as an example of a union rather than an intersection, it is quite possible that during easier phases of flight, such as cruise on a cloudless day, that Stress is not observed and it is not having an affect on the subjective Person in the subjective Environment. Flight conditions can change, and normally do. As the subjective Person begins the descent and landing phase, the air traffic controller announces that the weather is deteriorating with a thunderstorm in the vicinity. It is likely that Stress would emerge again and influence SA. There-

fore, even though sub-elements are bound to intersect, they are also free to form unions for a period of time and then revert to an intersection relationship as flight conditions change.

NTSB Accident Reporting: Discovering the Elements of Situation Awareness

As promised, what follows is a practical guide for determining the effects of individual elements of situation awareness on crew actions. Rather than suggest that SA was the culprit and leaving it at that, this guide shows how various elements played a part in resulting crew actions. If one were to use the accident referred to herein, it is likely that a simulator session (LOFT-like) could be created to examine the relevance of each element and how each element could be mitigated to preempt "loss of SA." A preceding lesson on how to mitigate the effects of those situation awareness elements involved in an airplane accident will prepare the student for the simulator session.

The accident of concern here involves American International 808, a Douglas DC-8-61, which crashed ¼ mile short of the runway at Leeward Point Airfield, Guantanamo Bay, Cuba on August 18, 1993 at 1656 eastern daylight time (Dismukes et al., 2007, p. 51). The following excerpt from the NTSB report appeared on page 51 of the text by Dismukes et al.

> …the impaired judgment, decision-making, and flying abilities of the captain and flight crew due to the effects of fatigue; the captain's failure to properly assess the conditions for landing and maintaining vigilant situational awareness of the airplane while maneuvering onto final approach; his failure to prevent the loss of airspeed and avoid a stall while in the steep bank turn; and his failure to execute immediate action to recover from a stall.

Here is the list of deficiencies noted by the NTSB. When analyzing an accident, listing deficiencies is an important step, after reading the entire report from beginning to end.

1. Impaired **j**udgment [J]
2. Impaired **d**ecision-**m**aking [DM]
3. Impaired **f**lying **a**bilities [FA]
4. Improper **a**ssessment of **c**onditions [AC]
5. Ineffective vigilance in **SA** [SA]
6. Failure to prevent **l**oss of **a**irspeed resulting in a stalled condition (basic aircraft control) [LA]
7. Failure to take action to **r**ecover from **s**tall [RS]

Looking for Latent Failures. The insinuation by the Board is that there is an intersection among all seven events, such that J ∩ DM ∩ FA ∩ AC ∩ SA ∩ LA ∩ RS. All elements are included, but associations between elements are not specifically explained. For example, are J and DM combined in some way to insinuate that J ∩ DM = J·DM? If J and DM are combined, then if J·DM were not impaired, would the outcome have been different? Intuitively one might say "yes." From a James Reason point of view (1990), the sequence of events toward active failure could have been preempted by altering just one action. In NTSB reports, no one can accurately guess if judgment and decision-making were the only set of impairments that could have averted an accident.

Reckoning Situation Awareness. "Failure to…maintain vigilant situation awareness" was cited as a precondition leading to the failures to maintain basis aircraft control and recover from the stall (p. 51). Based on what is known about situation awareness, this is the time to refer to the 22-element list (Dennehy & Deighton, 1997). While also considering the impairments and failures, examine the NTSB report to find specific impairments relative to the list.

Put a check mark next to the SA element that most resembles what happened just preceding the active failure (crash). Dismukes et al. (2007) have provided some help in this way by examining what happened and naming what happened. Starting on page 58, the authors examined the factors affecting crew performance. This examination was based on the NTSB report, but was not restricted by it. The authors mentioned workload specifically, but they also mentioned fatigue. Dismukes et al. said that fatigue was the main culprit in this accident. Fatigue could be blamed for poor workload management, and a general sense of being lost en route to the runway.

Table 3
Situation Awareness Checklist

1. Memory		12. Overall Awareness	X
2. Decision making ability	X	13. Perception	X
3. Awareness of the environment	X	14. Safe flying	X
4. Present-state knowledge	X	15. Knowledge of procedures	X
5. Knowledge of goals		16. Air-performance awareness	X
6. Knowledge of aircraft systems		17. Attention	X
7. Spatial awareness	X	18. Mental picture	X
8. Workload	X	19. Course training	
9. Stress	X	20. Future-state knowledge	X
10. Time perception		21. Information processing	X
11. Cockpit resource management	X	22. Aircrew attitude	

Although the NTSB report only listed 7 impairments or failures, after using the 22-element checklist for situation awareness, not only are the 7 impairments or failures taken into account, but now there are additional elements. The 22-element listing names some elements also found in the NTSB report. Decision making ability is double listed. Safe flying is double listed (safe flying and flying abilities). Present-state knowledge (SA elements) can be paired with assessment of conditions and awareness of environment (NTSB). Safe flying can also be tied to loss of airspeed and failure to recover from stall, as well as air-performance awareness and attention.

Other elements not accounted for in the NTSB report are spatial awareness (a brain-based awareness), workload, stress, cockpit resource management (which by itself is a host of other elements), overall awareness, perception, knowledge of procedures, mental picture, future-state knowledge, and information processing. Had our appraisal of the NTSB report ended with their 7 impairments or failures, we would have neglected to look at the remaining 10 elements. If we had stopped with the problem of maintaining vigilant situation awareness, we would have stopped with the term *vigilance*. However, vigilance is only a minor player in this story.

It is important to decide how the elements interrelate. Ask yourself this question, "are any of the elements grouped together so that they are a set within another set?" You already know that SA has sub-elements, so all the marked elements in Table 3 can be considered to be a subset of SA. Thus, one can say that [each specific element] ⊆ [situation awareness]. In combination, all the marked elements = SA. This association (subset and set) is important when creating objectives for a lesson on SA, with a companion simulator session.

All 22 elements were not involved in the lack of vigilant situation awareness. For example, memory, although a part of any cognitive process, was not singled out in the report and not insinuated in the additional report by Dismukes et al. (2007). One should be careful about including all elements in the accident analysis. It is recommended here that instructors confine their list to activities they can teach in the classroom or the simulator, or even the airplane. For example, aircrew attitude shows up in the 22-element list, but unless it is a critical component to an accident, it is difficult to reproduce this event in the simulator in the same way it occurred on the day of the accident.

There is still a problem with cockpit resource management in this report. Like SA, CRM is a complex mixture of many elements: to include communication, stress management, workload management, team work, situation awareness, and risk management. Even these CRM elements can be further reduced. However, how this additional analysis should be conducted is material for future studies. For this study, it is enough to show that NTSB reports do not completely tell the whole story. NTSB reports need a second level of analysis.

To adequately perform an analysis of the 22 elements of SA, it helps to know something about cognitive psychology. Each SA element is in itself a constellation of other terms and descriptions.

Creating Instructional Elements. Based on the quote from the NTSB report in the preceding section, it was possible to make the following assertion: J ∩ DM ∩ FA ∩ AC ∩ SA ∩ LA ∩ RS. However, after applying the SA checklist, it is also possible to make the following assertions:

[Present-state knowledge ∩ AC ∩ Awareness of Environment];
[Safe flying ∩ LA ∩ RS ∩ Air-performance ∩ Attention].

This implies that all events that intersect (∩) should be taught together. The instructor would have objectives like:

Objective 1: Given a flight from point A to destination B, determine the pilot's present-state knowledge, assessment of conditions, and awareness of the environment. Or

Objective 2: Given a flight from point A to destination B, determine the pilot's ability to fly safely, to maintain proper airspeed, to recover from unusual attitudes or stalls, to self-evaluate air –performance, and to pay attention.

By developing course objectives that congregate those elements of situation awareness that are most likely associative, the instructor will be able to assess these abilities in a more productive way. It is recommended that each instructor use the checklist in Table 3 when assessing simulator sessions or when developing testing instruments in class.

If elements do not act together—they do not intersect—then it is important to keep elements apart to examine each element by itself. For example, neither the NTSB report nor the SA checklist listed memory as an impairment or failure. However, it is not possible to think without using memory. Memory is absolutely necessary for pilots when in the act of piloting an airplane. However, if it is not conspicuous, it is better to treat memory as a separate function, where [all affected elements] ∪ [unaffected elements]. When developing scenarios, it is wise to make two lists: one for all affected elements of SA and another for all unaffected elements. The instructor should show the two lists to the students and confirm why the elements are segregated as they are. When determining what a pilot should do, it is wise to determine what a pilot should not do. Testing strategies should incorporate questions of inclusion and exclusion.

How Set Theory Affects Situation Awareness Training. As an instructor builds lesson plans on the topic of situation awareness, it is wise to keep in mind some of the ideas presented in this paper. It is also important to use Venn diagrams and set theory as illustrations of the abstractions that make up the construct known as situation awareness. When constructing enabling objectives, there are a number of points you need to make. First, let the class know that it is possible to name the sub-elements of SA. Refer to the study by Dennehy and Deighton (1997) as proof of some uniformed approach to investigating the elements of SA. Second, illustrate that it is possible to examine each sub-element, as though it were entirely separated from all other sub-elements. Use Probability Rules 1, 4, and 5 and propositions from set theory (Appendix A) to illustrate this point. Third, use set theory to show that sub-elements of SA can be included in a set called Actual SA or as the interactionists suggest (Dennehy & Deighton, 1997) sub-elements (subjective Person) are included in the subjective Environment. Fourth, inform the class that it is possible for one or more sub-elements of SA to go missing, and that these elements resist being found even if an expert is observing a phase of flight. Stress can be part of SA at times and at other times it might not be an observable part of SA. This opens up a discussion on the nature of each sub-element: how each can be mitigated or outright controlled. Next, illustrate that unions between sub-elements can form freely, depending on the situation, and even transform into intersections as the flight scenario changes. Use Endsley's (2000) dynamic model of SA to illustrate how relationships between and among sub-elements can change when the subjective Environment changes.

Conclusion

Sometimes classroom instructors settle for illustrations of situation awareness that are handed down from one instructor to the next, or instructors use texts that present a narrow view of SA, which can be very misleading. If SA is not studied properly, then pilots will not be equipped with the tools to affect the nature of situation awareness in the flight environment (Hawkins, 1987)

Students and instructors might shy away from discussions where philosophy, psychology, and statistics are used to present abstract constructs, such as situation awareness. Studies in philosophy help students form better arguments; psychology helps students appreciate the processes ongoing in the brain; and statistics do more than annoy undergraduates: a study of Venn diagrams and its underlying set theory provide illustrations of relationships between and among elements that play a vital role in pilot training.

A good background in set theory will help students debunk illustrations that misinform, rather than inform. Even though Pew's (2000) Venn diagram of situation awareness was an innocent portrayal of levels of SA, his illustration have been misleading. Even though models of SA present various perspectives on how SA is gained or lost (Bell & Lyon, 2000; Uhlarik & Comerford, 2002), none of the models present the intricacies of the interplay or intersection of sub-elements of SA as well as set theory. An understanding of the rules of probability and the propositions of set theory prevent disinformation about how SA works. It follows then that instructors would be less inclined to make confusing statements about SA, if the insides of SA were produced and studied, in the light of set theory and the powerful Venn illustration.

The practical guide for determining sub-elements of SA, as shown in the NTSB report of American International 808, and the follow-on analysis of Dismukes et al. (2007), indicated that NTSB findings may fall short. When situation awareness was separated from the 7 impairments and failures, 10 additional elements (within SA) appeared. If a pilot would have stopped with the list of seven, he or she would have failed to understand how SA played in the accident. It is recommended here that professors and flight center staff read *Limits of Expertise* (Dismukes et al., 2007) and use it to enhance flight training. Dismukes et al. showed how much more could be captured from an NTSB report by applying a second level of analysis. The SA checklist (Table 3) is just such a tool. Additional checklists should be developed for the concept of complacency and crew resource management.

References

Bell, H. H., & Lyon, D. R. (2000). Using observer ratings to assess situation awareness. In M. R. Endsley & D. J. Garland (Eds.) *Situation awareness analysis and measurement* (pp. 129-146). Mahwah, NJ: Lawrence Erlbaum and Associates.

Colle, H. A., & Reid, G. B. (1997). Mental workload measurement structures. In D. Harris (Ed.) *Engineering psychology and cognitive ergonomics: Vol. 1 transportation systems,* (pp. 253-260). Aldershot, Hants, England: Ashgate Publishing Ltd.

Dekker, S. W. A. (2005). *Ten questions about human error*. Aldershot, England: Ashgate Publishing.

Dennehy, K., & Deighton, C. D. B. (1997). The development of an interactionist framework for operationalising situation awareness. In D. Harris (Ed.) *Engineering psychology and cognitive ergonomics: Vol. 1 transportation systems,* (pp. 283-290). Aldershot, Hants, England: Ashgate Publishing Ltd.

Dismukes, R. K., Berman, B. A., & Loukopoulos, L. D. (2007). *The limits of expertise: Rethinking pilot error and the causes of airline accidents*. Aldershot, England: Ashgate Publishing.

Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., Nikolic, D., & Manning, C. A. (1999). *Situation awareness as a predictor of performance in en route air traffic controllers* (DOT/FAA/AM-99/3). Washington, D.C.: Office of Aviation Medicine.

Endsley, M. R., & Garland, D. J. (Eds.) (2000). *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum and Associates.

Endsley, M. R., & Rodgers, M. D. (1994). *Situation awareness information requirements for en route air traffic control* (DOT/FAA/AM-94/27). Washington, D.C.: Office of Aviation Medicine.

Endsley, M. R., & Rodgers, M. D. (1997). *Distribution of attention, situation awareness, and workload in a passive air traffic control task: Implications for operational errors and automation* (DOT/FAA/AM-97/13). Washington, D.C.: Office of Aviation Medicine.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.) *Situation awareness analysis and measurement* (pp. 3-32). Mahwah, NJ: Lawrence Erlbaum and Associates.

Hall, J. (2005a). *Tools of thinking: Understanding the world through experience and reason* (Part 1). Chantilly, VA: The Teaching Company.

Hall, J. (2005b). *Tools of thinking: Understanding the world through experience and reason*(Part 2). Chantilly, VA: The Teaching Company.

Hawkins, F. H. (1987). *Human factors in flight* (2nd ed.). Aldershot, England: Ashgate Publishing Ltd.

Hays, W. L. (1994). *Statistics* (5th ed.). Fort Worth, TX: Harcourt College Publishers.

Helmreich, R. L., & Foushee, H. C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit resource management*, pp. 3-45. San Diego, CA: Academic Press.

Hollnagel, E., Kaarstad, M., & Lee, H-C. (1999). Error mode prediction. *Ergonomics, 42*(11), 1457-1471.

Kasser, J. L. (2006). *Philosophy of science* (Part 1). Chantilly, VA: The Teaching Company.

Kreeft, P. (Speaker) (2004). *What would Socrates do*? Recorded Books, LLC, Barnes & Noble Publishing.

Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge University Press.

Pew, R. W. (2000). The state of situation awareness measurement: Heading toward the next century. In M. R. Endsley & D. J. Garland (Eds.) *Situation awareness analysis and measurement* (pp. 33-47). Mahwah, NJ: Lawrence Erlbaum and Associates.

Uhlarik, J., & Comerford, D. A. (2002). A review of situation awareness literature relevant to pilot surveillance functions (DOT/FAA/AM-02/3). Washington, D.C.: Office of Aerospace Medicine.

Figures
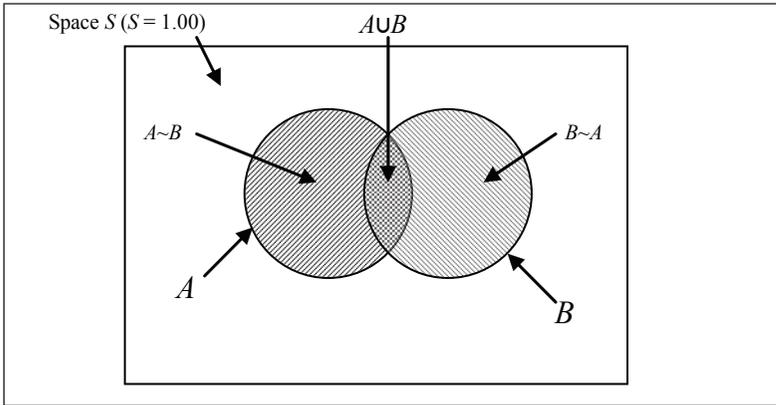
Figure 1. Relationships Between Events (Venn diagram)
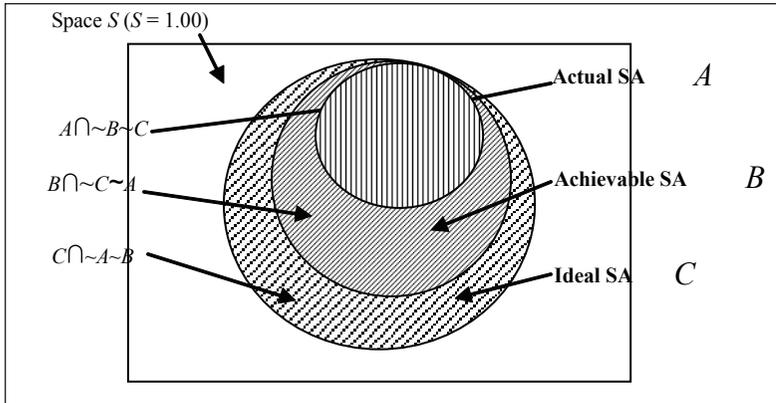
Figure 2. Interrelationships of SA without Intersections
(adapted from Pew, 2000, p. 36)
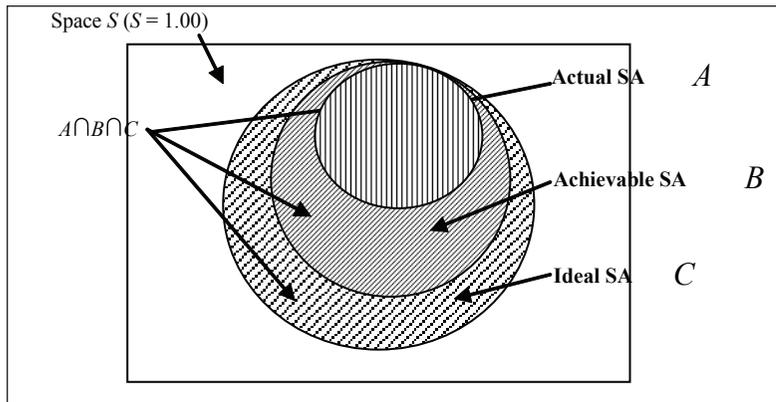
Figure 3. Interrelationships of SA with Intersections
(adapted from Pew, 2000, p. 36)
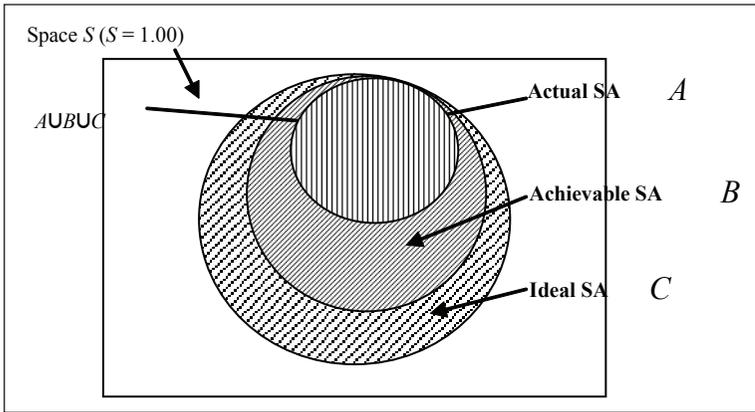
Figure 4. Interrelationships of SA Shown as Unions
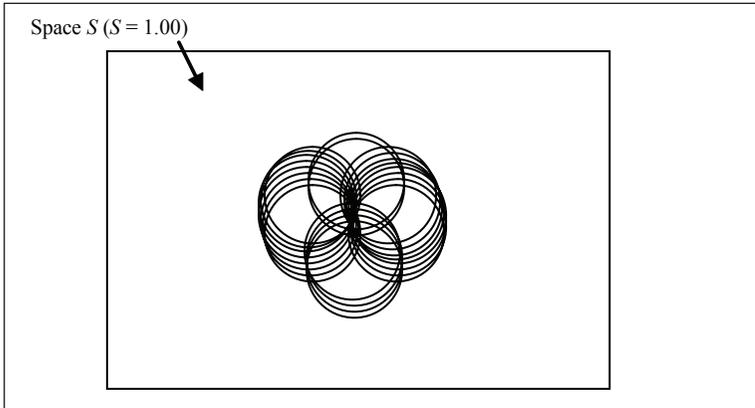
Figure 5. 22-item view of SA
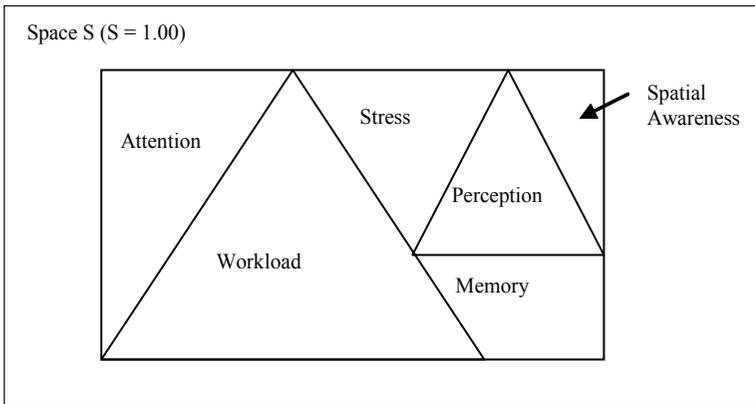
Figure 6. Mutually Exclusive Partitions of Space S

Appendix A: Mathematical Symbols*

Set Theory

A	a set, or event
A = {a,b,c,d}	set A includes the following members: a set specified by listing
B⊆A	the set B is included in the set A; B is a subset of A
B⊂A	set B is a proper subset of A, so that some elements of A are not in B
Ø	the empty set or "impossible" event; the null set
A∪B	union of sets or events A or B
A∩B	intersection of sets or events A and B
~A	complement of set A: the event "not A"
A − B	difference between sets A and B
A = B	sets A and B are equal or equivalent
(a,b)	ordered pair of elements, a from set A and b from set B

*Symbols and meanings were extracted from Hays, 1994, p. 1093

Probability Rule 1	Rationale
$p(\sim A) = 1 - p(A)$	$p(A \cup \sim A) = p(A) + p(\sim A)$ and $p(A \cup \sim A) = p(S)$ so that
Hays, 1994, pp. 22-23	$p(A) + p(\sim A) = 1.00$ and thus, $p(\sim A) = 1 - p(A)$

Probability Rule 2	Rationale
$0 \leq p(A) \leq 1.00$ for any event A	Only positive numbers that are ≥ 0 and ≤ 1 are within the probability range.
Hays, 1994, p. 23	

Probability Rule 3 $p(\emptyset) = 0$, for any S	Rationale
	Also known as the rule of the impossible event, we know that:
Hays, 1994, p. 23	If $p(S) + p(\sim S) = 1.00$ and since $\sim S = \emptyset$ then $p(\emptyset) = 0$

Probability Rule 4	Rationale
For any two events A and B in S, $p(A \cup B) = p(A) + p(B) - p(A \cap B)$	When A and B are mutually exclusive, $p(A \cap B) = 0$ So that
Hays, 1994, p. 24	$p(A \cup B) = p(A) + p(B)$

Probability Rule 5	Rationale
If the set of events A, ...,L constitutes a partition of S, then $p(A \cup ....\cup L) = p(A) + ...p(L) = 1.00$	When elements are mutually exclusive and exhaustive, the sum of their probabilities must be equal to 1.00.
Hays, 1994, p. 24	

** All probability rules and their rationale were taken from (Hays, 1994)

Screening Air Traffic Control Specialist (ATCS) Applicants for Psychopathology

Raymond E. King, David J. Schroeder, Carol A. Manning,
Paul D. Retzlaff, and Clara Williams

Civil Aerospace Medical Institute
FAA/CAMI
AAM-520
P.O. Box 25082
Oklahoma City, OK 73125
(405) 954-1163
raymond.king@faa.gov

Abstract

Applicants for air traffic control specialist (ATCS) positions have been assessed with the 16 Personality Factor (16 PF) test since 1965, with progressively decreasing effectiveness. The purpose of the present study was to assess the viability of using of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) in lieu of the 16 PF. A sample of 1,014 ATCSs in training voluntarily completed the MMPI-2 as part of a research program, after being cleared with the 16 PF. Those data are used to estimate the number of future candidates that would be referred for follow-up psychological evaluations, given varying MMPI-2 scale cut scores (T scores of 65, 70, 75, and 80). A final algorithm of 70T or above on scales 1, 2, 3, 4, 6, 7, and 8 as well as 75T on scale 9 is initially recommended. Initial cut scores are recommended that are mindful of the reality of a non-referred population and the needs of the operational air traffic control community.

Screening Air Traffic Control Specialist Applicants

Selecting applicants for safety-sensitive jobs such as air traffic control specialist (ATCS) from pools of promising individuals can be a difficult task because applicants and organizations usually have conflicting goals. Applicants are attempting to look their best to increase their chances of being hired and are likely to minimize any mental health issues. At the same time, organizations attempt to hire the most qualified and fit individuals, those who are most likely to successfully complete their training and become effective employees. Organizations face

two challenges when hiring. The first task is to "select in" those applicants who have the required positive attributes. Select-in methods determine who is best suited for completing the complex tasks associated with safe and efficient performance. The second task is to "select out" those applicants with negative qualities that would pose a safety risk or otherwise make success difficult due to their limited adaptability. Neither approach alone is sufficient. Select-out criteria often eliminate applicants with a psychiatric diagnosis (as currently defined in the *Diagnostic and Statistical Manual of Mental Disorder*, DSM-IV TR, APA, 2000) suggesting a lack of fitness. The result is the identification of a small, probably impaired subset of the candidate pool who it is not wise to hire, at least not without additional assessment. Aviation occupations require the highest standards of psychological suitability and fitness to ensure the public safety.

Due to the high stakes involved, Butcher (2002) issued a call for research into the mental health of commercial pilots and advocated the use of modern psychometric instruments. He made no mention, however, of other vital members of the aviation team such as air traffic control specialists (ATCSs). To ensure aviation safety, personnel in both occupations need to be alert, attentive, and ready to respond immediately to critical events. Although the Federal Aviation Administration (FAA) certifies the medical fitness of pilots and ATCSs, commercial aviators typically work for airlines or other private employers that bear the responsibility for selecting qualified applicants. The vast majority of ATCSs (commonly termed "air traffic controllers") are employees of the FAA. Thus, the FAA is responsible for developing the appropriate selection tools and conducting the selection screening.

In an effort to ensure the emotional health (fitness) of the ATCS workforce, the FAA has used the 16 Personality Factor (16 PF) test since 1965 during the medical assessment (select-out) process. This procedure is part of the Controller Health Program, which was initiated by FAA Order 9430.2 and currently outlined in FAA Order 3930.3A (see Appendix A). Candidates took both Form A (187 items) and Form B (187 items) of the 16 PF and thus responded to 374 items on a three-point scale. A "case identifier" scale, composed mostly of anxiety items from the 1967-1968 edition, has evolved over the years (Convey, 1984). The 38-item scale displays acceptable reliability. The 18 items from Form A had a Cronbach alpha of .71, and the 38-item full scale had a Cronbach alpha of .85 (King, Retzlaff, Detwiler, Schroeder, & Broach, 2003). Unfortunately, its clinical utility has been limited. Historically, only a very small percentage of job candidates have been identified by this scale (in 2007, 1 out of 2,010 applicants or .05%; in 2006, 3 out of 1,200 or .25%, and in 2005, 5 out of 843 or .59%).

Dollar, Broach, and Schroeder (2003) indicated that the 16 PF is somewhat effective in predicting who will go on to retire on disability but posited that other factors must be at play in determining disability retirements. It is possible that some individuals who would have gone on to retire on disability were screened out with the 16 PF, leading to a potential restriction in range.

There was interest in comparing the 16 PF with other psychological tests as early as 1971, when Smith gauged the item ambiguity (the degree to which an item elicits multiple interpretation) of the 16 PF to the Minnesota Multiphasic Personality

Inventory (MMPI). Smith found the items of the MMPI to be less ambiguous than those of the 16 PF. While both tests have been substantially revised since 1971, the FAA continues to use the 1967/1968 edition of the 16 PF. In 1996, Schwarzkopf, Buckley, and Pace urged replacement of the FAA's 16 PF procedure due to declining scientific interest, its sole focus on anxiety symptoms, and its "fake-ability." In their paper, written under contract to the FAA, they urged consideration of the MMPI-2. Indeed, the point of the current paper is to explore the feasibility of using the MMPI-2 for this *initial* screening of ATCS candidates to better identify those harboring symptoms suggestive of emotional instability and requiring additional assessment.

Graham (1990) noted that the MMPI-2 has been used in two ways in selection. It can be used to screen for psychopathology, and it can be used to predict the quality of an applicant's job performance by matching personal characteristics to job requirements. Graham wrote that police and nuclear power plant operators are best selected by eliminating persons with very elevated scores on one or more of the clinical scales, obviously following the former approach. Graham asserted that screening applicants for psychopathology is most justified when considering individuals for occupations involving susceptibility to stress, personal risk, and personal responsibility. He specifically delineated "air traffic controller" (p. 197) among such sensitive occupations.

Lowman (1989) provides useful information regarding some of the difficulties and concerns when conducting pre-employment screening for psychopathology, focusing mainly on the occupation of nuclear power plant workers. Westefeld and Maples (1998) reported on matching applicants, mostly in police departments, to successful occupational incumbents based on the MMPI-2. Such an approach is more in line with a select-in approach rather than a select-out approach as it suggests an optimal profile, rather than a need to meet minimal criteria (absence of psychopathology).

The major concerns about the future psychological functioning of ATCSs, who are relatively young at the time of entry into the workforce (by law, they must be less than 31 years of age), include mood disorders and other DSM-IV-TR (2000) Axis I disorders (termed "neurosis," and "psychosis" in FAA Order 3930.3A; see Appendix A), in addition to personality disorders. Hammen (2001) summarized the epidemiological research on depression and cited varying lifetime prevalence and concluded that prevalence estimates are influenced by demographic factors (with women and those with lower income and education levels having higher rates), as well as the method of assessment. McNally, Malcarne, and Hansdottir (2001) called for increased longitudinal research on the spectrum of anxiety disorders to better understand these conditions and their development during the course of a lifetime. Harvey (2001) explored the relatively rare prevalence (approximately 1% of the population) of schizophrenia and noted that it is particularly deleterious to occupational success when it develops in late adolescence/early adulthood. Finally, Geiger and Crick (2000) considered personality disorders, by definition an *enduring* pattern of maladaptation, and found that the diag-

nosis is typically not reliable; hence, prevalence estimates are problematic. Overall, the prevalence of psychopathology in the pool of candidates for ATCS positions is likely to be relatively low, particularly at the time the candidates are assessed, but the potential consequences cannot be discounted. There is therefore a growing interest in enhancing the FAA's ability to identify those ATCS candidates who possess psychological symptoms that could compromise aviation safety.

This study only considers the select-out portion of selection. Readers who are interested in the process the FAA uses to select-in applicants (the process to assess aptitude) are referred to King, Manning, and Drechsler (2006). The current study was designed to explore the feasibility of utilizing the MMPI-2 to replace the 16 PF as the initial (select-out) screen. Readers are first introduced to the MMPI-2 and the concept of T scores and then the research method and results are described. Findings are then translated in terms of the operational needs of the FAA.

The MMPI-2

Scales
There are numerous scales on the MMPI-2. The first scales to consider are "validity" scales and used to determine the test-taking "style" of the client. Ten main "clinical" scales focus on various psychopathologies. The scales have names, but these names are typically not used in the profession because they are quite archaic. Instead, clinical psychologists refer to the scales by their letters or number. Hence, a patient scoring high on the fourth scale would not be said to have scored high on "Psychopathic Deviate" but to have scored high on "*PD*" or "scale *4*."

Validity Scales
L ("*Lie*")
Elevation reflects a deliberate attempt of the individual to present him/herself in a positive light. There may be denial of minor flaws/weaknesses that most individuals would admit. Excessive elevations on this scale (relatively rare) render the profile invalid and hence uninterpretable.
F ("*Fake Bad*")
Used to detect atypical ways of responding. Test items are agreed to by few normal adults. High scores may reflect an invalid test profile due to malingering. *Not expected in a job applicant*; more likely in a person seeking the benefits of a patient (compensation, avoidance of jail).
K ("*Fake Good*")
A more subtle index of individuals who are trying to present themselves in a positive light. Tends to be elevated as education and socioeconomic status increases (and when used in an employment selection setting); denies psychopathology. Used to correct several of the clinical scales (Scales *1*, *4*, *7*, *8*, & *9*) by adding various amounts of *K* to get a more accurate reading of their functioning.

Clinical Scales
Scale *1* – *Hypochondriasis*
High scores reflect individuals who have an excessive number of vague nonspecific complaints and body concerns (gastrointestinal distress, fatigue, pain, and general weakness).

Scale *2 – Depression*

> Individuals with high scores often reflect depressive symptoms (depressed, blue, unhappy) and are generally lacking in self-confidence.

Scale *3 – Hysteria*

> Symptoms involve two dimensions. One reflects a general denial of physical health and includes rather specific somatic complaints. The other group involves a general denial of psychological or emotional problems.

Scale *4 – Psychopathic Deviate*

> High scores reflect difficulty in incorporating the values and standards of society; may involve asocial or antisocial behaviors, impulsiveness, and need for immediate gratification. May be a bit elevated in younger test takers as a normal function of late adolescence (Graham, 2006).

Scale *5 – Masculinity-Femininity*

> A legacy bi-directional scale (in other words, both low and high scores have meaning) that is not considered a clinical scale. Reflects interest and not sexual orientation. As it does not indicate psychopathology, it is not suggested for use in the medical screening of ATCSs.

Scale *6 – Paranoia*

> High scores reflect individuals with disturbed thinking, ideas of reference, suspiciousness, hostility, and paranoia.

Scale *7 – Psychasthenia*

> High scores reflect individuals experiencing a great deal of psychological turmoil and discomfort. They tend to be anxious, tense, and agitated. They are worrisome individuals that have difficulty concentrating.

Scale *8 – Schizophrenia*

> High scores are reflective of bizarre mentation, delusions, and possible hallucinations. Confused thinking, poor judgment, and alienation are common.
>
> Scale *9 – Hypomania*
>
> High scores are suggestive of overactivity, poor impulse control, irritability, and possible aggressive outbursts

Scale *0 – Social Introversion*

> Not a clinical scale. Individuals with high scores tend to be introverted, while low scorers are extroverted. As it does not indicate psychopathology, it is not suggested for use in the medical screening of ATCSs.

T score conversion

The number of items endorsed in the keyed direction on each scale (the "raw score") are converted to T scores by using the published norms. These are standard scores with a mean of 50 and a standard deviation of 10. Thus, a T score of 70 means that the individual scored two standard deviations above the mean of the population on which the test was normed for the scale in question. An elevation of two standard deviations may be better understood as the 95th percentile, meaning higher than 95% of the population. So, the higher the T score values, the more items the person taking the test endorsed, suggesting a relatively greater presence of traits consistent with a psychiatric disorder.
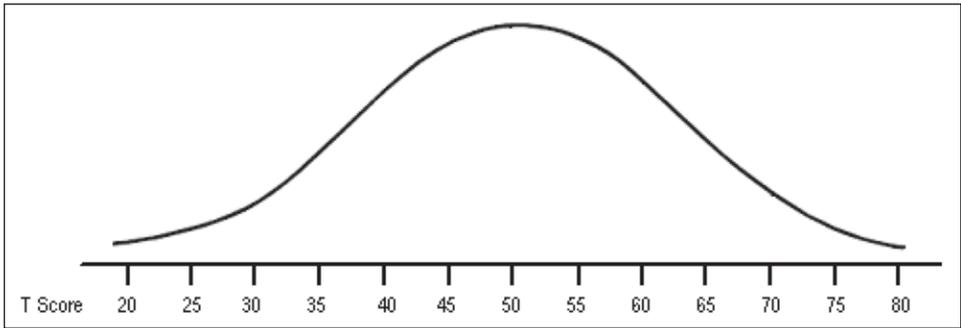
Figure 1. T-scores and the normal distribution

Method

The MMPI-2 was administered to a cohort of recently hired ATCSs during their first days of training at the FAA Academy in Oklahoma City, OK. These 794 male and 220 female fledgling air traffic controllers had all been hired as ATCSs and thus had been interviewed at an air traffic control facility and had passed a medical examination, which included passing the existing 16 PF case-identifier-procedure hurdle. All voluntarily agreed to participate in the present study. All participants had at least 12 years of education. Many participants had several years of college, as most were hired under the College Training Initiative program, which is hosted by 13 four-year and community colleges throughout the United States. Indeed, 544 of the 1,014 participants indicated that they had more than 12 years of education.[1]

Non-gender norms were used, as this research is intended to support personnel selection and gender-specific norms are prohibited by the Civil Rights Act of 1991. A consideration of gender was also unnecessary due to the lack of inclusion of scale 5 (Masculinity-Femininity), which specifically requires attention to the test taker's gender. Non-K corrected Clinical Scales are also reported here, as *K* corrections tend to be elevated in applicant populations (due to the tendency for positive impression management – also known as "faking good").

Results

Table 1 (which presents *K*-corrected clinical scales to allow comparison to the data published by Butcher, 1994) provides the means and standard deviations for the K-corrected MMPI-2 scores for the 1,014 ATC applicants; Figure 2 presents this information graphically for ease of comparison. Again, T scores represent the norms, with a mean of 50 and a standard deviation of 10. As this sample's data diverge from these, differences become apparent. In terms of validity scales, both *L* and *K* are about 7 points higher. This elevation is almost one standard deviation and approaches a significant elevation. It is apparent that both groups, ATCSs and pilot applicants, have placed themselves in a positive light and deny pathology.

---

1. Due to a misunderstanding of the demographics portion of the data collection, participants did not consistently report their total years of education.
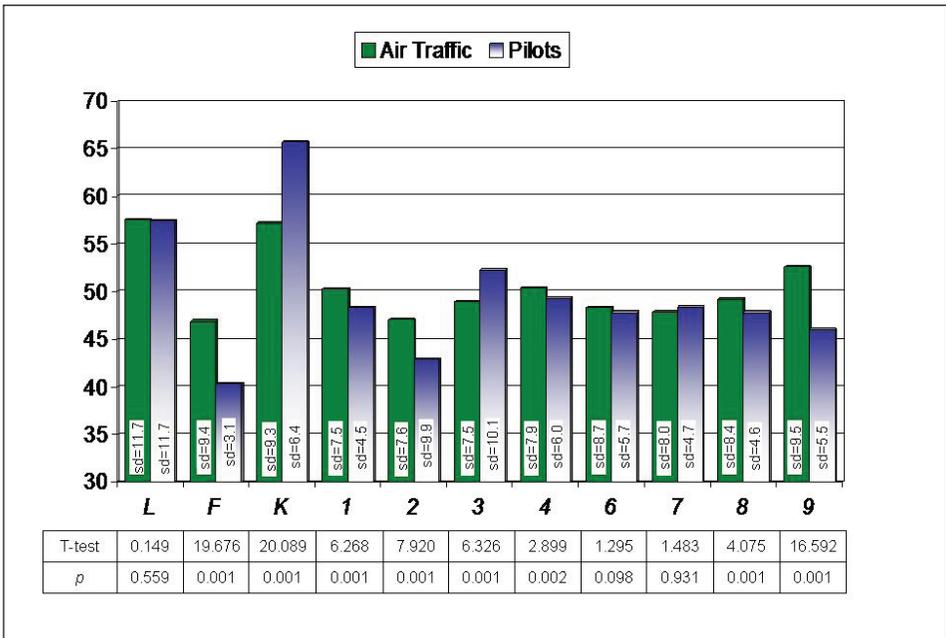
Figure 2. K-corrected means and standard deviations for air traffic control specialist (research) and pilot (applicant) samples.

Nevertheless, the clinical scales are remarkably similar to the general population normative group published in the MMPI-2 manual (Butcher, Graham, Ben-Porath, Tellegen, Dahlstrom, & Kaemmer, 2001). Here, clinical scales do not vary as much as 3 points from the norm. It would be expected that the sample would minimize pathology in the clinical scales, but they do not. Minimization of these scales would entail means down around 40, not at the mean of 50. The additive K corrections, however (.5 for scale *1*, .4 for scale *4*, 1 for scale *7*, 1 for scale *8* and .2 for scale *9*), may inflate some of the clinical scales to an extent.

Table 1
K-corrected means and standard deviations for air traffic and pilot samples.

	Air Traffic Mean (Standard Deviation)	Pilots Mean (Standard Deviation)	t	p
Validity Scales				
L	57.5 (11.7)	57.4 (11.7)	0.149	0.559
F	46.9 (9.4)	40.4 (3.1)	19.676	0.001
K	57.2 (9.3)	65.7 (6.4)	20.089	0.001
Clinical Scales				
1	50.3 (7.5)	48.3 (4.5)	6.268	0.001
2	47.1 (7.6)	42.9 (9.9)	7.920	0.001
3	48.9 (7.5)	52.3 (10.1)	6.326	0.001
4	50.4 (7.9)	49.3 (6.0)	2.899	0.002
6	48.3 (8.7)	47.8 (5.7)	1.295	0.098
7	47.9 (8.0)	48.4 (4.7)	1.483	0.931
8	49.2 (8.4)	47.8 (4.6)	4.075	0.001
9	52.6 (9.5)	46.0 (5.5)	16.592	0.001

Table 1 also provides the means and standard deviations for the pilots from Butcher's (1994) work. While the normative sample used to establish the T scores can serve as a reference group, Butcher's pilots can be considered an aerospace-specific control group of individuals of approximately the same age, although half a generation apart. This pilot sample data differs from that of ATCSs in other ways as well. While Butcher tested only men, the current sample is about 20% female. In addition, Butcher's data were collected before applicants were offered employment. As such, the current comparison is less than ideal. That being noted, Butcher's pilots scored at about the same level on the L validity scale as our sample. His pilots also scored about 16 points above the mean on the K scale (p<0.001). His sample appears to be more defensive than the ATCS sample. It is possible that the validity scales would be more similar if the groups were responding in more analogous settings, rather than research participants (ATCSs) and job applicants (pilots).

Looking at the clinical scales for ATCSs and pilots, there are more similarities than differences. Most scales are within a point or two across the samples. The only apparent differences are on 2 and 9, with both scales being higher in the ATCS sample (p<0.001).

Table 2 provides the means and standard deviations using non-K-corrected norms. Here it is noted that scales 7 and 8 are particularly lower than with K-correction, compared to the norms. This sample endorsed fewer items that reflect anxiety or disordered thinking than the general population normative group.

Table 2
Non-K-corrected means and standard deviations for ATCS sample.

Scale	Mean T (Standard Deviation)
1	45.7 (8.3)
2	47.1 (7.6)
3	48.9 (7.5)
4	47.5 (7.6)
6	48.3 (8.7)
7	43.0 (8.4)
8	44.1 (9.2)
9	50.8 (9.2)

In using the MMPI-2 as a screen for ATCS applicants, a practitioner's interest would not be to compare the applicant directly to a group average. Rather, it would be to compare an applicant to established cut scores on the various clinical scales. Table 3, therefore, provides the percentage of participants scoring at or above potential cut scores of 65T, 70T, 75T, and 80T. At the 65T cut-score level, only about 2-3% of participants are elevated on a given scale. The exception was on scale *9*, "*Hypomania*," where about 9% fell at or above this cut score. These elevations are more likely indicative of this young sample's generally high energy level than reflective of a high number of individuals with possible mood disorders.

Table 3

Non- K- corrected and non-gender norms: percentages above cut scores for ATCS sample.

Scale	≥ 65T (%)	≥ 70T (%)	≥ 75T (%)	≥ 80T (%)
1	1.0	0.4	0.2	0.1
2	1.7	0.4	0.2	0.1
3	2.0	0.8	0.3	0.1
4	1.7	0.7	0.3	0.1
6	3.4	2.0	1.0	0.3
7	1.8	0.8	0.3	0.1
8	2.7	2.0	0.8	0.6
9	9.3	4.2	1.6	1.1
1 or more	14.6	7.4	3.2	2.0
2 or more	3.7	2.0	1.0	0.4

Discussion

As a number of participants scored high on more than one scale, the percentage of participants identified by the test in general is not a simple total of the percentages for each scale. Summing across subjects with one or more scales at or above 65T, we identified about 15% of the participants. Participants with two or more elevated scales represented only about 4% of the sample.

At a less stringent 70T cut-score level, only about half as many participants were identified. Indeed, less than 1% were at or above that level for most scales, with the particular exception of *9*, which was at about 4%. About 7% had one or more scales elevated, and only 2% had two or more scales elevated.

The percentage above 75T continued to drop. Very few were identified by individual scales alone. About 3% had one or more high scales, and only 1% had two or more high scales.

Using a cut score of 80T resulted in very few applicants being identified. Here only 2% had one or more scales highly elevated, and only 0.4% had two or more. Most individual scales were identifying only 0.1%. Differences across the various cut scores point to one of the concerns associated with use of a screening tool. With lower criteria, a clinician is likely to identify a higher percentage of individuals who, in fact, would not represent a risk. In turn, at higher cut scores, a clinician is going to miss a number of individuals who are likely to prove to be at increased risk.

The most relevant statistic here is the percentage of participants with one or more scales at or above the cut scores. If the 65T cut score is used, then the assumption is that some 15% of the ATCS applicants have significant psychopa-

thology. This seems high for a relatively high functioning group. After all, this group has demonstrated a high degree of functioning by being hired, either by demonstrated ability or by passing a rigorous examination. If the 70T cut-score is used, then a psychopathology prevalence rate of about 7% is suggested. This number seems far more consistent with the probable psychopathology rate of this group of people. A cut score of 75T results in about 3% being identified. Finally, a cut score of 80T appears to be so high that only 2% are identified.

Many factors contribute to the establishment of cut scores. First, the best way to make decisions of this type is to collect research data before hand and compare them with actual multi-year outcomes. Unfortunately, this project has moved forward more quickly than that. As such, a rigorous research program should be delineated very quickly to allow for cut score changes in the future that are based on data.

In the meantime, there are probably two over-arching factors. The first is a practical issue and that is the number of applicants referred for follow-up, in-person evaluations. The second is the a priori probability of the presence of psychopathology within this population.

If applicants are to be referred for follow-up evaluation, the number of referrals must be sufficient to warrant a program, yet low enough to avoid huge costs on the high end. It is probably not worth having a program if fewer than 30 applicants are referred each year. Conversely, at about $1000 per evaluation, things get expensive and difficult to manage if there are more than 100 or so per year. If there are 1500 applicants per year, this logic suggests that cut scores resulting in between 2% and 7% would be a good target from a programmatic perspective.

The second approach is the epidemiological approach. Here an a priori estimate of psychopathology in this group is approximated, and the cut scores are set to that level. Some groups have more pathology and some less. In an inpatient psychiatric facility, nearly 100% have some sort of significant pathology. In an outpatient setting, many clients are seeking assistance for such things as marital problems and child problems, so the prevalence of significant psychopathology is probably quite low. In any case, individuals with severe psychopathology are not likely to pursue a career as an ATCS.

The *lifetime* prevalence of mental disorders in adults aged 18 to 54 was 29.4% between 1990 and 1992 and 30.5% between 2001 and 2003, (Kessler et al., 2005). The question that remains unanswered is the likelihood of an employee *developing* mental illness over the course of adulthood; a question that can only be answered by longitudinal research (Reifler, 2006). Use of the MMPI-2, or any other selection instrument, cannot completely eliminate the risk of an employee developing a mental illness over the course of a career, but it can help identify who is *currently* suffering from psychiatric symptoms and is a big step to the longitudinal research envisioned by the Controller Workforce Plan (FAA, 1965).

In samples such as ATCS candidates, where there are several prior screening hurdles and a certain self-selection process, it is doubtful that the level of true psy-

chopathology is much over 5%. At least 1% probably experience some degree of psychological discomfort, but it is hard to believe that as many as one in 10 would meet the criteria for a "mental illness" diagnosis.

Both of these approaches converge on a cut score solution that is very similar. A program should identify and refer between 2% and 7% of applicants. This cut point is programmatically and clinically logical and is likely to strike a balance of false positives and false negatives.

In deciding upon cut scores, there are two possible approaches. The first is to identify different cut scores for different scales. The second is to select a single cut score for all scales. The allure of choosing differing cut scores is that some of the anomalies in the table can be "smoothed out." For example, Scale *9* seems to pick up more participants than would be clinically suggested. Alternatively, a single cut score for all scales would be the most parsimonious solution but the least sensitive.

Then the question becomes, "What cut scores should be used?" Here the percentages in the "1 or more" row are the relevant data. If a cut score of 65 is used, 14.6% would be referred. This rate is probably too high. If a cut score of 80 is used, only 2.0% would be referred. This rate seems too low.

So the discussion should center on the use of either 70 or 75 (or a combination). With 1500 applicants, a cut score of 70 would result in about 100 (7.4%) being sent for follow-up evaluations. This rate is reasonable. It also has the benefit of resulting in enough follow-up evaluations for scale-specific outcome purposes.

A cut score of 75 would result in 3.2%, or about 40 of 1500 applicants. This rate would be less costly and easier to manage initially.

It should be remembered that as the program matures and data are collected, the data might support future adjustments in the cut scores. As such, cut scores should not be "fixed in stone."

Cut Score Decision
MMPI-2 cut scores of 70 appear to meet the needs of the agency best, with the exception of scale *9* where a 75 is more reasonable. Hence, the decision model will use a cut score of 70 and above for scales *1, 2, 3, 4, 6, 7,* and *8*. Since scale *9* appears to "over classify" the controllers, that scale will use a cut score of 75 and above. With these cut scores, any applicant with one or more scales above the cut score will be referred for second-tier psychological assessment.

This algorithm will identify 4.9% of applicants. This outcome is well within the parameters suggested. This number is low enough that second-tier costs will not become too great, the management of candidates will be efficient, and there will

be appropriate periods of time during which the medical status of candidates is undetermined.

Applicant Level Reliability

The purpose of this paper was to model the behavior of the MMPI-2 in ATCS applicants for initial psychological test screening. The data suggest that the ATCS participants in this study were remarkably close to the published norms as well as similar to the pilots reported by Butcher (1994). Overall, the vast majority of the participants in this study rendered profiles solidly within normal limits, when compared with the normative sample. It should also be noted that elevated MMPI-2 scales do not necessarily indicate the presence of a disqualifying medical condition, as situational circumstances can result in the temporary elevation of clinical scales without a concomitant presence of a psychiatric condition. Sorting these matters out is the function of the second-tier assessment.

Recommendations

The MMPI-2 should be given serious consideration to replace the 16 PF in the psychological screening of ATCSs. Cut scores of 65, 70, 75, and 80 were applied to the dataset, and the resulting numbers of identified participants were examined for clinical consistency and programmatic need. Therefore, we recommend using initial cut scores of 70 and above for scales *1*, *2*, *3*, *4*, *6*, *7*, and *8*, as well as a cut score of 75 and above for scale *9*. These cut scores resulted in no more than 4.9% of the sample being identified as requiring further psychological evaluation. While an argument could be made to set the cut scores at 65 as that represents the traditional clinical cut point of the MMPI-2, readers must bear in mind that the research participants and the ATCS job candidates represent non-referred populations. The previous method was referring to considerably less than one percent. Any screening program must guard against having too many false positives and cut scores can be adjusted as the data accumulates.

This approximately 5% appears to be relatively well represented across the eight MMPI-2 scales that are being used. Individual subject analysis suggests that some proportion of participants took the task less than seriously and responded randomly (King, Schroeder, Manning, Retzlaff, Williams, 2008). Actual ATCS applicants, however, will likely be more responsive to the testing situation; thus, we expect an identification rate of about 4% to 4.5%. This rate seems acceptable from both a clinical and an administrative perspective.

With between 1200 and 2000 applicants being evaluated per year over the next 10 years, this rate should result in between 48 and 90 follow-up, second-tier evaluations per year. This flow should be sufficient to justify the program but not so great as to require large budgetary requirements and an unmanageable flow of candidates awaiting medical clearance.

A series of further research studies are needed to refine the cut score algorithm as applicants go through the hiring and training process. Thus, we offer the following additional recommendations:

1) The applicant data, when the MMPI-2 is used with candidates with a tentative offer of employment, should be immediately analyzed to ensure the percentages above the cut scores are not vastly different from the research sample. If the "yields" begin to differ, then the cut scores should be reconsidered.
2) The MMPI-2 scales should be compared to the data from second-tier assessments and clinician recommendations as they become available. Differential cut scores could be used, for example, if it is shown that some scales tend to predict negative clinician recommendations better than others do.
3) The scores should also be compared with ATCS training outcomes. While there are many reasons for poor training outcomes, some amount of the outcome is certainly related to psychological functioning.
4) The scores should be compared with dismissals and psychological problems that are reported to medical authorities, as well as other outcome measures. While it will take several years to accumulate this type of data on individuals who have taken the MMPI-2, it will be instructive due to the very considerable costs that may be avoided in the future.
5) Future research should consider inclusion of other scales, to include content scales, particularly those dealing with alcohol and substance abuse. In the meantime, these conditions will be assessed during the candidates' medical examinations. Results of these independent assessments will prove useful for future research.
6) Consideration should be given to conducting research on using the Restructured Clinical scales developed by Tellegen, Ben-Porath, McNulty, Arbisi, Graham, & Kaemmer (2003) and the 338-item Reformatted Form, which is about to be released .

Conclusion

The assessment of psychological symptoms is best conceptualized as a two-tier process. The first step would be to test all candidates that have been tentatively offered employment in compliance with the Americans with Disabilities Act (1991). The first tier could use the MMPI-2 in lieu of the 16 PF. Those candidates that score above established cut scores would then be evaluated more thoroughly by a licensed psychologist. This approach would be consistent with FAA Order 3930.3A, which specifically prohibits medical disqualification based on a single psychological test (FAA, 1980).

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*, 4th Edition Text Revision. Washington, DC: APA.

Americans with Disabilities Act of 1990 (1991). Pub. L. No. 101-336, §2, 104 Stat. 328.Retrieved January 1, 2008, from http://www.ada.gov/pubs/ada.htm.

Butcher, J.N. (1994). Psychological assessment of airline pilot applicants with the MMPI-2. *Journal of Personality Assessment, 62*(1)*,* 31-44.

Butcher, J.N. (2002). Assessing pilots with the "wrong stuff": A call for research on emotional health factors in commercial aviation. *International Journal of Selection and Assessment*, *10*(1&2), 168-184.

Butcher, J.N., Graham, J.R., Ben-Porath, Y.S., Tellegen, A., Dahlstrom, W.G., & Kaemmer, B. (2001). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administering, scoring, and interpreting*. Revised Edition, Minneapolis, MN: NCS Pearson.

Civil Rights Act of 1991 (1991). Public Law No 102-166. Retrieved January 1, 2008, from http://www.eeoc.gov/policy/cra91.html

Convey, J.J. (1984). Personality assessment of ATC applicants. In SB Sells (Ed.), *Selection of air traffic controllers* (pp. 323–51). (FAA/AM-84/2). Washington, DC: FAA Office of Aviation Medicine.

Dollar, C., Broach, D., & Schroeder, D. (2003). *Personality characteristics of air traffic control specialists as predictors of disability retirement*. (DOT/FAA/AM-03/14). Washington, DC: FAA Office of Aerospace Medicine.

Federal Aviation Administration (1965, October 15). *FAA Order 9430.2 - Establishment of health program for agency air traffic control specialist employees*. Washington, DC: U.S. Government Printing Office.

Federal Aviation Administration (1980, May 9). *FAA Order 3930.3A - Air traffic control specialist health program.* Washington, DC: U.S. Government Printing Office.

Geiger, T.C., & Crick, N.R. (2000). A developmental psychopathology perspective on vulnerability to personality disorder. In RE Ingram & JM Price (Eds.), *Vulnerability to Psychopathology* (pp. 57 – 102). New York: The Guilford Press.

Graham, J.R. (1990). *MMPI-2: Assessing personality and psychopathology,* 2nd Edition. New York: Oxford University Press.

Graham, J.R. (2006). *MMPI-2: Assessing personality and psychopathology*, 4th Edition. New York: Oxford University Press.

Hammen, C. (2001). Vulnerability to depression in adulthood. In RE Ingram & JM Price (Eds.), *Vulnerability to Psychopathology* (pp. 226–57). New York: The Guilford Press.

Harvey, P.D. (2001). Vulnerability to schizophrenia in adulthood. In RE Ingram & JM Price (Eds.), *Vulnerability to psychopathology* (pp. 355–81). New York: The Guilford Press.

Kessler, R.C., Demler, O., Frank, R.G., Olfson, M., Pincus, H.A., Walters, E.E., et al. (2005). Prevalence and treatment of mental disorders, 1990 to 2003. *The New England Journal of Medicine*, *352*, 2515-23.

King, R.E., Manning, C.A., & Drechsler, G.K. (2006). Operational Use of the Air Traffic Selection and Training Battery. *International Journal of Applied Aviation Studies*, *6*(2), 207-217.

King, R.E., Retzlaff, P.D., Detwiler, C.A., Schroeder, D.J., & Broach, D. (2003). *Use of personality assessment measures in the selection of air traffic control specialists*. (DOT/FAA/AM-03/20). Washington, DC: FAA Office of Aerospace Medicine.

King, R.E., Schroeder, D.J., Manning, C.A., Retzlaff, P.D., & Williams, C.A. (2008). *Screening air traffic control specialists for psychopathology using the Minnesota Multiphasic Personality Inventory-2.* (DOT/FAA/AM-08/13). Washington, DC: FAA Office of Aerospace Medicine.

Lowman, R.L. (1989). *Pre-employment screening for psychopathology: A guide to professional practice*. Sarasota, FL: Professional Resource Exchange, Inc.

McNally, R.J., Malcarne, V.L., & Hansdottir, I. (2001). Vulnerability to anxiety disorder across the lifespan. In RE Ingram & JM Price (Eds.), *Vulnerability to Psychopathology* (pp. 322– 25). New York: The Guilford Press.

Reifler, C.B. (2006). Epidemiological aspects of college mental health. *Journal of American College Health Association, 54*(6)*,* 372-6.

Schwarzkopf, A., Buckley, M.R., & Pace, T.M. (1996). *16 PF screen evaluation* (DTFA-02-93-D-93088). Unpublished manuscript, University of Oklahoma.

Smith, R.C. (1971). *Personality assessment in aviation: An analysis of the item ambiguity characteristics of the 16 PF and MMPI* (FAA/AM-71/35). Washington, DC: FAA Office of Aviation Medicine.

Tellegen, A., Ben-Porath, Y.S., McNulty, J.L., Arbisi, P.A., Graham, J.R., & Kaemmer, B. (2003). *The MMPI-2 restructured clinical (RC) scales: Development, validation, and interpretation*. Minneapolis, MN: University of Minnesota Press.

Westefeld, J.S., & Maples, M. (1998). The MMPI-2 and vocational assessment: A brief report. *Journal of Career Assessment, 6*, 10-1

Appendix A

Excerpt from FAA Order 3930.3A

CHAPTER 4. INITIAL HIRE

40. GENERAL. The medical standards for initial employment prescribed under the Physical Requirements section of the Office of Personnel Management (OPM) Qualification Standards for the Air Traffic Control Series 2152 shall be applied to all applicants for initial ATCS employment.

41. PSYCHOLOGICAL TESTING. A comprehensive psychological test battery shall be administered to all ATCS applicants at the time of their pre-employment interview or during the interval between the interview and the medical examination. The administration of this test battery shall be the responsibility of local facility chiefs; test materials shall be provided by the Flight Surgeons. Completed answer sheets shall be sealed by the applicant, and transmitted by the facility to the Federal Air Surgeon: Attention AAM500. TEST RESULTS SHALL NOT BE THE SOLE BASIS FOR REJECTION OF AN ATCS APPLICANT. Personnel who conduct this testing shall instruct applicants to refrain from discussing the content of tests with other applicants.

42. SUPPLEMENTAL MEDICAL INFORMATION. The Flight Surgeon shall obtain and evaluate applicable military and Veterans Administration medical records through established regional procedures. If military medical records are not received within 120 days following request, the Flight Surgeon may grant conditional medical clearance pending receipt and review of the additional medical

g. Psychiatric.
The applicant must have no established medical history or clinical diagnosis of any of the following:
    (1) A psychosis;
    (2) A neurosis;
    (3) Any personality disorder or mental disorder that the Federal Air Surgeon determines clearly indicates a potential hazard to safety in the Air Traffic Control System. The determinations will be based on the medical case history (including past social, and occupational adjustment) supported by clinical psychologists and board certified psychiatrists, including such psychological tests as may be required as a part of medical evaluation as the Federal Air Surgeon may prescribe.

h. Substance Dependency.
A history, review of all available records, clinical and laboratory examination will be utilized to determine the presence or absence or substance dependency, including alcohol, narcotic, and non-narcotic drugs. Wherever clinically indicated, the applicant must demonstrate an absence of these on thorough psychiatric evaluation, including any clinical or psychological tests required as part of the medical evaluation.

Drift Cues from a Tactile Belt to Augment Standard Helicopter Instruments

Ian P. Curry, Arthur Estrada,
Catherine M. Grandizio, and Bradley S. Erickson

U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362-0577

Abstract

Studies have shown that spatial disorientation (SD) plays a significant role in both the number and outcome of rotary wing accidents. Recent work has confirmed this and highlighted the contribution of brownout conditions and aircrew fatigue in accident causation. Current standard Army aircraft cockpit displays do not provide drift information leaving the pilot guessing the direction and magnitude of the aircraft's drift vector when close to the ground. This information is critical to the safe landing of helicopters in brownout conditions. The few helicopters with instrumentation that provide drift information do so via visual displays which require the focus of an already visually-saturated pilot. This effort tested a system that provides drift information through the tactile sense via a belt around the waist in a fatigued aviator model. Analysis of the study data showed that the tactile belt significantly improved drift control during takeoff and reduced drift error during hover. In fatigued pilots (awake for 31 continuous hours), all measures of drift were better with the belt versus without the belt. Fatigued pilots reported a significant reduction in visual and physical workload with the belt. Results indicate that the belt significantly improved pilot perception of drift and situation awareness and reduced mental stress.

Previous studies have shown that spatial disorientation (SD) plays a significant role in both the number and outcome of rotary wing class A-C accidents (Braithwaite, Groh, and Alvarez, 1997; Durnford, et al., 1995). More recent work confirms this finding and also highlights the contribution of brownout conditions and aircrew fatigue in accident causation (Curry and McGhee, 2007). In the years 2003-2005 the U.S. military lost in excess of $500M and 90 lives due to SD mishaps.

Early work on tactile displays in the sixties by Bliss et al. (1970) concentrated on replacing the orientation information from lost vision in the blind with that provided by touch. This work showed that the tactile sense could provide at least as much orientation information as sight although the reaction to that information was slightly slower, being in the order of 200 ms as opposed to 75 ms (Van Erp and Van Den Dobbelsteen, 1998). In a previous belt-area tactile display used in a tracking task (Schmid and Bekey, 1978), this translated to a response time of 0.25 seconds (s) versus 0.10 s for a visual response. The authors suggest this slight delay is due to the conduction velocity of the nerves concerned and is of no practical significance in the application proposed in this study.

Normal balance and orientation on the ground are provided by correct visual, inner ear, and skin/muscle/joint sensations. However, in aviation, the inner ear and skin/muscle/joint senses often provide false orientation cues. Visual information, the primary source of flight information, is usually reliable. However, using vision for orientation is intermittent, since vision must also be used for mission related information inside the cockpit. Understandably, the typical spatial disorientation accident occurs when the visual system is temporarily distracted or in reduced visibility.

Current standard Army aircraft cockpit displays do not provide drift information leaving the pilot guessing the direction and magnitude of the aircraft's drift vector when close to the ground. This information is critical to the safe landing of helicopters in brownout or whiteout conditions. The few helicopters with instrumentation that provide drift information do so via visual displays requiring the focus of an already visually-saturated pilot. To reduce the pilot's reliance on visual information during complex flight operations, the tactile situation awareness system (TSAS) was developed to provide information via the under-utilized sense of touch (McGrath et al., 1998; 2004). Providing tactile information allows the pilot to maintain orientation while looking away from the aircraft instrument panel. The full TSAS array consists of a custom fit, upper-body covering suit, shoulder straps, and a seat. All three components contain lines of tactors which respond to hardware and software in the aircraft and provide information on the aircraft's drift direction and magnitude. Unfortunately the system is bulky, hot, expensive, and difficult to maintain and, therefore, not a realistic option in the harsh field environments in which Army Aviation operates. Proof of concept flights for the TSAS were conducted in a UH-60 helicopter and resulted in improved aircraft control, increased pilot situational awareness, and a reduction in pilot workload (Raj et al., 1998; McGrath et al., 2004). Although successful, the expense of fitting each pilot with a custom TSAS vest remains prohibitive. The purpose of the current study was to

test a smaller, omnifit tactile system (TSAS-Lite) that provides drift information via a belt (8 tactors placed every 45°) around the pilot's waist.

The current study is different than previous studies in that it examined if tactile inputs limited to the waist proved as effective in providing helicopter drift information as the larger, more expensive TSAS. We hypothesize that the combination of cockpit instrument visual information and tactile drift information will provide the participants with a more complete "situational picture" when hovering, taking off, or landing in areas of limited visibility. This more complete "situational picture" will, in turn, reduce or eliminate inadvertent drift and the accidents that ensue. A successful demonstration could have significant implications, namely, the drift information provided by the TSAS-Lite belt could be integrated into pilot take-off and landing training and procedures.

Study Objective

The objective of this study was to assess whether tactile feedback delivered via a system worn around the waist could provide sufficient orientation information for fatigued pilots to perform helicopter flight maneuvers near the ground in a degraded visual environment.

Methods

The study was a within-subjects, repeated measures design and was conducted by the U.S. Army Aeromedical Research Laboratory (USAARL) personnel using the laboratory's UH-60 helicopter. Flight performance and subjective feedback provided TSAS performance and pilot perception metrics, while recurring cognitive tests established and confirmed levels of fatigue. Eligible participants included both men and women (military and civilian) between the ages of 19-55 years. The upper age limit of participants was restricted to 55 years based on research showing that total sleep time and other sleep parameters change dramatically in middle-aged individuals. Only healthy active duty, Reserve, National Guard, and Department of Defense civilian rotary wing pilots rated and current in the UH-60 helicopter were used in this study. Any pilots with prior experience with TSAS were excluded from the study.

The study consisted of four, two-day experimental periods (Table 1) during which two volunteers were randomly assigned per period. Hence, eight participants meeting the inclusion criteria participated. In order to assess the TSAS-Lite system in fatigued pilots, no sleep was permitted during the two-day experimental period resulting in 31 hours of continuous wakefulness. All volunteers were medically screened prior to taking part in the study and were taken through a comprehensive informed consent process before any experimental procedures commenced. Table 2 illustrates the diversity of the study's sample population.

Table 1
Testing Schedule (Two-day Experimental Period)

	Day 1	Day 2
00:00		
01:00		
02:00		
03:00		Cognitive testing
04:00		Shower
05:00		Breakfast
06:00	In-Process	
07:00	Cognitive testing	Cognitive testing
08:00	TSAS-Lite Simulator Training	
09:00		
10:00		Lunch
11:00	Lunch	Test Flight
	Test flight	
12:00	Cognitive testing	Cognitive testing
13:00		Sleep
14:00		
15:00		
16:00		
17:00	Cognitive testing	
	Dinner	
18:00		
19:00		
20:00		
21:00		
22:00		
23:00	Cognitive testing	

Table 2
Participant Demographics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Age	8	28	26	54	38	9.27362
UH60 Flight Hours	8	3026	54	3080	1219.875	1260.52035
Rotary wing Flight Hours	8	5855	145	6000	1984.875	2038.88313

Prior to data collection, each volunteer was trained in the use and interpretation of cues from the TSAS-Lite belt (Figure 1). The training consisted of a one-hour session in a UH-60 simulator on the first morning of the study. The experimental procedure consisted of two test flights with a safety pilot, the first on the afternoon of Day 1 and the second 24 hours later on the afternoon of Day 2 (Table 1). Each test flight consisted of one hover, two takeoff and climb-outs, and two approach and landings (Table 3) in the JUH-60 Research Black Hawk helicopter.

Figure 1. TSAS–Lite belt.

Table 3
Flight Maneuvers

	Maneuver	Maneuver Standards	Perform
1	Stationary Hover	Maintain Heading, Altitude (10 feet above ground level), & Position	Once during flight
2	Takeoff & Climb-out	Maintain ground track, continuous acceleration to 80 knots indicated airspeed, climb to 200 feet above ground level for traffic pattern.	Twice during flight
3	Approach & Landing	Maintain ground track and continuous deceleration to terminate to the ground the designated landing point.	Twice during flight

Both subjects/pilots were onboard the aircraft during each test flight. However, while one participant was flying, the other was seated in the passenger area unable to hear any comments regarding the flight and/or data collection. Upon completion of the first pilot's set of flight maneuvers, the pilots exchanged seats and the second pilot performed the flight maneuvers. In order to simulate a

degraded visual environment, the helicopter's chin bubble was occluded with bubble wrap and the flying subject/pilot wore a set of frosted goggles (Figure 2) which limited clear vision to the cockpit instruments. There were four test conditions: rested with a fully-functioning belt; fatigued with a fully-functioning belt; rested with a nonfunctioning belt; and fatigued with a nonfunctioning belt. All participants experienced all of the conditions with the condition order balanced across the sample to avoid any order effects. Psychometric testing was undertaken throughout the experimental period as per the schedule in Table 1.

Figure 2. Frosted goggles.

Description of the TSAS-Lite System

The prototype TSAS-Lite belt system uses the sense of touch to provide drift information to aircraft operators. The TSAS-Lite system accepts data from the aircraft via the ASN-128D Doppler Global Navigation System to obtain the aircraft position, velocity, and vector. Drift information is then displayed via the electromagnetic tactors located on the belt. During the flight maneuvers, the location of the signaling tactor was used to indicate direction of helicopter horizontal drift motion, and the tactor pulse pattern (frequency) was used to indicate the velocity of the helicopter drift.

The system consists of a commercial, off-the-shelf (COTS) PC-104 central processing unit (CPU) (Real Time Devices CMC6686GX233HR-128), a custom 8-channel tactor driver board and eight electromechanical tactors (Engineering Acoustics, Inc.). The tactors provide a vibrating stimulus at 90Hz +/- 20% with three rates of firing depending on pre-set ground speeds (0-15 knots [kts]: 200 milliseconds [ms], 15-30 kts: 600 ms, 30-45kts:1000ms). The sensation is similar in intensity to a standard pager vibration. The prototype belt is a flexible omnifit neoprene with Velcro fastenings and is worn sufficiently tight around the belt area to provide tactor contact while maintaining comfort. The CPU and tactor drive electronics are housed in a water resistant sealed housing, with data, tactor and operator switch interfaces. For operational use, the system could interface to existing military GPS units or COTS sensors. The system requires only digital data from position or direction sensors.

The testing schedule is shown in Table 1. The cognitive testing refers to a battery of computer-based tests that have established normative data for comparison with our subjects and was useful in establishing fatigued states. The tests consisted of the following:

Evaluation of Risks Questionnaire (EVAR)

Impairments in judgment are often apparent in situations where an individual engages in behavior where the risks far outweigh the probable advantages. The propensity to engage in or avoid risky behavior and situations was assessed by a brief 24-item paper and pencil questionnaire that has been used effectively to measure individual variability in risk assessment in previous research with Special Operations Forces (Sicard et al., 2001). Individuals mark a point along a 100 millimeter (mm) bipolar visual analogue scale to indicate their preference for various types of risky activities. Administration time was approximately 5 minutes.

Visual Analogue Scale (VAS)

The VAS consists of eight 100-mm lines centered over the adjectives 'alert/able to concentrate', 'anxious', 'energetic', 'feel confident', 'irritable', 'jittery/nervous', 'sleepy', and 'talkative' (Penetar et al., 1993). The extremes of each line correspond to ratings of 'not at all' and 'extremely.' Scores consist of the distance of the participant's mark from the left end of the line (in mm). The task was presented via computer.

Profile of Mood States (POMS)

The POMS (McNair, Lorr, & Droppleman, 1992) is a 65-item adjective checklist that measures current mood states along six subscales: tension-anxiety, anger-hostility, depression-dejection, vigor-activity, fatigue-inertia, and confusion-bewilderment. Volunteers rated themselves from 1 (not at all) to 5 (extremely) for each mood-related adjective.

Psychomotor Vigilance Task (PVT)

Participants completed a 10-minute PVT. A pushbutton response to the visual stimulus (presented with an inter-stimulus duration of 1-10 s) was required. The PVT is a standard laboratory tool for the assessment of sustained performance and is methodologically reliable and relatively versatile (Loh et al., 2004). The PVT data consists of reaction times from stimulus onset to response and includes the number of responses in excess of 500 ms.

Flight Data

The flight data were divided into three phases; hover, take-off, and approach to landing. All data were gathered through the Aeromedical Instrumentation System (AIS). This system is unique to the USAARL UH-60 and gathers data in six degrees of freedom allowing full analysis of the flight performance. The take-off and approach data of particular interest was in drift (unwanted lateral movement from a horizontal azimuth). The approach and take-off portions of the flights

were measured on lateral deviation (drift) from a direct flight-path to or from a designated point and both were expressed as integrations of the acceleration in meters per second (m/s). In the landing phase, drift produces many dynamic rollover type accidents and is the parameter not represented in the information provided by the flight instruments of the majority of helicopters. The hover data in lateral drift, heading and altitude represents an error from a set datum, and the final output is a root mean square error derived from a score produced automatically by the AIS. The hover portion was simply measured in meters from the datum. In addition, altitude data (ft) was collected in all conditions.

Post-flight Questionnaire

After each flight the subjects completed a questionnaire asking them to detail their perception of drift, mental stress, cognitive demand, situation awareness, visual workload and physical workload. The responses were made on a 100 mm Likert scale.

Results

All statistical analyses were conducted using SPSS® 12.0 with significance set at an alpha level of .05 for all statistical tests. There were two main areas of analysis: a) cognitive testing to establish a fatigue effect (using one-way repeated measures ANOVA) and b) flight data to assess performance effects (using two-way repeated measures ANOVA).

As expected, the cognitive assessments indicated a fatigued condition on Day 2. Across the sessions, the VAS sleepy and energetic scales revealed a significant fatigue effect ($p < 0.01$ in both cases) over time (Figure 3).
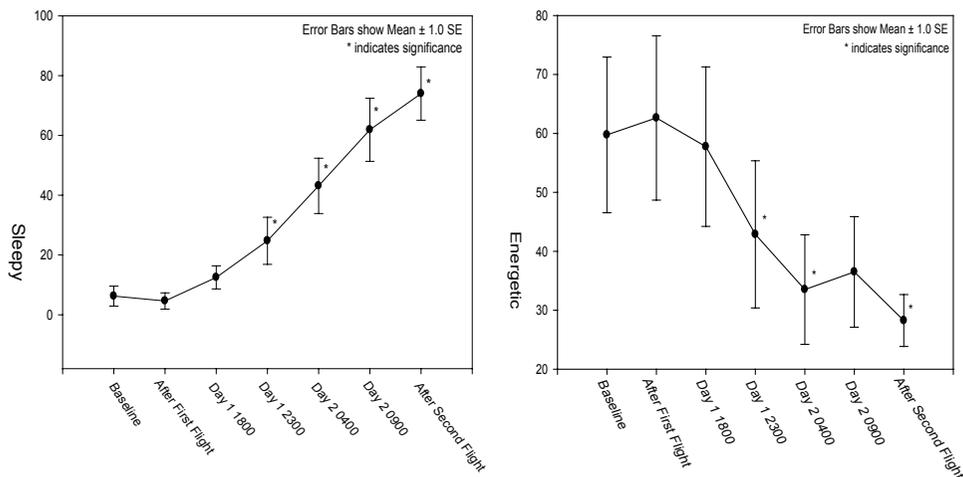
Figure 3. VAS sleepy and energetic measures.

The self confidence measure of the EVAR declined significantly ($p = 0.004$) across the sessions (Figure 4), while the POMS achieved significance ($p < 0.05$ in all cases) across the sessions indicating a decline in mood states (Figure 5).
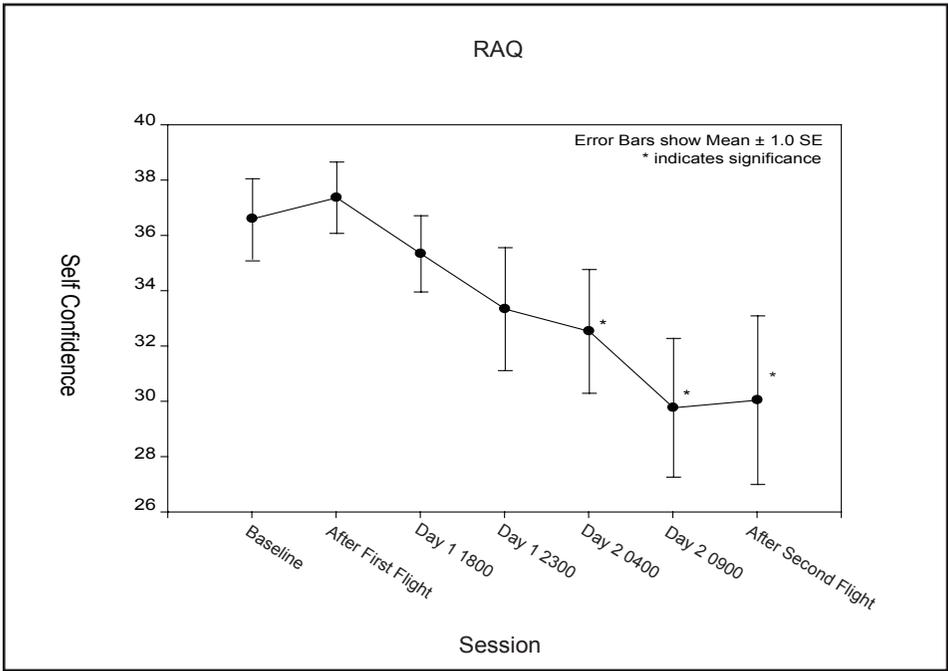
Figure 4. EVAR self confidence measure.
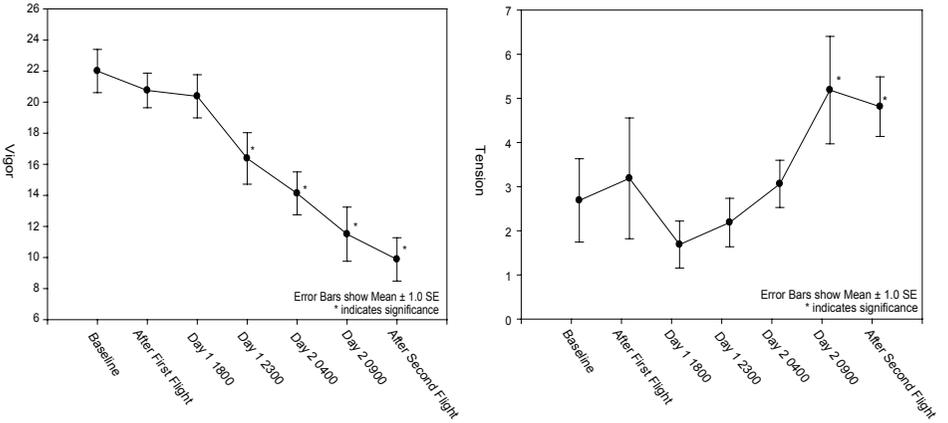
Figure 5 continued on next page
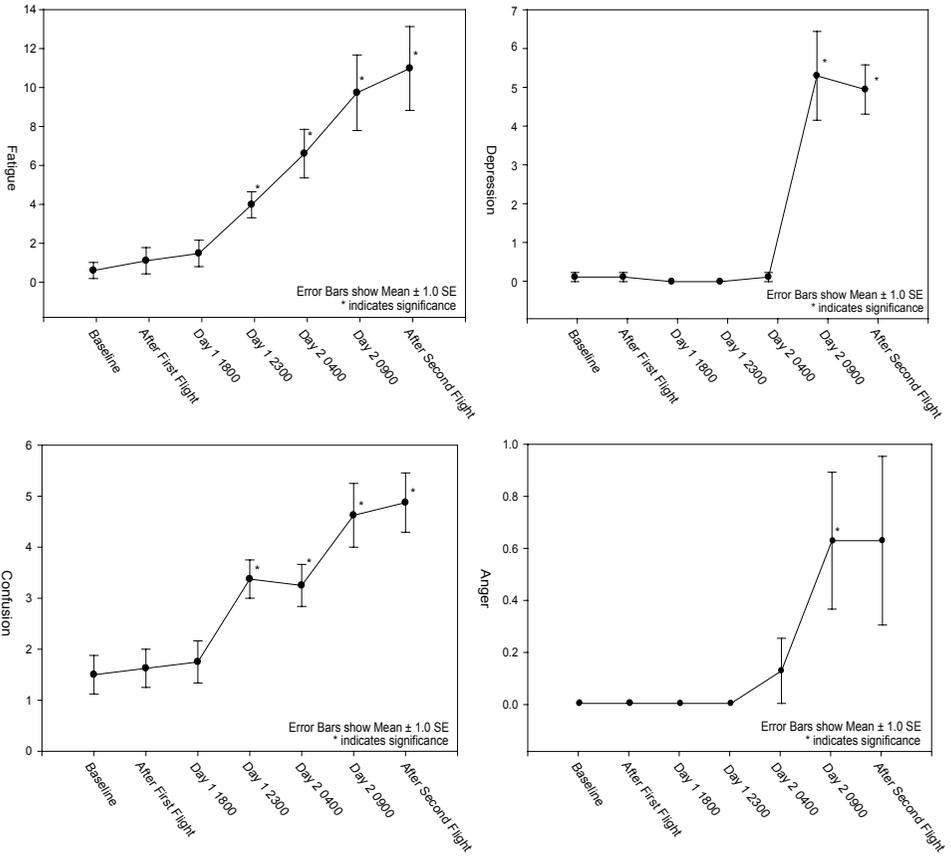
Figure 5. POMS results.

The PVT results presented in Figure 6 imply a tendency toward slower reaction times and a greater number of lapses greater than 500 ms over time. The degree of performance decline was not significant, however, due to the great variance on Day 2.
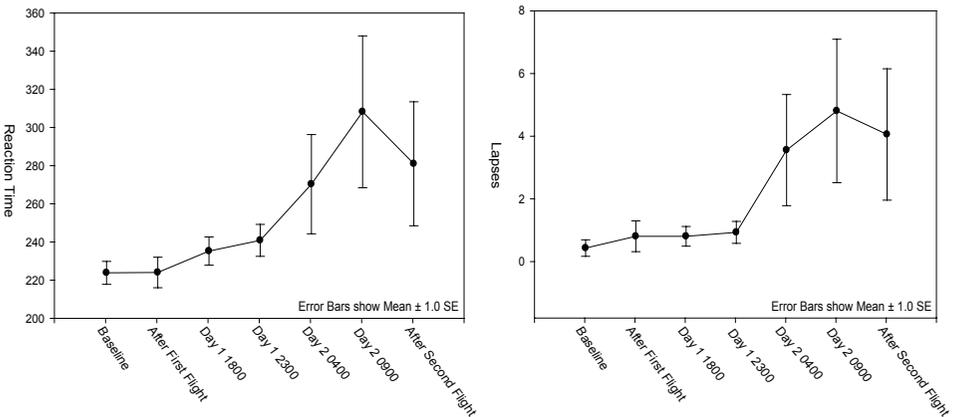
Figure 6. PVT data.

As for the flight performance data, analysis showed that the tactile belt significantly improved drift control during takeoff (*p*=.046) by rested aviators, yet demonstrated no significant differences in drift rate during approach between any condition of belt activity or fatigue (Figure 7). Hover performance did show a significant improvement in drift control with the belt active as opposed to inactive (*p*=0.027)(Figure 8).
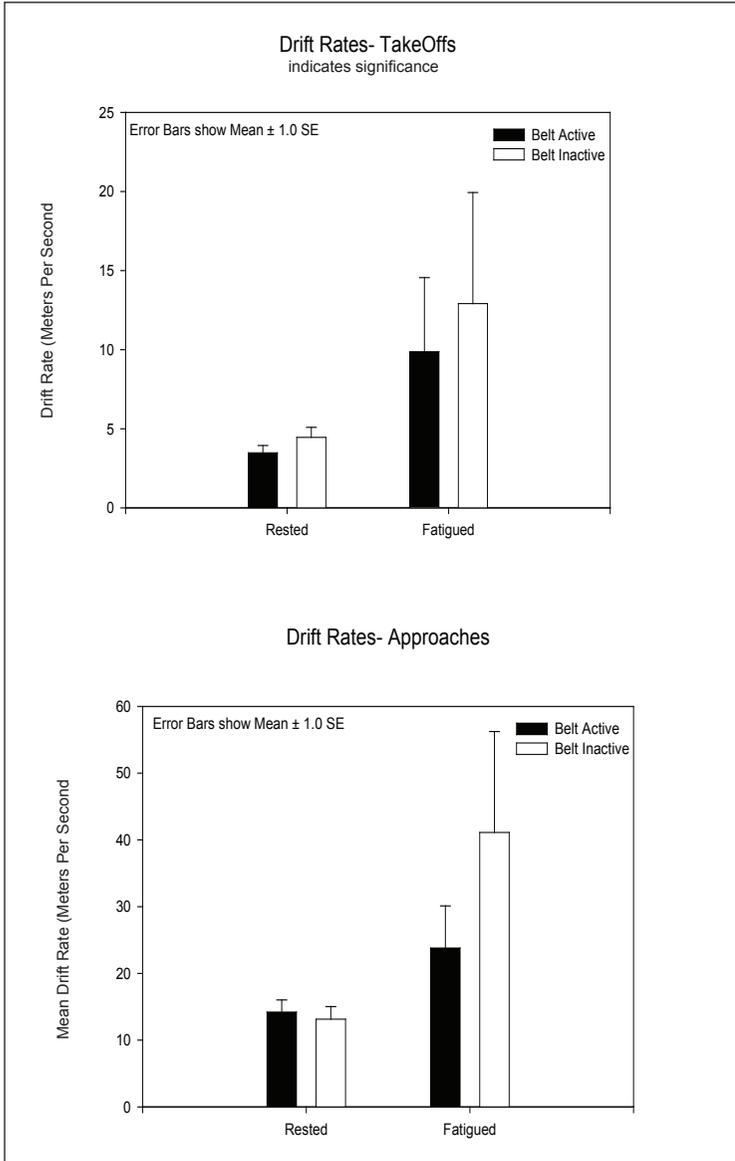
Figure 7. Drift during take-off and landing in rested and fatigued pilots.
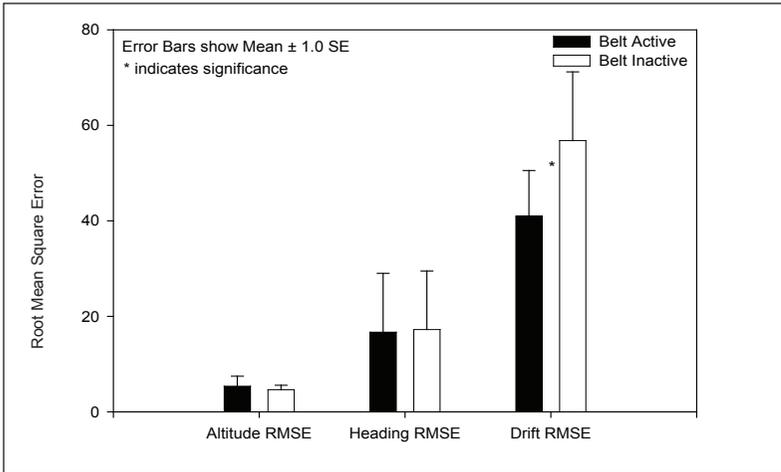
Figure 8. Hover performance.

Analyzing the questionnaire data there were main effects for fatigue, visual workload (p=0.032) and physical workload (p=0.041). In Figure 9, the post-flight questionnaire data for the belt condition main effects are summarized; all were statistically significant (p<0.01) in favor of the active belt.
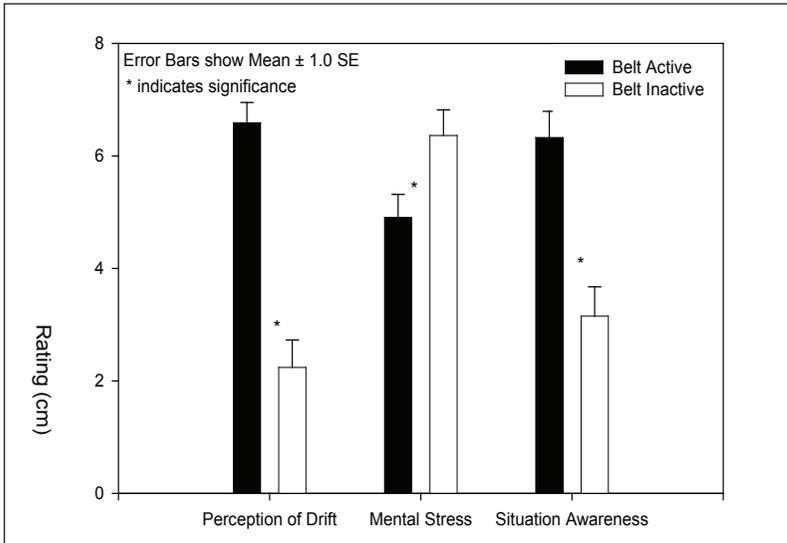
Figure 9. Post flight questionnaire results.

Discussion

This study placed pilots in a situation where their situational awareness was compromised by limiting their orientation cues to information available to them from a standard cockpit display and then adding in a novel tactile cue for aircraft drift. The pilots were then asked to fly 'blind' maneuvers near the ground while

rested and then again when subjected to significant fatigue. The psychometric testing showed significant effects of fatigue with sleepiness and fatigue scores increased and measures of energy and vigor decreased over the course of the study. The pilots' emotional state also changed significantly, specifically as sleep deprivation increased, anger, tension and depression increased and self-confidence decreased. The objective PVT data also showed consistent tendencies to a longer reaction time and more lapses with increased time awake although neither achieved statistical significance. Despite this finding, the overwhelming impression is that these participants were significantly fatigued by the 31 hours of wakefulness they underwent, which relates directly to the current operational conditions military pilots are experiencing

In general, it is expected that poor performance will occur when situational awareness is incomplete or inaccurate (Endsley, 1995), and situational awareness can be compromised when external visual references are poor. Most previous attempts to ameliorate the effects of compromised visual conditions on flight performance have been exclusively based on giving the pilot a visual display with which to maintain his spatial awareness and orientation. The results of the flight portion of the study show that a simple tactile system can provide enough orientation information to the pilot to enable a safe landing with no external reference. In addition, the tactile sense and the interpretation of that sense does not seem to be significantly impaired by fatigue or the stress of the situation. These results are significant for the hover (fatigued and rested) and for takeoff under rested conditions. The results also demonstrate a strong positive impression for both the take-off phase under fatigued conditions and the approach phase of flight.

An important area of the study was to gain the impressions of a selection of pilots with varying levels of experience and they consistently reported that the ability to spend more time visually on other displays while gaining drift information from the tactile instrument resulted in increased situational awareness and reduced workload and mental stress.

Conclusions

The results indicate that modern aircraft instrumentation provides virtually all the pieces of the orientation puzzle with the exception of drift information. When that drift information is added via tactile feedback the limited-display TSAS-Lite can provide increased aircraft control and safety in the critical areas of low speed maneuver near the ground in degraded visual conditions. In fatigued pilots, following 31 hours of sleep deprivation, the TSAS-Lite display helped augment traditional aircraft instruments in an intuitive, non-visual manner. Analysis of the study data showed that the tactile belt significantly improved drift control during takeoff and reduced drift error during hover. In fatigued pilots, all measures of drift were better with the belt versus without the belt. In addition, fatigued pilots reported a significant reduction in visual and physical workload with the belt. The system's value was evident in the consistently positive pilots' perceptions of the system during the course of the study

This study's findings demonstrate the promise of tactile displays and support the continued development of future applications of tactile systems to better orient the aviator and possibly any vehicle operator to a world they cannot fully visualize.

Disclaimer

The opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the U.S. Army and/or the Department of Defense.

References

Bliss, J.C, Katcher, M.H, Rogers, C.H, & Shepard, R.P. (1970). Optical-to-tactile image conversion for the blind. *Man Machine Systems, IEEE Transactions, 11*(1), 58-65.

Braithwaite, M.G, Groh, S, & Alvarez, E.A. (1997). *Spatial disorientation in U.S. Army helicopter accidents: An update of the 1987-92 survey to include 1993-95*. USAARL Report 97-13. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Curry, I.P. & McGhee, J.S. (2007, May). *Spatial Disorientation as a cause of mishaps in combathelicopter operations*. Proceedings of the Aerospace Medical Association Annual Meeting, New Orleans, LA.

Durnford, S.J., Crowley, J.S., Rosado, N.R., Harper, J., & DeRoche, S. (1995). *Spatial disorientation: A survey of U.S. Army helicopter accidents 1987-92*. USAARL Report 95-25. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Loh, S., Lamond, N., Dorrian, J. Roach, G. & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-min duration. *Behavior Research Methods, Instruments, & Computers, 36*(2): 339-346.

McGrath, B.J., Estrada, A., Braithwaite, M.G., Raj, A.K., & Rupert, A.H. (2004). *Tactile situation awareness system flight demonstration final report*. USAARL Report 2004-10. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

McGrath, B.J., Suri, N., Carff, R., Raj, A.K., & Rupert, A.H. (1998). *The role of intelligent software in spatial awareness displays*. 3rd Annual Symposium Exhibition on Situational awareness in the Tactical Air Environment, Patuxent River, MD.

McNair, D.M., Lorr, M., & Droppleman, L.F. (1992). *Manual: Profile of Mood States-Revised*. San Diego, CA: Educationand and Industrial Testing Service.

Penetar, D., McCann, U., Thorne, D., Kamimori, G., Galinski, C., Sing, H., Thomas, M., & Belenky, G. (1993). Caffeine reversal of sleep deprivation effects on alertness and mood. *Pscyhopharm, 112*, 359-365

Raj, A.K., Suri, N., Braithwaite, M.G., & Rupert, A.H. (1998). The Tactile Situation Awareness System in rotary wing aircraft: Flight test results. *Proceedings of the RTA/HFM Symposium on Current Aeromedical Issues in Rotary Wing Operations,* 1-7.

Schmid, H.P. & Bekey, G.A. (1978). Tactile information processing by human operators in control systems. *IEEE Transactions on Systems, Man and Cybernetics, SMC-8, 12,* 860-866.

Sicard, B., Jouve, E., & Blin, O. (2001). Risk propensity assessment in military special operations. *Military Medicine, 166*(10), 871-874.

Van Erp, J.B.F. & Van Den Dobbelsteen, J.J. (1998). *On the design of tactile displays*. Defense Report number TD 98-0266.

General Aviation Pilots' Attitudes
Toward Advanced Cockpit Systems

Stephen M. Casner

National Aeronautics and Space Administration

Abstract

A previous study of general aviation pilots using advanced cockpit systems demonstrated a link between pilots' attitudes and beliefs about advanced systems and their behavior and performance in the cockpit (Casner, 2005). To understand these attitudes and beliefs, a survey was administered to 134 general aviation pilots. The survey explored topics such as pilots' general attitudes toward advanced cockpit systems, how pilots believe these systems affect workload and awareness, pilots' preferences for cockpit systems, pilots' perceptions of risk, long-term effects on pilot skill, and the effects of advanced cockpit systems on the number of errors pilots make as well as the overall safety record. The results show that general aviation pilots hold generally positive attitudes about advanced cockpit systems and exhibit a strong preference for using them. Pilots recognize potential pitfalls associated with advanced cockpit systems but are more likely to ascribe the problems to other pilots than they are to themselves. Overall, general aviation pilots' attitudes were mostly similar to those of airline pilots with a few notable exceptions. A number of contradictory attitudes point out the need for specific future studies to clarify the effect of attitudes and beliefs on pilot behavior and ultimate safety outcomes.

In a previous study, pilots who used GPS and moving map displays estimated their navigational awareness to be greater than that of pilots who navigated using a sectional chart and pilotage in a conventional cockpit (Casner, 2005). Believing their awareness to be superior in the presence of a GPS and moving map, these pilots appeared to assume a less active role in the navigation process. When put to a practical test of navigational awareness, these pilots in fact performed worse than pilots who used pilotage to navigate, and quickly lowered their awareness estimates. Pilots who navigated using a sectional chart and pilotage performed better than they expected, and subsequently raised their estimations of their own

awareness (Casner, 2005). These results suggest that pilots' beliefs and attitudes about advanced cockpit systems can sometimes be powerful determinants of pilot behavior and performance in the cockpit.

The purpose of this study was to capture pilots' beliefs and attitudes about advanced cockpit systems that are rapidly becoming commonplace in general aviation aircraft. The survey was designed with four goals in mind.

A first goal of the survey was to measure pilots' basic attitudes toward and beliefs about advanced cockpit systems. For example, would pilots rather fly an advanced cockpit aircraft than a conventional aircraft? Do advanced cockpit systems make pilots feel safer? If so, which systems most contribute to a feeling of greater safety? Do pilots believe they can navigate more accurately with GPS and lower their workload by using an autopilot? Do pilots believe that advanced cockpit systems will increase or decrease the number of errors that they make? Do pilots believe that advanced cockpit systems will result in an increase or decrease in the number of accidents?

A second goal was to discover if pilots view advanced cockpit systems as affecting themselves differently than the way they perceive the same systems affecting other pilots. It is well known that people sometimes inflate estimations of their own abilities (Kruger and Dunning, 1999; Sulheim, Ekeland, and Bahr, 2006), and believe that negative outcomes are less likely to happen to them then they are to others (Bodner et al, 2000; Burger and Burns, 1988; Greening and Dollinger, 2005).

A third goal of the survey was to discover any relationships between pilot experience and attitudes toward advanced cockpit systems. Does experience with using advanced cockpit systems change one's attitudes toward them?

A fourth goal of the survey was to compare the attitudes and beliefs of general aviation pilots with those of airline pilots who operate advanced cockpit systems found in commercial jet transports. Three previous surveys done with airline pilots facilitate this comparison. Wiener (1985) surveyed pilots of an early-generation automated airplane (MD-80) equipped with a navigation computer and autopilot, but with standard flight and navigation instruments. BASI (1998) and Hutchins et al (1999) surveyed pilots of later generation airplanes that contained a full suite of glass-cockpit avionics.

Survey Methodology

Participants

One hundred thirty-four general aviation pilots were recruited on a volunteer basis from California Bay Area flight schools, fixed-based operators, and flight safety seminars. Criteria for inclusion were that each participating pilot held at least an FAA private pilot certificate with an airplane category and class rating. Student pilots were excluded from the sample, as were pilots who currently worked in Part 121 airline service.

Apparatus

A paper and pencil survey was used to collect the data. The survey contained four items that asked participants about their flight experience and certificates and ratings held, 52 attitude probes, and five multiple-choice or supply-type questions. Roughly, half of the attitude probes were worded in a positive tone while the other half were worded in a negative tone. Each attitude probe made a short statement about advanced cockpit systems and asked pilots to respond using a five-point Likert scale: (Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree).

Procedure

Pilots were approached in person by the experimenter and asked to participate in the survey. All pilots that agreed to complete the survey did so at the same time they were recruited. There was no time limit for completing the survey and most pilots finished within approximately 15 minutes. Pilots were told that their survey responses would remain anonymous. Pilots received a NASA coffee mug as compensation for completing the survey.

Results

Pilot Demographics

Certificates Held

Table 1 shows the proportion of surveyed pilots who held each type of FAA certificate along with the proportions for all U.S. (non-student) pilots (AOPA, 2007).

Table 1
Certificates and ratings held by survey participants and all active U.S. pilots.

Certificate/Ratings	% of Sample	% Active U.S. (2006)
Private and Private with Instrument Rating	66.4%	46.2%
Commercial	26.1%	25.5%
ATP	7.5%	28.3%
Flight Instructor	13.4%	14.5%

Note that the proportion of private pilots in our sample is roughly 20% larger than that of the population at large, and the proportion of airline transport pilots is roughly 20% smaller. Indeed, the goal was to exclude pilots who work in airline service from the sample. Note that certificates held is at best a correlate of which type of job any pilot presently holds. That is, holding an airline transport certificate does not guarantee that a pilot works in airline service, nor does working in airline service guarantee that a pilot holds an airline transport pilot certificate.

Flight Time

Pilots who completed the survey had a mean of 1,589 hours of total flight time

(SD = 2,841). The median total flight time was 650 hours. The minimum and max-imum flight times were 76 hours and 19,309 hours.

Pilots had a mean of 496 hours (SD = 919) in airplanes that contained at least a panel-mounted GPS navigation computer. The median was 155 hours. The min-imum and maximum were 0 and 5,000 hours.

Pilots had a mean of 59 hours (SD = 370) in glass-cockpit airplanes: those that feature a GPS navigation computer, electronic flight instruments, moving map dis-play, autopilot, etc. The minimum and maximum were 0 and 4,200. The median was 0, meaning that more than one-half of the pilots surveyed had no time logged in glass cockpits. Since this makes for a rather non-normal distribution, all correla-tional comparisons made with this variable throughout the paper were performed using only those pilots who had at least one hour of flight experience in glass cock-pits.

Responses to Survey Items

The following organizes the survey items and results into nine topic areas:

1. General Attitudes about Advanced Cockpit Systems
2. Workload
3. Awareness
4. Learning
5. Retention
6. Error
7. Safety
8. Preferences for In-Flight Use
9. Overall Preferences

Descriptive statistics are reported using the coding convention that a response of Strongly Agree is scored as 5, and a response of Strongly Disagree is scored as 1. In the case that a survey item was left blank, no score was recorded.

1. General Attitudes about Advanced Cockpit Systems
Six survey items queried pilots about their general attitudes toward advanced cockpit systems.

"I look forward to new kinds of advanced cockpit systems."
Responses from general aviation pilots in this survey differed striking from the airline pilots surveyed in a previous study with regard to this statement. Note that two results are shown for the survey done by Wiener (1985). Wiener's pilots were surveyed once after completing an initial training program on a jet transport equipped with early-generation cockpit automation (MD-80), and then again after acquiring some experience with the aircraft. A chi-square test revealed that GA pilots surveyed here were significantly more enthusiastic than Wiener's airline pilots: $X^2(8)$ = 86.79, p < .01, as shown in Figure 1 (M = 4.17, SD = 0.67). Note that Wiener's airline pilots had significantly more overall flight experience (M = 14,568, SD = 7,966) than the general aviation pilots surveyed here (M = 1,589, SD = 2,841).
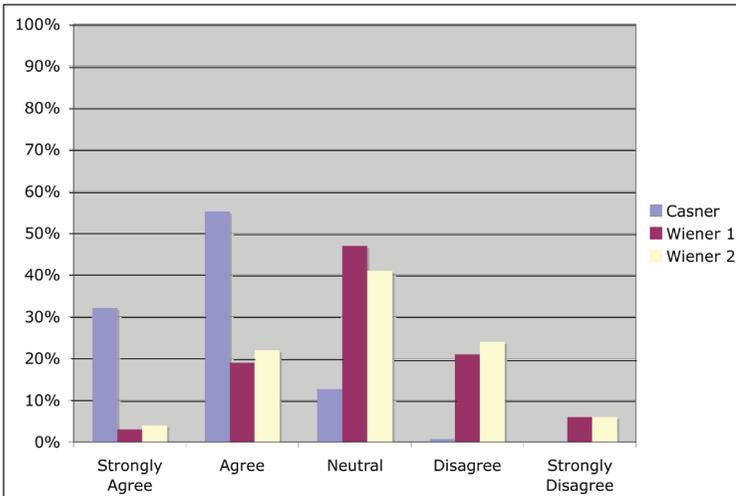
Figure 1. "I look forward to new kinds of advanced cockpit systems."

"They've gone too far with advanced cockpit systems."
    Pilots in this study disagreed with this statement as shown in Figure 2 (*M* = 2.14, *SD* = .80). Figure 2 shows this result along with the results of two earlier surveys of airline pilots' opinions (BASI, 1998; Wiener, 1985). Airline pilots in both of these studies disagreed with the statement to roughly the same extent as our general aviation pilots.
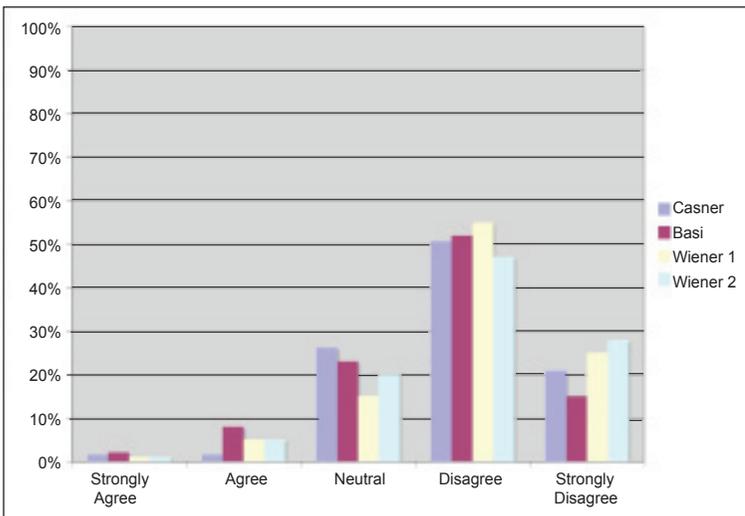
Figure 2. "They've gone too far with advanced cockpit systems."

"I sometimes feel more like a 'button pusher' than a pilot"
    General aviation pilots were neutral and leaned toward disagreement with this attitude as shown in Figure 3 (*M* = 2.63, *SD* = .93). Disagreement with

this probe was significantly correlated with higher flight time in glass cock-pits: $r(132) = -0.30$, $p < .05$, and flight time in airplanes with at least a GPS: $r(132) = -0.15$, $p < .05$. This result is consistent with that obtained in Wiener's survey of MD-80 pilots, and with that obtained in the Hutchins et al (1999) survey of pilots of later-generation automated airliners (B-757, A-320, etc.).
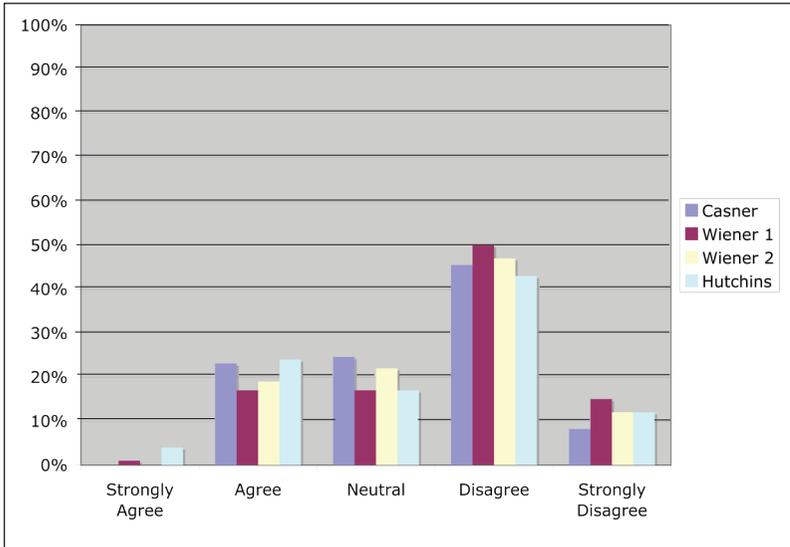
Figure 3. "In an advanced cockpit, I sometimes feel more like a "button pusher" than a pilot."

Figure 4 shows the results for three additional survey items that probed pilots' general attitudes toward advanced cockpit systems. There was a modest but significant correlation between responses to the first probe and total flight time: $r(132) = .16$, $p < .05$). Greater flight experience was associated with agreement with the attitude that advanced cockpit systems are becoming too complex.
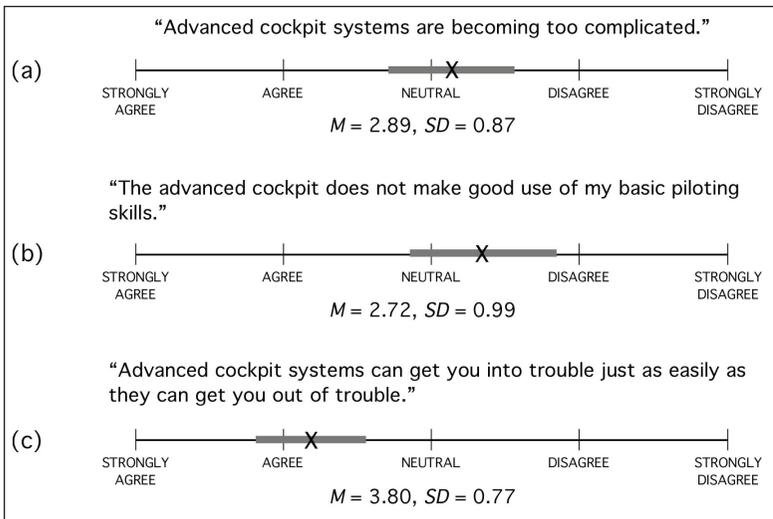
Figure 4.

Summary. The results suggest that general aviation pilots hold generally positive attitudes toward advanced cockpit systems. The results further suggest that these attitudes might be tempered both by overall flight experience and by experience in advanced cockpit aircraft. Greater experience appears to be associated with greater concerns about the complexity of the systems.

2. Workload

Five survey items explored pilots' attitudes toward the pros and cons of how advanced cockpit systems might affect pilot workload.

"Using the autopilot lowers my workload"
"Navigating using GPS lowers my workload"
"I can better control my workload in an advanced cockpit"

The strongest agreement with any item in the survey was seen in response to the statement that using an autopilot lowers workload, as shown in Figure 5 ($M = 4.5$, $SD = .56$). Agreement was significantly correlated with total flight time: $r(132) = 0.15$, $p < .05$, time in airplanes that contained at least a GPS: $r(132) = 0.18$, $p < .05$, and pilot certificate held: $r(132) = 0.31$, $p < .01$. Pilots also agreed that using GPS lowers workload ($M = 3.98$, $SD = .73$), and that they can better control workload in an advanced cockpit ($M = 3.72$, $SD = .83$).
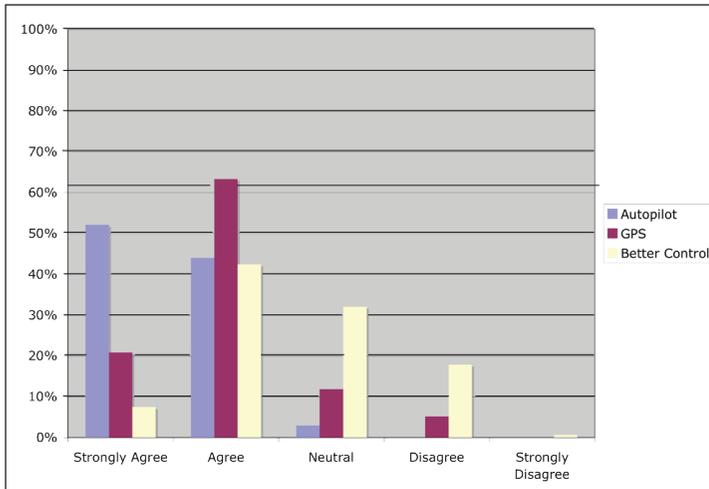
Figure 5. Three attitude probes about the effect of advanced cockpit systems on pilot workload.

"I sometimes spend more time setting up and monitoring the autopilot than I would just hand-flying the aircraft."

General aviation pilots were neutral and to some degree split with respect to this probe as shown in Figure 6 ($M = 2.67$, $SD = 1.02$). General aviation pilots responded similarly to airline pilots surveyed in previous studies. Agreement was modestly correlated with time spent in glass cockpits: $r(132) = 0.22$, $p < .05$.
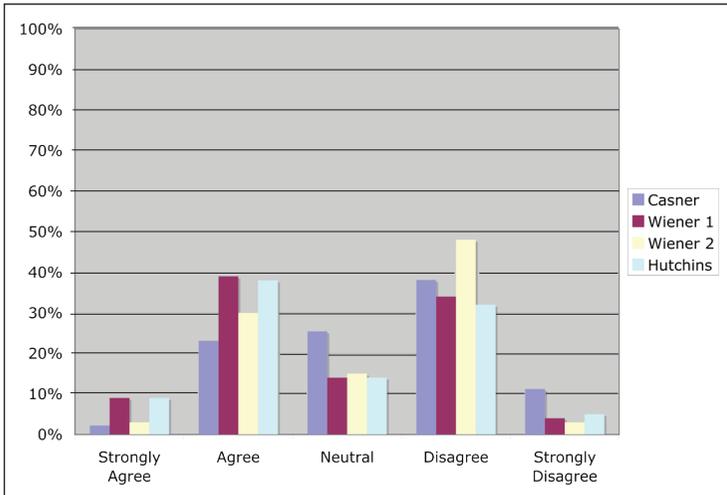
Figure 6. "I sometimes spend more time setting up and monitoring the autopilot than I would just hand-flying the aircraft."

"There are too many alerts and warning noises in the advanced cockpit"
    Pilots responded neutrally to this statement (M = 2.86, SD = .87).

    Summary. These findings suggest that pilots have a strong general belief that autopilots and navigation computers are effective tools for reducing pilot workload. Pilots responded neutrally to two probes that suggested potential disadvantages associated with two advanced cockpit systems with respect to pilot workload.

3. Awareness
    Eight survey items probed pilots' attitudes toward the impact of advanced cockpit systems on pilot awareness.

"My situational awareness is better in an advanced cockpit"
    Eighty-five percent of all pilots surveyed either agreed or strongly agreed with this statement (M = 4.15, SD = .82). Agreement with this statement was modestly correlated with time spent in airplanes that contained a GPS navigation computer: $r(132)$ = .16, p < .05.

"The pilot that uses pilotage (a sectional chart) is going to have better navigational awareness than one who uses a GPS and moving map display"
        Pilots were neutral but leaned toward disagreement with this statement (M = 2.75, SD = 1.09). These attitudes are consistent with those found by Casner (2005): pilots felt that having a GPS and moving map was superior to having a paper chart.

"If you turn off my GPS and moving map during a flight, I might be lost" vs. "For some pilots, turn off their GPS and moving map during a flight, and they might be lost."
        Figure 7 shows responses to these two survey items designed to compare pilots' beliefs about the potential to become lost when their GPS and moving map display failed. The means for the two items were (M = 2.02, SD =

.88 and *M* = 3.79, *SD* = .70). A significant difference was found between what pilots believed about themselves and what they believe to be true of other pilots: *t*(132) = -18.13, *p* < .005. There is a sizeable literature on what psychologists have termed *unique invulnerability,* the perception that an individual is less vulnerable to negative events than others. Unique invulnerability has been demonstrated in a variety of situations including skydiving (Bodner et al, 2000), the decision to use birth control (Burger and Burns, 1988), and preparing for natural disasters (Greening and Dollinger, 2005).
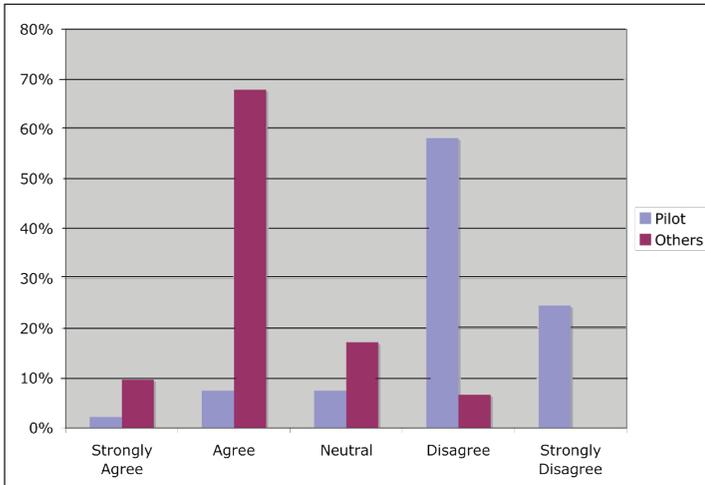
Figure 7. "If you turn off my GPS and moving map during a flight, I might be lost vs. For some pilots, turn off their GPS and moving map during a flight, and they might be lost."

"When I have a traffic alerting system on board, I look out the window less often"
<div align="center">*vs.*</div>
"Pilots who use traffic alerting systems have a tendency to look out the window less often."

Figure 8 shows the responses to the first of several pairs of statements aimed at comparing pilots' beliefs about themselves and their beliefs about other pilots. Items such as these were distributed randomly throughout the survey to minimize the chances that participants would sense that such a comparison was being made. Again, there was a significant difference between the responses for these two probes (*M* = 2.68, *SD* = .86 and *M* = 3.23, *SD* = .85) that queried pilots about a tendency to look out the window less often when a traffic alerting system is available: *t*(133) = -5.24, *p* < .005. Pilots more readily ascribed the problem of decreased vigilance to other pilots than they did to their own behavior. Interestingly, pilots who have spent more time in glass cockpit aircraft trended away from agreement with these two statements: *r*(132) = -0.40, *p* < .01 and : *r*(132) = -.20, *p* < .05.
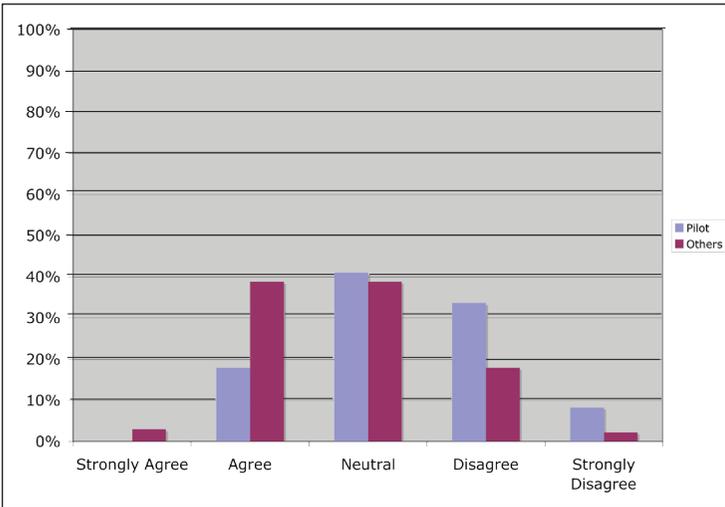
Figure 8. "When I have a traffic alerting system on board, I look out the window less often." vs. "Pilots who use traffic alerting systems have a tendency to look out the window less often."

"I always know what mode the GPS and autopilot are in"

General aviation pilots agreed with this statement as shown in Figure 9 (M = 3.56, SD = 1.03). Agreement with this attitude was correlated with time spent in airplanes that contained at least a GPS navigation computer: $r(132)$ = .16, $p < .05$. Figure 13 shows responses from the GA pilots as well as responses collected from airline pilots in three previous surveys. In an empirical study of pilots using advanced cockpit in flight, Casner (2004) demonstrated that lack of mode awareness is often a persistent problem for general aviation pilots who are learning to fly instrument procedures using GPS to navigate.
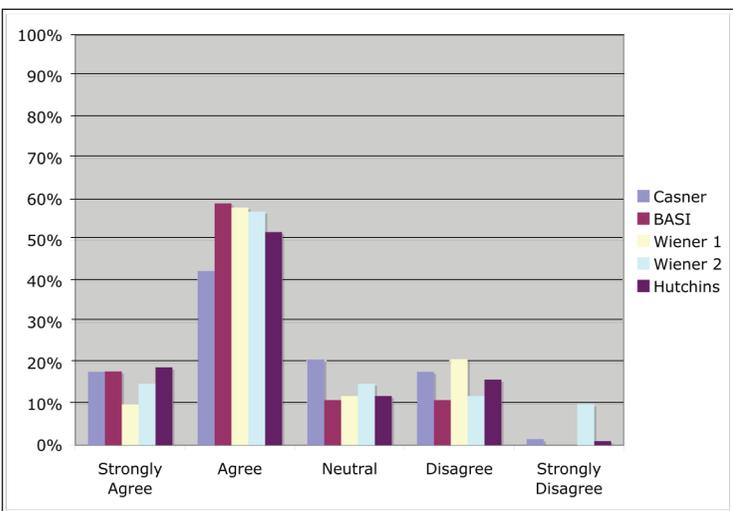
Figure 9. "I always know what mode the GPS and autopilot are in."

"It worries me that the GPS, autopilot, or other system may be doing something that I don't know about."

Pilots were neutral and leaned toward disagreement with this statement as shown in Figure 10 ($M = 2.68$, $SD = 1.05$). Figure 10 shows these responses along with those from a previous survey of airline pilots (BASI, 1998).
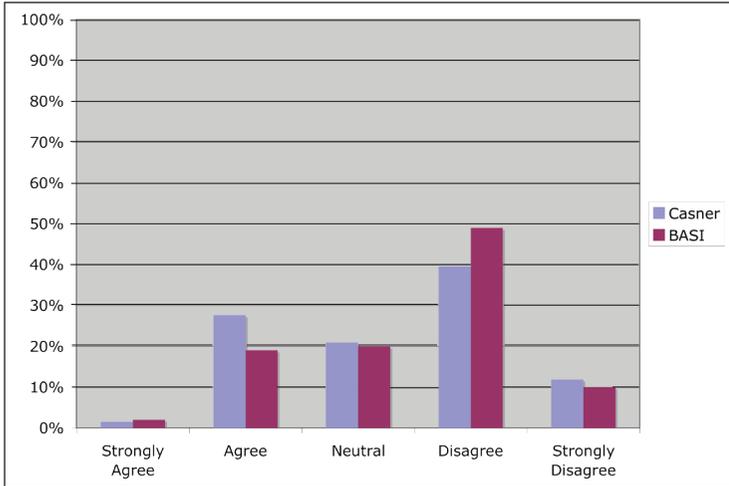
Figure 10. "It worries me that the GPS, autopilot, or other system may be doing something that I don't know about."

Summary. Pilots have a strong general belief that advanced cockpit systems raise their level of awareness. This attitude seems to include not only navigational awareness, but also an awareness of the operating modes and other behaviors of advanced cockpit systems. Pilots seem to recognize an "out-of-the-loop" phenomenon associated with the use of advanced cockpit systems but are more likely to ascribe the problem to other pilots than they are to themselves. Similarly, pilots acknowledge a tendency to look out the window less often when a traffic alerting system is available but are much more likely to ascribe the problem to other pilots.

4. Learning

Eleven survey items were designed to measure pilots' attitudes about how advanced cockpit systems might affect the way pilots train and maintain proficiency.

"There are still features of the advanced cockpit that I don't understand"

Pilots generally agreed with this statement as shown in Figure 20 ($M = 3.66$, $SD = .90$). As expected, agreement with this probe was negatively correlated with time in glass cockpits: $r(132) = -.30$, $p < .01$, and airplanes that contained at least a GPS navigation computer: $r(132) = -.23$, $p < .01$. Figure 11 presents the responses to this probe along side the results of three previous surveys done with airline pilots. In contrast to the general aviation pilots, airline pilots leaned toward disagreement with this probe in

the three earlier surveys. Of particular interest are the responses of Wiener's first group of airline pilots who were also relatively new to advanced cockpit systems (Wiener 1 in Figure 11).
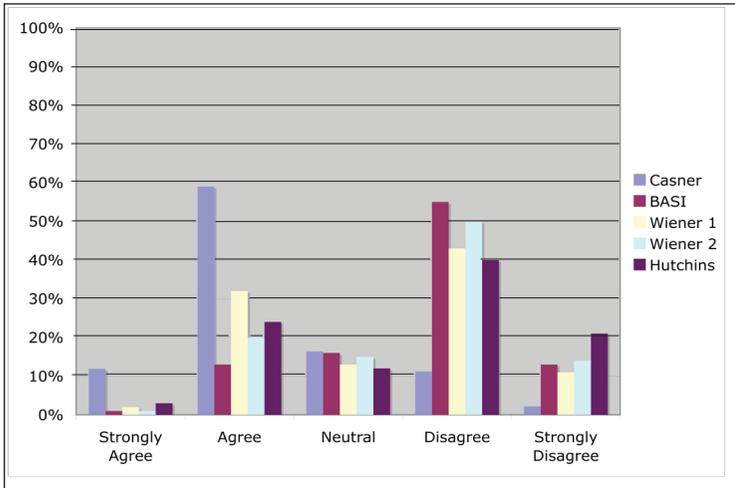
Figure 11. "There are still features of the advanced cockpit that I don't understand."

Figure 12 shows the results of four other attitude probes related to learning to use advanced cockpit systems. Of particular interest are the responses to the probe shown in Figure 12(a). Seventy percent of all pilots surveyed foresee a significant problem with versus using only advanced cockpit aircraft to train new pilots. There were no significant correlates between responses to these items and flight experience.
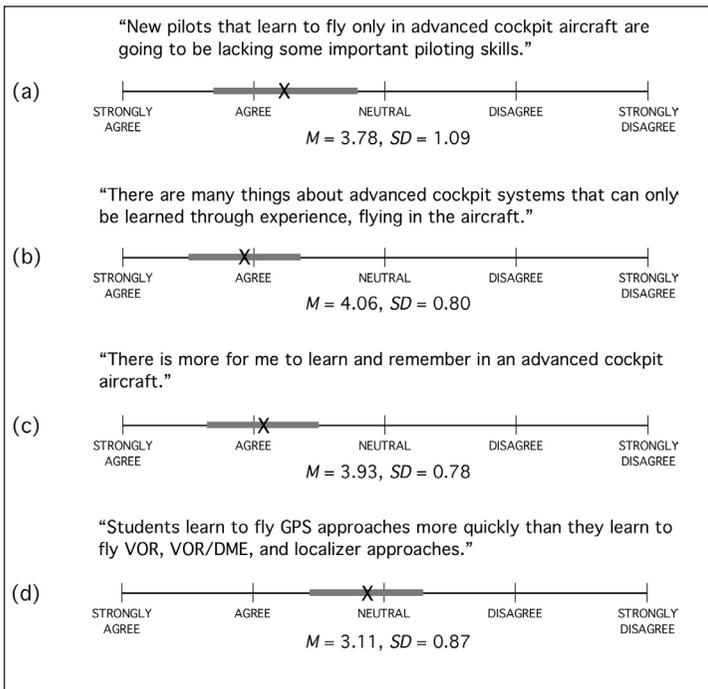
Figure 12.

"I found everything I needed to know about advanced cockpit systems in the manufacturer's technical manuals"

General aviation pilots disagreed with this statement, as shown in Figure 13 (M = 2.47, SD = .84), and agreement was correlated with time in glass cockpits $r(132)$ = -.20, p < .05, and airplanes that contained at least a GPS navigation computer: $r(132)$ = -.16, p < .05. The results shown in Figure 13 are presented with responses to similar probes collected from airline pilots in a previous study (BASI, 1998). Note that airline pilots were somewhat divided on this issue.
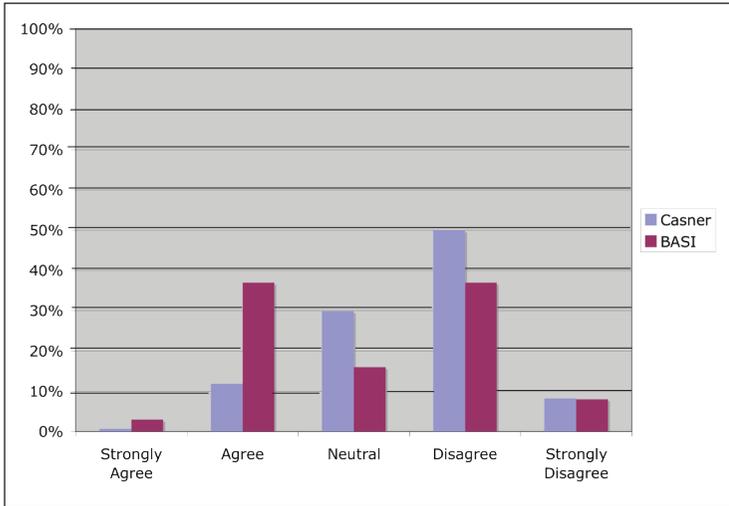
Figure 13. "I found everything I needed to know about advanced cockpit systems in the manufacturer's technical manuals."

Figure 14 shows the results of five survey items that probed pilots' attitudes toward the role of the FAA in training and certifying pilots who operate advanced cockpit aircraft.

"The FAA has provided pilots, students, instructors, and evaluators with sufficient guidance about flying advanced cockpit aircraft."

(a) STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 2.48$, $SD = .77$

"The FAA should publish a new pilot handbook to cover advanced cockpit systems."

(b) STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 3.83$, $SD = 0.85$

"The practical test standards should be expanded to include skills specific to advanced cockpit aircraft."

(c) STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 3.77$, $SD = 0.88$

"The FAA pilot knowledge tests should include questions about advanced cockpit systems."

(d) STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 3.75$, $SD = 0.82$

"Pilots should not be allowed to act as PIC in advanced cockpit aircraft unless they get an endorsement similar to the one required for high-performance and complex airplanes."

(e) STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 3.47$, $SD = 0.96$

Figure 14.

Interestingly, 58% of all pilots surveyed either agreed or strongly agreed with the need for additional training and a logbook endorsement for pilots who wish to operate advanced cockpit aircraft.

Summary. Pilots seem to believe there are unique learning challenges posed by the advanced cockpit, and may believe that using advanced cockpit aircraft only to train new pilots may be inappropriate. Pilots trended toward feeling more knowledgeable when experience using advanced cockpit systems was greater. Pilots generally feel that the FAA should assume a stronger role in the training and certification of pilots who operate advanced cockpit aircraft.

5. Retention
Four survey items probed pilots' attitude toward retaining knowledge and skills related to advanced COckpit systems.

"I am concerned that I might become too dependent on GPS, autopilots, and other advanced cockpit systems." vs. *"I am concerned that today's pilots may become too dependent on GPS, autopilots, and other advanced cockpit systems"*
Responses to these two attitude probes are shown in Figure 15 ($M = 3.17$, $SD = 1.1$ and $M = 3.43$, $SD = 1.1$). There was a small but significant differ-

ence between what pilots believe about themselves and other pilots: $t(133)$ = -1.90, $p < .05$.
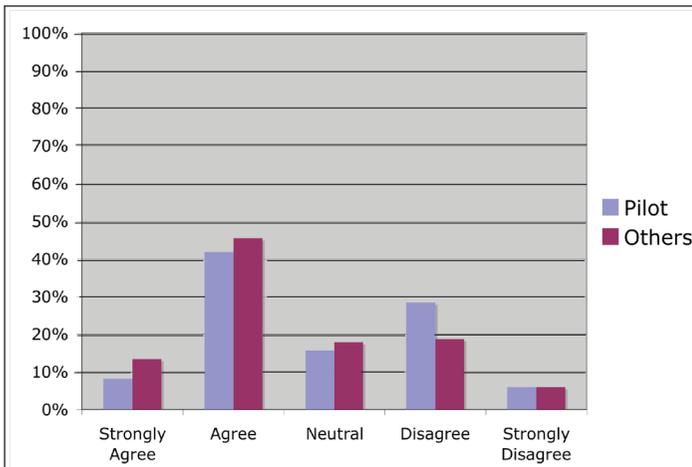
Figure 15. "I am concerned that I might become too dependent on GPS, autopilots, and other advanced cockpit systems." vs. "I am concerned that today's pilots may become too dependent on GPS, autopilots, and other advanced cockpit systems."

"I am concerned that flying advanced cockpit aircraft will cause my basic flying skills to deteriorate"

Pilots responded neutrally to this statement as shown in Figure 16 ($M = 2.79$, $SD = 1.08$). A similar probe was used in previous surveys of airline pilots and the results are shown in Figure 16.
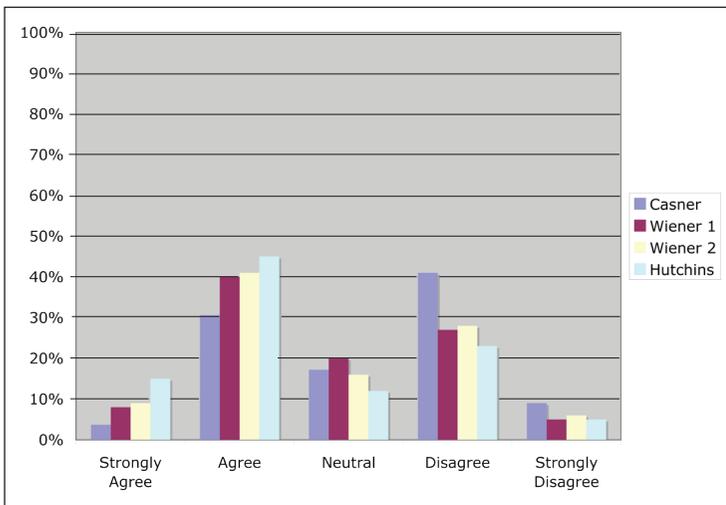
Figure 16. "I am concerned that flying advanced cockpit aircraft will cause my basic flying skills to deteriorate."

"I need to fly more often to maintain proficiency in an advanced cockpit than I do in a conventional aircraft"

Pilots were neutral but learned toward agreement with this statement ($M = 3.40$, $SD = .88$).

Summary. Roughly, half of the pilots surveyed felt that advanced cockpit systems require extra practice to maintain proficiency, although this attitude does not appear

(a) "Advanced cockpit systems are going to reduce the number of errors that pilots make."

STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 3.18$, $SD = .97$

(b) "I am less likely to make a navigational error or bust an altitude in an advanced cockpit."

STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 3.58$, $SD = .95$

(c) "Using GPS is going to result in fewer navigational errors."

STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 4.19$, $SD = .78$

(d) "Advanced cockpit systems create opportunities to make new kinds of errors."

STRONGLY AGREE — AGREE — NEUTRAL — DISAGREE — STRONGLY DISAGREE
$M = 4.07$, $SD = .60$

Figure 17.

"Incorrect data entered by mistake is easy to detect in the advanced cockpit"

Pilots were divided in their responses to this statement as shown in Figure 18 ($M = 2.86$, $SD = .81$). Figure 18 shows these pilots' responses along with the similar responses collected from airline pilots in a previous survey (BASI, 1998).
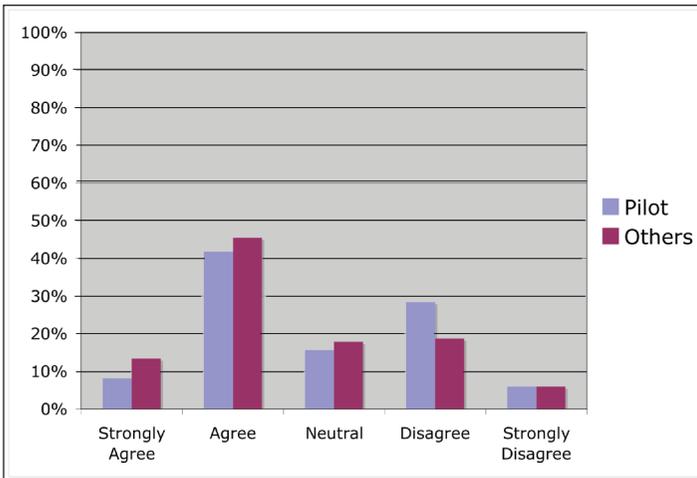
Figure 18. "Incorrect data entered by mistake is easy to detect in the advanced cockpit."

Summary. Pilots differ in their beliefs about how advanced cockpit systems will affect pilot error but do appear to generally believe that GPS will help eliminate specific navigational errors. Some pilots believe that advanced cockpit systems help them discover mistaken entries more easily while other pilots do not.

7. Safety
Seven probes explored pilots' beliefs about how specific advanced cockpit systems would ultimately affect safety.

"I feel safer in an advanced cockpit aircraft than I do in a conventional aircraft"
"I feel safer in any aircraft that has a parachute (ballistic recovery system) for the airframe"
Pilots responded neutrally to both of these statements as shown in Figure 19 ($M = 3.2$, $SD = .97$ and $M = 2.66$, $SD = 1.0$).
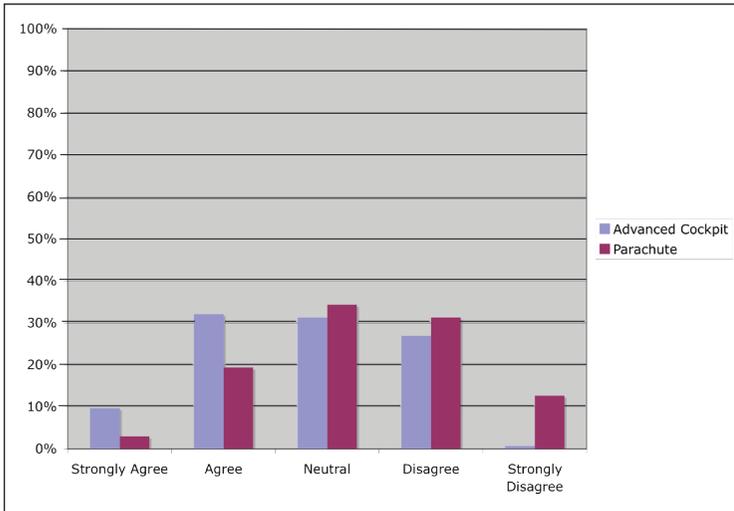
Figure 19. Two survey items that probed pilots about perceived safety.

"Terrain displays in the cockpit are going to reduce the number of controlled flight into terrain (CFIT) accidents"
"Traffic alerting systems are going to reduce the number of mid-air collisions"
"Cockpit weather systems are going to reduce the number of weather-related accidents"
"GPS is going to reduce the number of accidents"

Pilots did agreed with the first three of these statements and were divided in their responses to the fourth statement as shown in Figure 20. The means for the responses for the three items were $M = 3.80$, $SD = .85$; $M = 3.71$, $SD = .76$; $M = 3.49$, $SD = .91$; and $M = 3.08$, $SD = .96$.
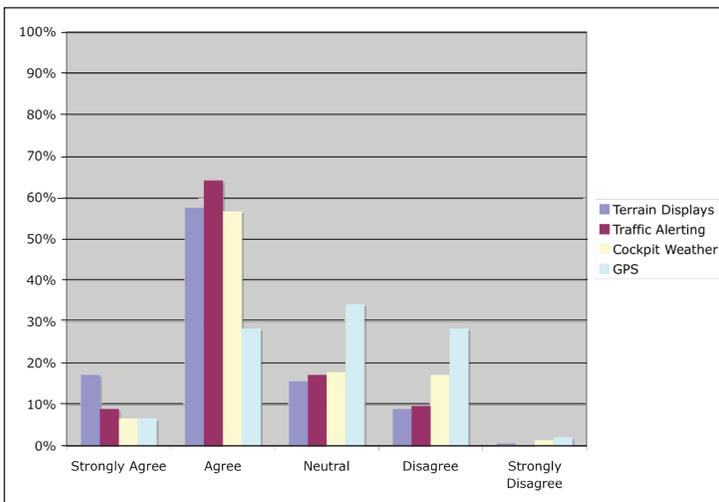
Figure 20. Four survey items about the impact of specific advanced cockpit systems.

Agreement with the first two statements was positively correlated with time in glass cockpits: $r(132) = .24$, $p < .05$ and $r(132) = .25$, $p < .05$. Agreement with the second two statements was positively correlated with time in airplanes having at least a GPS navigation computer: $r(132) = .17$, $p < .05$, and $r(132) = .20$, $p < .05$.

"Some pilots will misuse advanced cockpit systems to stretch the boundaries of safety"

Pilots agreed with this cautionary statement ($M = 3.88$, $SD = .63$). 80% of pilots surveyed either agreed or strongly agreed with this statement.

"How do you think advanced cockpit systems will affect the number of aircraft accidents?"

Figure 21 shows pilots' responses to this question. Overall, pilots agreed with the attitude that advanced cockpit systems will "somewhat decrease accidents."
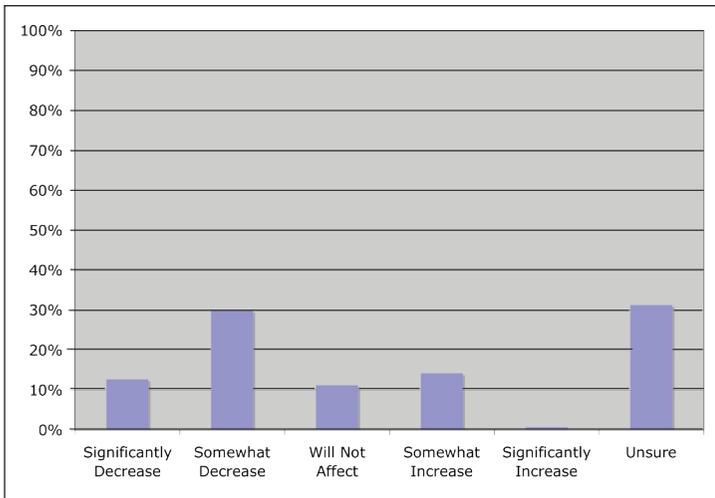
Figure 21. "How do you think advanced cockpit systems will affect the number of aircraft accidents?"

This statement yielded the largest number of significant correlations with other survey items. These correlations help provide insight into the reasons how and why pilots believe advanced cockpit systems will affect the safety record. Beliefs about how advanced cockpit systems will affect the accident rate correlated with survey items that stated that advanced cockpit systems would result in increased navigational awareness: $r(132) = .42$, $p < .01$; increased mode awareness: $r(132) = .22$, $p < .05$; fewer overall errors: $r(132) = .36$, $p < .01$; fewer navigational errors: $r(132) = .22$, $p < .05$; and decreased pilot workload: $r(132) = .38$, $p < .01$. This same measure correlated with attitudes that GPS: $r(132) = .23$, $p < .05$; terrain displays: $r(132) = .25$, $p <)$; weather systems: $r(132) = .26$, $p < .01$; and parachute systems: $r(132) = .26$, $p < .01$ would lead to fewer accidents. This

same measure was negatively correlated with attitudes that advanced cockpit systems did not make good use of their basic piloting skills: $r(132) = -.30$, $p < .01$; that pilots were too dependent on advanced systems: $r(132) = -.30$, $p < .01$; that advanced systems could sometimes get pilots into trouble: $r(132) = -.30$, $p < .01$, and that some pilots would use systems to stretch the boundaries of safety: $r(132) = -.20$, $p < .05$. These data suggest that there are pilots who have a "positive mindset" toward advanced cockpit systems.

Summary. Pilots were divided in the belief that they are safer in aircraft that contain advanced cockpit systems, 43% believed that advanced cockpit systems would reduce the number of accidents to some degree. Pilots associated specific factors (e.g., lower workload, fewer errors, increased awareness) and specific systems (e.g., terrain displays, traffic alerting systems, weather systems) with potential reductions in the number of accidents. Pilots acknowledged the potential for misusing advanced cockpit systems to do things they would not typically do without them.

8. Preference for In-Flight Use
Six survey items explored pilots' preferences for when to use advanced cockpit systems during flight.

"I prefer to use the autopilot during periods of high workload"
"I prefer to hand-fly the aircraft (autopilot off) during periods of low workload."
These two statements were designed to measure pilots' overall attitudes toward autopilots and workload. Pilots responded in agreement with the first statement ($M = 4.34$, $SD = .76$) and neutrally to the second statement ($M = 3.38$, $SD = 1.02$) as shown in Figure 22.
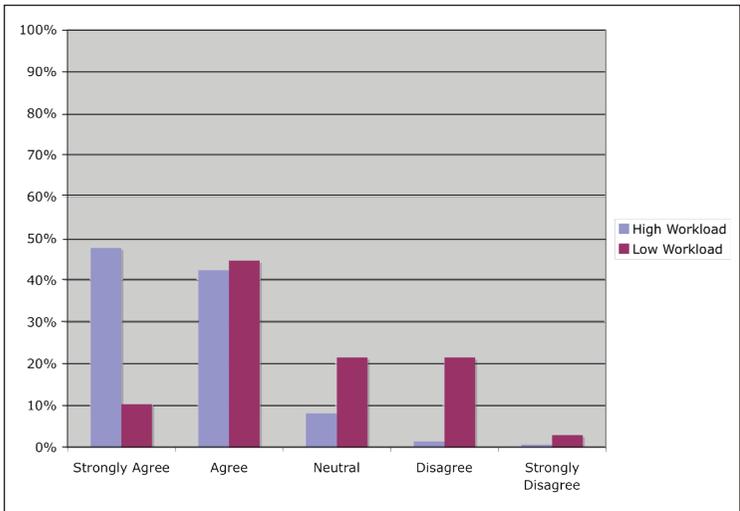
Figure 22. Pilot preferences for using the autopilot during high and low workload periods.

"I prefer to use the autopilot when flying en route"
"I prefer to use the autopilot when flying an instrument approach"
"I prefer to use the autopilot during a missed approach procedure"

Given that pilots claim to prefer to use the autopilot during high workload situations, these statements probe pilots' preferences for specific phases of flight that have different workload characteristics. Interestingly, pilots' agreed that they prefer to use the autopilot during the generally low-workload en route phase (M = 4.03, SD = .82), but responded neutrally to the statements about the two higher workload phases (M = 3.37, SD = 1.02; M = 2.98, SD = .96) as shown in Figure 23. These responses are contradictory to pilots' responses to the two previous items. This result is similar to the findings of a simulator study done by Kirlik (1993) in which pilots exhibited a preference for using the autopilot mostly during periods of lower workload.
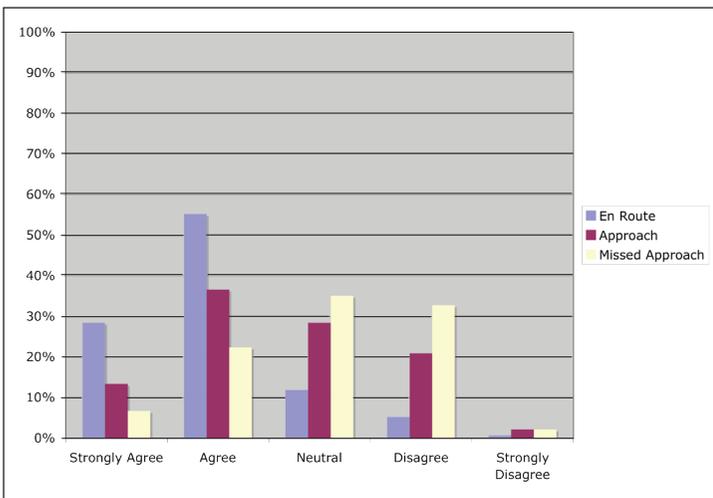
Figure 23. Pilots' preferences for using the autopilot during different phases of flight.

"I would rather use GPS than VORs to navigate"

Pilots agreed with this statement (M = 4.28, SD = .77). 86% of all pilots surveyed either agreed or strongly agreed with this statement.

9. Overall Preferences
"If you could have ONE advanced system in your cockpit, it would be a(n):"
"If you could add a SECOND advanced system to your cockpit, it would be a(n):"
"If you could add a THIRD advanced system to your cockpit, it would be a(n):"

These item asked pilots to choose, in order of decreasing preference, three advanced cockpit systems to have in their own aircraft. Pilots were presented with a list of six advanced cockpit systems to choose from: (GPS, autopilot, moving map, terrain warning, hazardous weather display, traffic-

alerting system). The results, shown in Table 2, indicate that a GPS navigation computer was the first choice for most pilots, while an autopilot and moving map display were the second and third favorite choices, respectively.

Table 2
Pilots' preferences for advanced cockpit systems

Preferences	GPS	Autopilot	WX Display	Moving Map	Terrain Display	TCAS
1st Choice	72	33	0	15	3	12
2nd Choice	37	44	15	23	2	14
3rd Choice	2	24	33	38	17	20

"If you could choose, what percentage of your flying time would you spend in advanced cockpit aircraft and what percentage in conventional cockpit aircraft?"
The median response for advanced cockpit aircraft was 90% with a mean of 74.1% (SD = 30.5), a minimum of 0, and a maximum of 100. The median response for conventional aircraft was 10% with a mean of 25.9% (SD = 30.42), a minimum of 0, and a maximum of 100.

Pilots' preferences for advanced vs. conventional cockpits yielded significant correlations with many of the same attitude probes. Higher percentages of preferred time in advanced cockpit aircraft were correlated with the belief that advanced cockpit systems would increased navigational awareness: $r(132)$ = .36, $p < .01$; increased mode awareness: $r(132)$ = .21, $p < .01$; fewer overall errors: $r(132)$ = .22, $p < .01$; and fewer navigational errors: $r(132)$ = .23, $p < .01$. This same measure correlated with attitudes that terrain displays and parachute systems would lead to increased safety: $r(132)$ = .26, $p < .01$, $r(132)$ = .17, $p < .05$. This same measure was negatively correlated with attitudes that advanced cockpit systems did not make good use of their basic piloting skills: $r(132)$ = -.22, $p < .01$; and that pilots were too dependent on advanced systems: $r(132)$ = -.20, $p < .05$.

It comes as little surprise that responses to the two surveys items that queried pilots about the impact of advanced cockpit systems on the accident rate, and their preferences for using them instead of conventional systems were significantly correlated with one another: $r(132)$ = 0.4, $p < .01$.

Summary. Pilots expressed a clear preference to spend almost all of their time in aircraft that are equipped with advanced cockpit systems. Pilots exhibited a strong preference for using GPS to navigate and an autopilot to control the aircraft, and these two systems were pilots' first choices for inclusion in their own cockpit. Pilots seem to hold the general belief that autopilots are an invaluable tool during times of high workload, but oddly do not prefer to use them in specific high-workload situations. Correlations among pilot preferences and attitudes toward the likely impact of advanced cockpit systems on safety suggest that pilots' attitudes and preferences are consistent.

Summary and Conclusions

The results indicate that general aviation pilots hold generally positive attitudes toward advanced cockpit systems. Pilots indicated that they preferred to fly a median of 90% of their time in advanced cockpits. Correlations between this stated preference and survey items that probed specific issues associated with advanced cockpit systems suggest that pilots' general preferences are supported by specific reasons. Pilots seem to prefer to fly advanced cockpit aircraft because they believe advanced cockpit systems offer specific benefits such as lower workload and increased awareness. Pilots believe that the advanced cockpit will help decrease accidents, and that terrain, weather, and traffic systems are three systems that will help reduce the number of accidents. Responses to survey items about learning to use advanced cockpit systems suggest that pilots acknowledge learning challenges posed by these systems and are open to the idea of investing additional training time or requirements to master them.

Pilots acknowledge potential pitfalls of advanced cockpit systems but also seem to believe that they are less susceptible to these pitfalls than are other pilots. This result suggests that an important part of training pilots to use advanced cockpit systems should focus on helping pilots to understand potential traps associated with advanced cockpit systems and to help pilots more accurately assess their own vulnerabilities.

Pilot experience modestly accounted for some of the variation in pilots' responses, indicating that pilots' attitudes likely do change as they accumulate experience with advanced cockpit systems. Pilots with greater experience with advanced cockpit systems feel less like "button pushers," believe that they understand the equipment better, and are less dissatisfied with manufacturers' technical manuals.

The comparison of general aviation pilots with the airline pilots tested in previous surveys showed a striking similarity in attitudes between these two groups. This result suggests a strong similarity between the issues that confront the pilots who operate these two types of aircraft. This is an important finding since there is already a trove of literature on pilot interaction with advanced cockpit systems found in the jet transport cockpit (Wiener and Nagel, 1988; Billings, 1997; Parasuraman and Mouloua (1996).

Comparing Beliefs and Attitudes with Performance and Behavior

The payoff for measuring pilots' attitudes toward advanced cockpit systems comes when we are able to identify areas in which attitudes are misaligned with reality and lead to unexpected or undesirable behaviors and outcomes. In this study, we have only identified pilot attitudes. The responses to the survey items point out the need for studies that directly compare the attitudes measured here with pilot behavior and performance in the advanced cockpit.

Pilots' contradictory attitudes toward the usefulness of autopilots are a first topic to explore. An empirical study of pilots operating through all phases of flight would help understand when and why pilots would invest the time required to configure and engage the autopilot and when they would opt for manual control of the aircraft.

The general aviation setting provides a unique opportunity to test pilots' attitudes toward skill atrophy because of extended use of advanced cockpit systems. Because of the highly scripted nature of airline flight operations, previous studies of airline pilots have been limited to surveys of airline pilots' opinions about skill atrophy. A future study might attempt to document the extent to which general aviation pilots use automated systems, query pilots about their confidence in their own instrument flying skills, and then put those skills to a practical test.

Another future study might document and compare the errors made by skilled general aviation pilots while flying in advanced vs. conventional cockpits.

Yet another study might focus on pilots' belief that traffic, weather, and terrain systems are likely to reduce the accident rate. If pilots feel that these systems provide an extra margin of safety, are pilots likely to accept additional risk (Wilde, 2001)? Numerous other studies have demonstrated a close link between the likelihood of accident involvement and risk perception (O'Hare, 1990) and other attitudinal factors (Hunter, 2006).

Studying the relationship between pilot experience with advanced cockpit systems and their attitudes toward them was naturally limited in that these systems are relatively new to the general aviation fleet. It would be interesting to measure how general aviation pilots' attitudes change as they acquire significant amounts of experience with these new systems.

References

AOPA (2007). *FAA certificated pilots 1929-2006*. Aircraft Owners and Pilots Association. Retrieved from www.aopa.org/special/newsroom/stats/pilots.html.

BASI (1998). *Advanced technology aircraft safety survey report.* Bureau of Air Safety Investigation. Australian Capital Territory.

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

Bodner, D., Cochran, C. D., and Blum, T. L. (2000). Unique invulnerability measurement in skydivers: Scale validation. *Psi Chi Journal of Undergraduate Research, 5*(3), 104-108.

Burger, J. M., and Burns, L. (1988). The illusion of unique invulnerability and the use of effective contraception. *Personality and Social Psychology Bulletin, 14*(2), 264-270.

Casner, S. M. (2005). The effect of GPS and moving map displays on navigational awareness while flying under VFR. *International Journal of Applied Aviation Studies, 5*(1), 153-165.

Casner, S. M. (2004). Flying IFR with GPS: How much practice is needed? *International Journal of Applied Aviation Studies, 4*(2), 81-97.

Greening, L., and Dollinger, S. J. (2005). Illusions (and shattered illusions) of invulnerability: Adolescents in natural disaster. *Journal of Traumatic Stress, 5*(1), 63-75.

Hunter, D. R. (2006). Risk perception among general aviation pilots. *International Journal of Aviation Psychology, 16*(2), 135-144.

Hutchins, E., Holder, B., and Hayward, M. (1999). *Pilot attitudes toward automation.* University of California, San Diego.

Kirlik, A. (1993). Modeling strategic behavior in human-automation interaction: Why an "aid" can (and should) go unused. *Human Factors, 35*(2), 221-242.

Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134.

O'Hare, D. (1990). Pilots' perceptions of risks and hazards in general aviation. *Aviation, Space, and Environmental Medicine, 61*, 599-603.

Parasuraman, R. and Mouloua, M. (1996). *Automation and human performance.* Mahwah, NJ, Lawrence Erlbaum Associates.

Sulheim, S., Ekeland, A., Bahr, R. (2007). Self-estimation of ability among skiers and snowboarders in alpine skiing resorts. *Knee Surgery, Sports Traumatology, Arthroscopy, 15*(5), 665-670.

Wiener, E. (1985). *Human factors of cockpit automation: A field study of flight crew transition.* NASA Contractor Report 177333.

Wiener, E. and Nagel, D. C. (1988). *Human factors in aviation.* Academic Press.

Wilde, G. J. S. (2001). *Target risk 2: A new psychology of safety and health.* PDE Publications.

Perturbing the System: Emergency and Off-Nominal Situations under NextGen

Barbara K. Burian, Ph.D.

NASA Ames Research Center
Mail Stop 262-4
Moffett Field, California 94035
706-348-1617
650-604-5344 (messages only)
Barbara K. Burian@nasa.gov

Abstract

Emergencies and off-nominal situations will challenge the safe and efficient operation of NextGen. This paper focuses on three issues: 1) defining the terms "emergency" and "off-nominal," 2) identifying the full-range of emergency and off-nominal situations and their effects on the functioning of human operators, technologies, procedures, and NextGen operations, and 3) determining performance capabilities, limitations and external pressures affecting human response to these situations.

In the national airspace system (NAS) envisioned under the Next Generation Air Transportation System (NextGen)—one that will be gradually introduced over the next several years and fully in place by 2025—performance-based services will be provided to aircraft flying 4D trajectories from take-off to landing.[1]   Tightly spaced arrivals and departures, achieved through the use of new procedures and advanced technologies, will facilitate super-density operations at and around airports, thereby increasing capacity. Access to and use of enroute flow corridors by sufficiently equipped aircraft, as well as dynamically defined airspace, will also facilitate the orderly movement of the greatly increased number of aircraft expected to be flying in the future NAS. This tightly-coupled system will work best when unperturbed by poor weather conditions, equipment failures, human errors, or emergencies. But, of course, poor weather conditions, equipment failures, human errors, and emergencies occur in today's NAS and will continue to occur after NextGen has been implemented.

---

[1] "4D trajectory" refers to the precise navigation of an aircraft within 4 dimensions: along the x, y, and z axes at specific points in time.

The Joint Program Development Office (JPDO) personnel and others who have been developing the NextGen Concept of Operation (JPDO, 2007) understand these facts. For example, the NextGen operating concepts and proposed research emphasize inclement weather, the nemesis of smooth and timely aviation operations, so that in the future, limitations due to poor weather conditions can be adapted to or overcome and will have less impact on operations.

In contrast, the NextGen Concept of Operations currently gives less treatment to emergency and off-nominal situations. (JPDO, 2007). Under NextGen, what is an "emergency"? What is an "off-nominal"? Are emergencies considered a subset of off-nominal situations or are the two categories separate and distinct? Will (or should) response to off-nominals and emergencies differ because of how they are defined? What factors need to be considered when developing procedures for responding to these situations? These are just a few questions the aviation industry, operators, regulators, and researchers need to answer as they develop the technologies, concepts, and procedures necessary for the implementation of NextGen.

This paper explores: 1) how emergencies and off-nominal situations are addressed within the NextGen Concept of Operations, 2) the need to consider the full range of emergency and off-nominal conditions that may occur and their potential effects and ripple effects on NextGen operations, 3) human response to high stress, workload, and emergencies, and 4) related research issues. This exploration must begin with a definition of terms.

Emergency, Abnormal, and Off-nominal Situations

Clear definitions and a common understanding of what constitutes emergency and off-nominal situations are essential if NextGen technologies and procedures are to deal with them adequately. The terms "off-nominal" and "emergency" are used throughout the NextGen Concept of Operations (JPDO, 2007), but they are neither defined nor appear in the document glossary. Because both terms are often used together (i.e., "emergency *and* off-nominal situations") it can be presumed that the authors of the NextGen Concept of Operations see these terms representing two distinct and separate categories of situations rather than one (emergencies) being a subset of the other (off-nominal). So, beginning with that dichotomy, we must determine what constitutes each category and how current situations, which are classified somewhat differently, would be categorized under NextGen.

The usage of "*off-nominal*" in the document could be interpreted to include any situation that is out of the idealized norm of NextGen operations, with the exception of emergencies. Thus, an off-nominal situation might include such things as the performance of a missed approach or go-around by an aircraft, dispatching an aircraft with a piece of non-critical equipment (Minimum Equipment List [MEL'ed] [2]), delays in ground operations due to bad weather, or a backlog in approaches to

---

[2] "MEL'ed" equipment aboard an aircraft is not working as intended and is typically disabled, if necessary, until a certain time when repairs are required to be made. The Minimum Equipment List (MEL) states which equipment may be disabled in this way without affecting the dispatch status of the aircraft. Equipment that has been "MEL'ed" has been determined to not adversely affect the airworthiness of an aircraft or safety of flight.

a particular runway (something that is not supposed to occur if NextGen is operating as envisioned). These off-nominal situations under NextGen are also considered off-nominal (but common) in today's classification scheme.

Situations aboard aircraft, which have historically been classified as *abnormal*[3] by the industry, would generally be categorized as "off-nominal" under NextGen. These situations are important for the crew to detect and address, but are not time critical and do not pose a significant threat to life or the airworthiness of the aircraft. They typically involve the failure of technology or equipment that does not reach the level of an emergency, such as the overheating of an air conditioning pack. Thus, the NextGen off-nominal category includes not only the completion of atypical procedures (a missed approach/go-around, dispatching with MEL'ed), or the breakdown of operations (delays and backlogs), but also the malfunction of technologies and equipment.

Consequently, NextGen off-nominal situations will differ along two important dimensions: "character" and regularity of occurrence. In terms of character, some off-nominals will require significant attention and intervention and may have important implications for operations, such as the malfunction of some types of equipment or technologies. However, other off-nominal conditions may be relatively benign and may have few requirements for attention and intervention, such as dispatching an aircraft with some equipment MEL'ed. Character is not a static trait and may vary as a function of other conditions. For example, dispatching an aircraft with an inoperative thrust reverser (i.e., thrust reverser is MEL'ed) will likely require more crew attention and planning when attempting a landing on a wet runway in a crosswind. Conversely, landing that same aircraft with a MEL'ed thrust reverser on a dry runway with winds straight down the runway may not be an issue.

With regard to the second dimension, some current off-nominal situations occur with a fair degree of regularity and may continue to be relatively common under NextGen, such as the previous example (see the "Common Off-Nominal but not Abnormal" portion of Figure 1). Other off-nominal situations should be far rarer under NextGen, such as delays in terminal operations, and should comprise the bulk of off-nominal NextGen situations.

To summarize, under NextGen the label "off-nominal" could be applied to many conditions currently identified as abnormal, in addition to all types of situations that fall outside of the conception of nominal or idealized NextGen operations or procedures, but do not meet the criteria for an emergency (see Figure 1).

---

[3] Historically, aircraft situations that deviated from normal operation were termed either "abnormal" situations or "emergencies." At least one US air carrier uses the term "irregular" rather than "abnormal," and the Federal Aviation Administration sometimes uses the term "non-normal" in place of "abnormal." Several years ago, Boeing adopted the term "non-normal" but uses this term to apply to what had previously been both abnormal *and* emergency conditions. Therefore, because the term "non-normal" has two different meanings within the industry, it is not used in this paper to avoid confusion (except in the section where Boeing's adoption of the term is discussed).
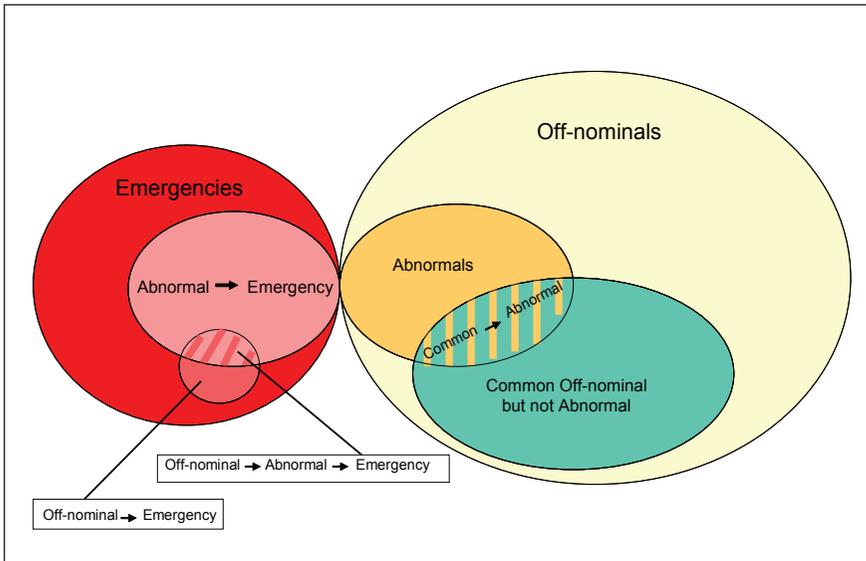
Figure 1. Emergency, Abnormal, and Off-nominal Situations under NextGen
(Common off-nominal conditions are those that occur quite frequently but are not,considered abnor-
mal, such as dispatching an aircraft with equipment MEL'ed. Off-nominal conditions can, sometimes
contribute to another situation becoming one that is abnormal or an emergency.)

Emergencies are generally understood to be situations with significant poten-
tial for injury, loss of life, and/or severe damage to aircraft, equipment, or infrastruc-
ture (Burian, Barshi, & Dismukes, 2005; Stokes & Kite, 1994). This conception is
not likely to change under NextGen. Often emergencies are time-critical events
and response must be immediate to avert catastrophe.

In the off-nominal discussion above, abnormal conditions were located within
the realm of NextGen off-nominal situations. Unfortunately, the conception of how
to classify today's abnormal situations under NextGen is not so cut and dried.
Some abnormal conditions may actually become emergencies if left unattended.
Hence, in Figure 1, circles representing abnormal situations can be found within
both the off-nominal *and* emergency categories, although only abnormal situations
which evolve into emergencies belong in the emergency realm. Thus, a situation
may begin as an off-nominal under NextGen and jump the boundary to become an
emergency

Further complicating the classification of situations as either off-nominal or
emergency is the same underlying condition might be considered either an
abnormal (hence, an off-nominal condition) or an emergency depending upon the
context and circumstances in which it occurs. For example, a fuel leak discovered
while on approach to landing would most likely *not* be considered an emergency
(but rather an abnormal, i.e., off-nominal situation) whereas the same fuel leak
discovered at cruise altitude during flight over an ocean most likely would be con-
sidered an emergency. Thus, under NextGen the contextual factors surrounding a
situation must be considered if for some reason it becomes important to categorize
that situation as either off-nominal or an emergency.

The final complication in this discussion is that just as abnormal situations may become emergencies, some off-nominal conditions may contribute to other off-nominal situations becoming abnormal or emergencies (using today's classification scheme). For example, an aircraft might be dispatched with the interphone between the cabin and cockpit inoperative. Because different patterns of chimes are often used for routine communication from the cockpit to the cabin, dispatching the aircraft this way might be considered simply off-nominal. However, should the cabin crew hear an unusual noise or smell smoke, the lack of an easy way to communicate these observations to the flight deck crew turns these situations into ones that are abnormal or emergencies.

As can be seen in Figure 1, the realm of possible off-nominal conditions should be much larger than the realm of emergency conditions. This is in part because emergencies occur with far less frequency than off-nominal situations. Additionally, the off-nominal category includes not only most of the possible malfunctions that may occur in equipment and advanced technologies but also most conceivable breakdowns in the operations upon which NextGen system is so dependent.

Although clarity in nomenclature regarding emergencies and off-nominals is required, making a distinction between emergencies and off-nominals may *not* be necessary or even desired when the events are unfolding. Several years ago, engineers at Boeing decided to use the term "non-normal" to refer to all situations historically characterized as either "emergencies" or "abnormals" (D. Boorman, personal communication, January 19, 2001). They thought that the same situation could be either an emergency or an abnormal depending upon other contextual factors, as described earlier. Additionally, collapsing these situations and the procedures for responding to them into one category—non-normal—eliminated the need for pilots to recall if a situation was categorized as an abnormal or an emergency by procedure developers when attempting to locate the procedure for responding to it. Similar considerations will need to be undertaken by NextGen developers—what is the best way of thinking about NextGen emergencies and off-nominals to ensure ease and speed of procedure location and response?

The NextGen Concept of Operations and Emergency and Off-nominal Situations

The current version of the NextGen Concept of Operations (JPDO, 2007) does not go into detail beyond the two high-level categories of emergencies and off-nominal situations. This document specifies seven main areas in which NextGen technologies and operations must be responsive to emergencies and off-nominal situations, and these seven areas cut across the eight key capabilities of NextGen:
1. Information Availability and Sharing:
   a. Network-Enabled Operations
   b. Surveillance Information Services
   c. Flight Data Management System
2. Net-Centric Infrastructure: Voice by Exception
3. Flow Corridors

4. Flow Contingency Management
5. Functional Task Allocation
6. The Role and Training of Air Navigation Service Providers
7. Response to Security Threats and Attacks

The range of emergency and off-nominal conditions addressed within each of these seven areas, described below, includes one or more of the following:

- emergency or off-nominal conditions aboard aircraft in the air or on the ground
- emergency or off-nominal conditions regarding the functioning of the NextGen architecture (which includes the failure, malfunction, degradation, or errors of technologies and equipment, personnel, and operating procedures)
- emergency conditions related to a security breach or threat

Information Availability and Sharing.
It is envisioned that airport operators, air navigation services, flight operators, and other stakeholders will be able to identify, monitor, and respond to emergency and off-nominal situations better through *Network-Enabled Operations*, *Surveillance Information Services*, and alerts generated by *Flight Data Management Services*. Communication regarding the situation, response, and needs will be "automatically routed to the appropriate user groups" and made available to "adjacent jurisdictions and relevant regional and/or national entities" enabling them to "provide the most efficient support possible" (JPDO, 2007, p. 3-11).

Under NextGen, a level of required communications performance (RCP) for all types of data and voice communication will be mandated and contingency operations that will ensure continuity of operations for the air traffic system will be established, which will include standardized responses to emergencies and all hazards. (JDPO, 2007).

Net-Centric Infrastructure: Voice by Exception.
Although datalink will be the preferred method of communication between the flight deck and ground under NextGen, "voice will be used in cases of emergency such as [one affecting the] safety of flight (e.g., a situation where a conflict or midair collision is imminent and voice [communication] will preclude an incident)" (JDPO, 2007; p. 4-2).

Flow Corridors
The NextGen Concept of Operations (JDPO, 2007) specifies that procedures will exist that allow aircraft with declared emergencies to safely exit the flow corridor, which is reserved for aircraft flying 4-D trajectory assignments. The unstated assumption is that aircraft with emergencies may no longer meet the equipment or performance requirements allowing use of the flow corridor and/or may need to alter their flight plans (i.e., divert) or require additional services from air navigation service providers.

Flow Contingency Management
Flow Contingency Management "is the process that identifies and resolves congestion or complexity resulting from blocked or constrained airspace or other

off-nominal conditions" (JPDO, 2007; p. 2-16). The process aims to affect as few flights as possible to deal with a constraint, and a variety of strategies may be used including "establishing multiple trajectories and/or flow corridors to reduce complexity, restructuring the airspace to provide more system capacity, or allocating time-of-arrival and -departure slots to runways or airspace" (p. 2-16).

Functional Task Allocation

The NextGen ATM system will rely heavily on the use of technologies and automation both in the air and on the ground, and the NextGen Concept of Operations requires appropriate allocation of roles and tasks to humans and automation (JPDO, 2007). Part of the decision about which tasks to fully automate will depend on ensuring "that service providers and flight operators perform well and can respond to off-nominal and emergency events when required" (JPDO, 2007, p. 2-11). Under NextGen, fail-safe modes and back-up functions that do not fully depend upon human intervention are supposed to be in place to provide multiple "layers of protection to allow for graceful degradation of services in the event of automation failures" (p. 2-11).

Additionally, an area identified for research in the Concept of Operations document is the exploration of "which NextGen systems should be fully automated without relying on human intervention for off-nominal situations" (JPDO, 2007, p. C-4). Thus, the use of automation should not impede human intervention in emergency and off-nominal situations. Automation is identified as both a potential source of off-nominal situations and as a potential solver of off-nominal conditions.

The Role and Training of Air Navigation Service Providers

It is expected that a significant portion of the training of air navigation service providers under NextGen will be devoted to the management of emergency and off-nominal operations using varying levels of automated support, including the possibility of *no* automation support (JPDO, 2007).

Response to Security Threats and Attacks

The NextGen Concept of Operations includes consideration of some requirements for adequate identification and detection of security threats and attacks, including chemical, biological, radiological, nuclear, high-yield explosives, and cyber attacks against the NextGen computer network and infrastructure. The document specifies that the emergency response to these types of situations must be "appropriately rehearsed to ensure that the responders are fully prepared and informed for any contingency" (JPDO, 2007, p. 6-7). The secure exchange of information is specified as a need during these events (JPDO, 2007) but is also required during normal operations.

<div align="center">What is Missing?</div>

As discussed, the NextGen Concept of Operations (JPDO, 2007) specifies several objectives regarding how to manage emergency and off-nominal situations and their effects. However, several of these objectives have not received

sufficient consideration. Other issues were not addressed at all. The remainder of this paper explores two of the areas that require further investigation: 1) the full range and effects of emergency and off-nominal situations, and 2) human responses to stress, workload, and emergencies.

Full Range and Impact of Emergency and Off-nominal Situations

One of the most significant issues requiring further attention is the need to explore, in-depth, the full range of potential emergency and off-nominal situations and the many ways in which these situations may affect NextGen operations. An emergency or off-nominal condition can originate in a human, a technology or equipment (software or hardware), or in a poorly designed operating procedure. The effects of these conditions may be constrained to one area or may be propa-gated among several. A wide array of situations will occur, some more frequently than others will. For example, consider just a few situations that might occur involving a single aircraft: a general aviation aircraft is lost, an aircraft is caught unexpectedly in a microburst while on approach, an aircraft experiences a gear malfunction on take-off and must burn off fuel for several hours before returning to the departure airport to land (e.g., JetBlue 292), an aircraft experiences a time critical emergency while at cruise altitude (e.g., in-flight fire) and must divert to an unfamiliar airport (e.g., FedEx 1406), an aircraft lands with a passenger on-board who is suspected to have a highly contagious disease (e.g., Cathay Pacific 451), an aircraft crash lands at an airport, thereby closing one or more runways (e.g., Continental 1943). Each of these situations varies in the degree that outside inter-vention may be necessary and the types of effects each have on the functioning of the individual aircraft, as well as the functioning of the ATM system.

It is appropriate to develop procedures so aircraft experiencing an emergency or off-nominal condition can *depart* flow corridors. However, what about inade-quately-equipped aircraft that need to *enter into* or disrupt the flow of other aircraft in a corridor, due to a time-critical emergency? On average, three aircraft divert every day due to smoke in the cockpit (Shaw, 2000); diversions will continue to be a reality under NextGen and procedures must be in place to accommodate them. It is also common for an aircraft to need a "piece of sky" to circle in a holding pat-tern to allow time to figure out an on-board anomaly, complete checklists, or burn off fuel; this also is unlikely to change under NextGen. The concept of a "graceful degradation of services" has often been applied broadly to more than automation failures by those developing NextGen concepts, technologies, and procedures. Unfortunately, there is nothing graceful in the "degradation of services" at an air-port when an accident suddenly closes down runways, or if, aircraft at multiple airports need to be quarantined. The NextGen Concept of Operations must foresee all these and other types of emergency and off-nominal conditions, and develop procedures to manage them.

Not all off-nominals will originate with aircraft. For example, an air navigation service provider who is severely fatigued will be vulnerable to confusion and loss of situation awareness. Initially, the origin and effect of the off-nominal condition, fatigue, are confined to the human operator. However, because of this fatigue, the air navigation service provider could be delayed in responding to the needs of a particular flight or may not attend appropriately to a degradation in automation functioning, thereby adversely affecting the spacing of multiple aircraft departing a

busy airport. Thus, the operation of single aircraft or the operation of the system as a whole is affected. Similarly, the malfunction of NextGen automated technologies will affect not only the functioning of ATM systems/operations, but also the responses required by individual aircraft and the human operators who must react and adjust.

It is essential to examine not only *what* is affected, but *how* these effects and potential ripple effects may be manifested. For example, imagine that Flight 123 is at cruise altitude in a flow corridor on a flight from Airport A to Airport B when the pilots notice a slight decrease in cabin altitude, indicating an increase in cabin pressure. In keeping with the earlier definition of terms, this is an off-nominal situation and, at this point, involves only the single aircraft (the locus of both the occurrence of the off-nominal condition and its effect). The pilots observe that the cabin altitude is very slowly continuing to decrease and then a BLEED LEAK BODY message appears on the crew alerting display. About this same time, a potential conflict with another aircraft is alerted and a datalink message from their airline dispatcher is received.

The crew must manage multiple tasks concurrently, their workload and stress have increased – potential conflicts with other aircraft and problems with pressurization and bleed leaks are both attention grabbers. Thus, the crew's performance is affected as they fixate on one or two of these tasks to the exclusion of others, causing a ripple effect.

The monitoring pilot accesses the electronic checklist for the bleed leak condition and notes that the bleed air system should automatically isolate the leak. Both pilots continue to monitor the cabin altitude as the Engine Bleed Switch OFF lights and the Isolation Switch CLOSED lights on the overhead panel automatically illuminate and extinguish. The flight attendants call to let them know that passengers are complaining of being too hot. While the monitoring pilot explains the situation to the flight attendants, the flying pilot attends to the potential conflict with another aircraft and the message from dispatch remains unread. The flight crew discusses other possible actions, such as starting the auxiliary power unit (APU) to supplement hydraulic power that may be lost, but using the APU necessitates descent to a lower altitude. At this point, the crew is reluctant to declare an emergency – it is possible that the aircraft automation will stabilize the situation, and they do not want to drop to a lower altitude (out of the flow corridor) or divert unnecessarily. However, if their situation is not rapidly stabilized, they run the risk of air conditioning smoke or fire as hot bleed air continues to be vented into the cabin. This off-nominal condition will have transitioned to a time-critical emergency, one likely demanding an immediate diversion.

Are procedures in place for the crew to use in the event that automated technology (isolation of the bleed leak) fails? What might be the consequences of the crew's fixation and subsequent shedding of other tasks? How and when should air service navigation providers be informed and get involved? Will an air service navigation provider be available, if, for example, inclement weather requires

human involvement in adjusting flow corridors and the flights of other aircraft? How will diversions be handled and how will possible disruptions to the overall system be managed? If it is best that an emergency aircraft land at its intended airport of arrival, should it remain in the flow corridor, and if so, what should be done with other aircraft that precede it in the corridor?

What should become apparent through reading this brief scenario and the associated questions is that a great deal of thinking is required about the effects that off-nominal and emergency conditions have upon human performance, the use of and dependence upon automation, and the functioning of the overall system. Assumptions about available resources and the functioning of procedures must be made explicit and tested. Researchers and developers must iteratively ask, "Well, what if…?" to account for as many exigencies as possible. The capabilities and limitations of humans and automation to respond to emergencies and off-nominals must be considered (more on this in the next section) along with all of the tasks that must be completed, the demands that completion of these tasks place on humans, technology, and operational procedures, and how these demands should be accommodated and facilitated under NextGen.

Human Responses to Stress, Workload, and Emergencies
Much has been learned through the study of aviation incidents, accidents, and human response to emergencies and other unexpected events, high workload and stress (e.g., Burian, 2005, 2006; Burian & Barshi, 2003; Kochan, Breiter, & Jentsch, 2007; Staal, 2004, Stokes & Kite, 1994; Tremaud, 2002; Woods & Patterson, 2001). This wealth of information should inform the development and design of NextGen technologies and procedures to ensure adequate and appropriate response to emergency and off-nominal situations. For example, in a recent analysis of 19 major accidents in the U.S. where pilot error was determined by the National Transportation Safety Board to be a contributory or causal factor (Dismukes, Berman, & Loukopoulos, 2007), six clusters of errors were identified, four of which are relevant to emergency and off-nominal situations:

1. Inadequate execution of emergency procedures under challenging conditions
2. Inadequate response to rare situations
3. Judgment in ambiguous situations that hindsight proves wrong
4. Deviation from explicit guidance or standard operating procedures

Several cross-cutting issues that contributed to these accidents were noted, including the need to manage multiple tasks concurrently and deal with greatly increased workloads, emergency and off-nominal situations that required very rapid responses (this was the case for two-thirds of the 19 accidents analyzed), equipment failures and design flaws (also occurring in two-thirds of the accidents), and information cues that were absent or displayed in a misleading fashion to the flight crews (Dismukes, Berman, & Loukopoulos, 2007). NextGen technology and procedure developers will find the detailed analyses and countermeasures provided in Dismukes, et al. and similar resources to be quite informative. Anticipated advances in technologies and automation will not obviate the lessons to be learned in these studies and their relevance to dealing with emergency and off-nominal conditions under NextGen.

Although NextGen automation will manage many common off-nominal conditions, or those that can be anticipated in advance by design engineers, situations that are uncommon, unanticipated, or ambiguous will still likely require human intervention. In fact, the ability to formulate and carry out a plan in the face of ambiguous or incomplete information is one of the areas in which humans far outshine the performance or abilities of automation. Even so, emergency and off-nominal situations, unless frequently encountered or trained, tend to require effortful cognitive processing (Hendy, Farrell, & East, 2001), which increases the time needed for human response. This will be true for pilots responding to an emergency and for air service navigation providers who assist or respond to an off-nominal situation affecting the larger system.

Studies of human performance have found that under stress (common during emergencies and off-nominal situations) attention tends to narrow so that only the few cues perceived to be the most threatening or salient are attended to (e.g., Bundesen, 1990; Wickens, 1984). This tunneling of attention means that other cues relevant to the situation may go unnoticed. Working memory is also negatively affected by stress (Baddeley, 1986), so humans experiencing high degrees of stress will likely have more difficulty pulling together disparate information from multiple sources and making sense of it – something often required in ambiguous or confusing emergency situations (Burian, Barshi, & Dismukes, 2005; Hendy, Farrell, & East, 2001).

So what does this have to do with the NextGen Concept of Operations? Earlier in this paper a case was made for examining the full range of emergency and off-nominal situations and their diverse effects. Degradation of human performance capabilities under stress is one of those effects. "Too often, procedures to be followed in the event of emergencies are written as if the humans following them will be operating at 100% peak efficiency" (Capt. David Keeling, personal communication, August 10, 2001). A few concerns regarding human performance during off-nominal situations do appear in the NextGen Concept of Operations. For example, in a discussion about function allocation between automation and humans, the need to ensure human operators' situation awareness in the event of automation failure is recognized. However, a much broader range of human capabilities and limitations when responding to emergency and off-nominal situations must be considered. Such consideration will allow appropriate countermeasures and mitigations to be developed, validated, and implemented in NextGen to accommodate and lessen the impact of these human limitations, and provide a foundation for developing strategies to take full advantage of unique human capabilities to set priorities and solve problems..

So far, this discussion has focused largely on the cognitive performance of individuals when faced with stress. However, it should be remembered that the larger socio-cultural environment in which NextGen operations will take place, will also affect how emergency and off-nominal conditions are handled. For example, pilots are often reluctant to declare an emergency in today's aviation environment. The reasons for this reluctance include fear of drawing the attention of

regulators, not wanting to have to complete lengthy reports, and possibly even ego--i.e., I don't need any help; I can handle this! (Barshi & Kowalski, 2003). Similarly, company and economic pressures affect the decisions pilots make about how to respond to some off-nominal conditions, such as bypassing alternate airports in order to divert to one where maintenance services are available (Burian & Barshi, 2003).

We cannot know at this point what socio-cultural forces will come into play by the time NextGen ATM operations are implemented, and we cannot be certain how these forces will influence human responses to emergencies and off-nominal situations, subtly or not so subtly. Nonetheless, it is incumbent upon those developing technologies and procedures for emergency and off-nominal situations under NextGen to keep in mind that such pressures will exist and, to the extent possible, anticipate the consequences of these pressures and develop appropriate countermeasures and mitigations.

Conclusion

Currently different populations within the aviation industry use different terms to describe the same situation (e.g., abnormal, irregular), the same term to describe different kinds of situations (e.g., non-normal = abnormal; non-normal = emergency + abnormal), and terms pertaining to emergency or off-nominal situations that may not be commonly used or understood by others in the industry (e.g., urgency situation). The terms "emergency" and "off-nominal" used in the NextGen Concept of Operations (JPDO) are neither defined nor described. To ensure that responses to NextGen emergency and off-nominal situations are adequate and effective, researchers and developers need to adopt a common nomenclature and must be clear about what is meant when that nomenclature is used.

What are the ranges of off-nominal and emergency situations that must be accommodated and what kinds of problems and ripple effects can be anticipated? What are the capabilities and limitations that human operators, technologies, and procedures bring to these situations and how might socio-cultural pressures play a role? This paper has only introduced the need to address these questions. Multiple techniques and approaches are available to assist researchers and developers in anticipating the effects of off-nominals and emergencies on NextGen operations such as cognitive walk-throughs (Wharton, Rieman, Lewis, & Polson, 1994), modeling (Foyle & Hooey, 2008), formal methods (Clarke & Wing, 1996), and failure mode and effects analyses (FMEA; Stamatis, 2003). Other techniques that take a solely systemic view of hazard analysis, such as the Systems-Theoretic Accident Modeling and Process (STAMP) approach developed by Leveson (2003) and her colleagues, will also be needed for anticipating and assessing the effects of emergencies and off-nominals on the entire NextGen ATM system.

Through research, we have some knowledge about the effects of stress and workload on human performance (e.g., tunneling, reduced working memory capacity); however, much of this knowledge has been derived from carefully controlled laboratory studies. These findings need to be extended and validated in the real world-operating environment. This knowledge must then be applied by NextGen researchers and developers when designing procedures and technolo-

gies to accommodate human limitations and take maximal advantage of human capabilities when responding to the stress and high workload that will naturally accompany emergency and off-nominal situations under NextGen.

In the NextGen concept of Operations (JPDO, 2007) several important objectives have been stated with regard to the functioning of the air traffic management system of the future; it is now necessary that all who are working on developing NextGen concepts, technologies, and procedures go much further in their consideration of emergency and off-nominal situations and their effects.

Acknowledgements

References

Baddeley, A. D. (1986). *Working memory.* Oxford, England: Oxford University Press.

Barshi, I. & Kowalski, T. (2003). Declaring an emergency: fact and fiction. In the *Proceedings of the 12th International Symposium on Aviation Psychology,* Dayton, OH, April 14-17.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*(4)*,* 523-547.

Burian, B. K. (2005). Do you smell smoke? Issues in the design and content of checklists for smoke, fire, and fumes. In *Proceeding of the International Society of Air Safety Investigators 2005 Conference*. Fort Worth, TX: ISASI.

Burian, B. K. (2006). Aeronautical emergency and abnormal checklists: Expectations and realities. In the *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, San Francisco.

Burian, B. K., & Barshi, I. (2003). Emergency and abnormal situations: A review of ASRS reports. In the *Proceedings of the 12th International Symposium on Aviation Psychology*. Dayton, Ohio: Wright State University Press.

Burian, B. K., Barshi, I, & Dismukes, K. (2005). *The challenge of aviation emergency and abnormal situations.* NASA TM-2005-213462. Moffett Field, CA: NASA Ames Research Center.

Clarke, E. M., & Wing, J. M. (1996). Formal methods: State of the art and future directions. *ACM Computing Surveys (CSUR), 28*(4), 626-643.

Dismukes, R. K., Berman, B. A., & Loukopoulos, L. D. (2007). *The limits of expertise: Rethinking pilot error and the causes of airline accidents.* Aldershot, Hampshire, England: Ashgate.

Foyle, D. C., & Hooey, B. L. (Eds.). (2008). *Human performance modeling in aviation.* Boca Raton, FL: CRC Press.

Hendy, K. C., Farrell, P. S. E., & East, K. P. (2001). An information-processing model of operator stress and performance. In P.A. Hancock and P.A. Desmond, (Eds.), *Stress, workload, and fatigue* (pp. 290-302). Mahwah, NJ: Lawrence Erlbaum Associates.

Joint Program and Development Office (JPDO). (2007). Concept of operations for the next generation air transportation system, Version 2.0. Retrieved September 12, 2007, from http://jpdo.gov/library/ NextGen_v2.0.pdf.

Kochan, J. A., Breiter, E. G., & Jentsch, F. (2007). Surprise and unexpectedness in flying: Database reviews and analyses. In *Proceedings of the 51st Annual Meeting of the Human Factors and Ergonomics Society*, 335-339.

Leveson, N. G., (2003). A new approach to hazard analysis for complex systems. *In Proceedings of the International Conference of the System Safety Society.*

Shaw, J. (2000). *A review of smoke and potential in-flight fire events in 1999.* Washington, DC: Society of Automotive Engineers.

Staal, M. (2004). *Stress, cognition, and human performance: A literature review and conceptual framework.* NASA TM-2004-212824. Moffett Field, CA: NASA Ames Research Center.

Stamatis, D. H. (2003). *Failure mode and effect analysis: FMEA from theory to execution* (2nd ed., Rev.). Milwaukee, WI: ASQ Quality Press.

Stokes, A. & Kite, K. (1994). *Flight stress: Stress, fatigue, and performance in aviation.* Burlington, VT: Ashgate.

Tremaud, M. (October 2002). *Operational and human factors involved in situations beyond the scope of published procedures*. Presentation given at the Airbus 16th Human Factors Symposium, Singapore.

Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994). The cognitive walkthrough method: a practitioner's guide. In J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods.* (pp. 105-140). New York: John Wiley & Sons, Inc.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention* (pp. 63-101). New York: Academic Press.

Woods, D. D. & Patterson, E. S. (2001). How unexpected events produce and escalation of cognitive and coordinative demands. In P.A. Hancock and P.A. Desmond, (Eds.), *Stress, workload, and fatigue* (pp. 290-302). Mahwah, NJ: Lawrence Erlbaum Associates.

Performance Assessment of Strategic Team Training in Simulated Air Traffic Control.

Christopher P. Barlett, Christopher L. Vowels
John D. Raacke, and James Shanteau
Kansas State University

Department of Psychology
1100 Mid-Campus Drive
Manhattan, KS, 66506
email: cvowels@ksu.edu
phone: 785-410-7429

Abstract

The current study was conducted in order to determine if strategic team training at an air traffic control task would increase long-term performance compared to teams that received factual training. Half of the teams engaged in strategic training prior to longitudinal assessment of performance in CTEAM (a simulated air traffic control task), whereas the other teams received factual training. Those teams that received strategic training consistently had lower delay times than the factually trained teams. Over the initial set of trials there was a speed-accuracy tradeoff, but this did not continue over subsequent trials; strategically trained teams continued to outperform factually trained teams on delay time, but performed similarly on accuracy for the remainder of the trials. Strategies were aimed at reducing the time to complete the necessary components of the CTEAM task and the results reflect this training benefit. Overall, these results imply that the strategic training method elicited positive long-term effects on performance in the dynamic decision environment.

Performance Assessment of Strategic Team
Training in Simulated Air Traffic Control

Dwyer (1984) defines a team as two or more individuals who have a specific role and specialized skills or knowledge that are used in an interdependent way to accomplish a common task. Cooke, Salas, Cannon-Bowers, and Stout (2000) incorporate a division of labor into their definition, in which team members perform heterogeneous tasks. Therefore, a complete definition of a team utilizes heterogeneity of team membership and interdependency to accomplish a given task.

The interdependency of members illustrates the key difference between a group and a team. If a group of people is given multiple tasks to complete and one group member fails at a given task, then it is necessary for another group member to compensate for this failure. In contrast, team members cannot compensate for another member's poor performance. If one team member fails at a given task, the entire team performance suffers. This difference between a group and a team is important, as a group places emphasis on independent individual performance, while a team emphasizes individuals functioning as one cohesive unit.

Controller Teamwork Evaluation Assessment Manual (CTEAM)

Edwards (1962) defined dynamic decision making by three factors. The first is that a series of actions must be made over time to achieve an overall goal. The second is that these actions are interdependent, such that later decision making depends on earlier actions. Third, the decision environment changes both spontaneously and as a consequence of earlier actions.

Computer microworlds, such as CTEAM, reflect the three criteria proposed by Edwards. CTEAM (Bailey, Broach, Thompson, & Enos, 1999) is a dynamic air-traffic-control microworld that contains four sectors, one for each participant to control. [See screen shot of C-TEAM in Figure 1.] Performance in each individual sector is dependent upon performance in adjoining sectors, and thus, CTEAM is a team task.
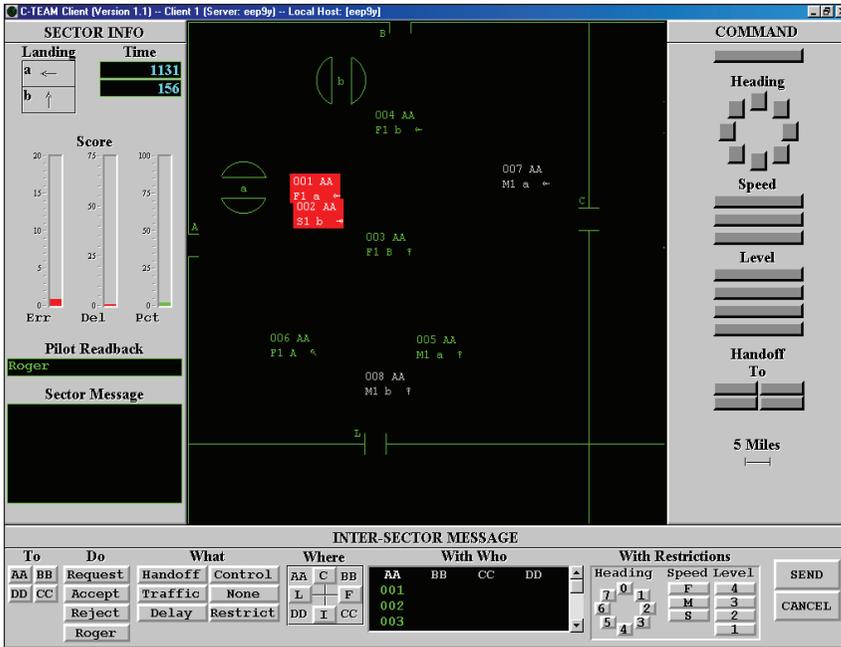
Note: On the right is the command action control panel by which operators controlled heading, speed, and altitude. On the left are feedback indicators for each of the performance variables: number of errors, delay time, and percent of planes landed. In the center, is the amount of visible airspace available to an operator.

Figure 1. Screen shot of CTEAM microworld simulation.

The objective of CTEAM is to direct aircraft to their appropriate destination as efficiently and safely as possible. The current study utilized a four-sector (team) version of CTEAM with a single controller for each sector. CTEAM is especially appropriate for studying team performance because it allows for experimental control in creating scenarios. When using this microworld platform, the researcher is able to manipulate, a) how long the scenario lasts, b) how many aircraft will appear in each sector throughout the scenario, c) where each aircraft will appear, and d) the location of obstacles and airports within the sector.

All the scenarios in the current study lasted 28 minutes, which has been demonstrated to be adequate time to complete each scenario (Raacke, 2003). Each of the four sectors contains two airports and two gates to pass aircraft from one controller to another. Finally, all of the aircraft started in the same relative location within each sector, equidistant from their respective airports. CTEAM saves all scenarios in their entirety allowing an exact replay of events in real-time.

The CTEAM controllers have the ability to direct aircraft by changing their direction (heading), speed, and altitude on a command toolbar. In addition, each operator receives feedback on their performance via a display of error bars for each of the three performance variables (time delay, operational errors, and percentage of aircraft landed).

The first performance variable assessed is the time delay for each aircraft. The CTEAM program estimates the optimal enroute time for each aircraft to reach its respective airport. Deviations from this calculated optimal time route adds time delay to the team's delay score, which has a cumulative negative impact on team performance. Thus, if the controllers do not hand off aircraft from one team member to another, if the aircraft encounters a barrier, or if the aircraft do not reach their respective airports in the optimal time, then there will be higher time delay. Likewise, if an aircraft crashes, then that aircraft is automatically assigned a time value at the total length of the scenario. Thus, time delay is also directly related to the number of crashes.

The second performance variable is the number of operational errors that occurred. Operational errors are defined by a count of the number of aircraft that fly into restricted airspace, including separation errors. Separation errors are defined as when one aircraft gets too close to another aircraft at the same altitude or when an aircraft gets too close to barriers. Note that a separation error does not necessarily mean that the aircraft crash into each other. When a separation error is committed, the aircraft turns red, warning that a crash may occur. Both separation errors and crashes are considered operational errors; the more errors that are committed, the more deleterious the effect on the team's overall performance.

The final performance variable is percentage of aircraft reaching their destination; the higher the score, the better the performance. CTEAM calculates the number of aircraft in each sector along with their destination sector. Once an air-

craft reaches its airport, the team percentage increases for the landing of that aircraft. This performance variable is inversely related to the number of aircraft crashes; therefore, as more aircraft crash, the team percentage of aircraft reaching their final destination is lower.

Each aircraft had a data block next to it (Figure 1), which contained information about direction, speed, altitude, and the route that the aircraft was to follow; the controller could manipulate each of these components. The number of aircraft for each sector varied with the difficulty of the scenarios: the low-density scenarios contained 24 aircraft (6 for each sector), the medium-density contained 36 aircraft (9 for each sector), and the high-density scenarios contained 48 aircraft (12 for each sector). Practice scenarios incorporated portions of the low, medium, and high-density scenarios, such that these scenarios became increasingly difficult as time continued; however, these scenarios are not included in the analysis.

At the team level, the task for participants is to activate their aircraft and hand them off to their teammates in order for the aircraft to reach their destination as quickly as possible, while making the fewest errors. Participants are able to view their airspace, and a portion of the airspace of two of their teammates. Since participants only have control over their respective airspace and rely on their teammates to operate effectively within each of theirs, this makes for a true team task. That is, if a team member does not hand off their aircraft, or crashes aircraft, then the entire team score will be poor.

Training on CTEAM

The purpose of training on any task is to improve performance. For instance, Gaeth and Shanteau (1984) were able to improve the accuracy of soil judges by training them to reduce the influence of irrelevant materials (e.g., excess moisture) on soil classification. Moreover, they found that hands-on "laboratory" training was more effective than formal "lecture" training.

Training is different from practice. Practice effects result from repeated exposure to the task. Training involves the accomplishment of certain performance outcomes specifically targeted by instruction or other means. Of course, it is possible to observe both training and practice effects: the former from instruction and the latter through repetition. A relevant example of practice comes from Thomas, Willems, Shanteau, Raacke, and Friel's (2001) study of air traffic controllers. Unfortunately, there was no control group, making interpretability difficult.

An example of a training study on individuals can be seen in Friel, Thomas, Raacke, and Shanteau (2002). In that study, half of the participants received strategic training on how to improve individual CTEAM performance and the other half received factual training. The strategic training involved showing participants strategies that were gathered from expert air traffic controllers. One such strategy was to fly the aircraft in a *diamond* formation instead of a box formation, which helps aircraft reach their destination faster. Another strategy, called the *"string of pearls,"* involved lining up aircraft one right after the other, like a string of pearls. A third strategy, placed all aircraft at one particular altitude going counterclockwise with aircraft at a *different altitude* going clockwise. Overall, there were eight different strategies.

Results from Friel et al (2002) showed that individuals who received the strategic training made significantly fewer operational errors and performed better than those who received factual training.

Overview of the current study

The purpose of the current study was to apply the individual training approach used by Friel, et al. (2002) to teams. In addition, the teams used in the current study will be compared to those from Raacke (2003) that used only factual training procedures for the participants, i.e., providing information on how to read the computer screen, how to issue directional commands, and how to manage the sector. Therefore, based on the past literature and the results from Raacke (2003), the following research questions were developed:

RQ1: Will strategic team training positively affect delay time, such that strategically trained teams will have lower delay time than factually trained teams?

RQ2: Will strategic team training positively affect number of errors committed, such that strategically trained teams will commit fewer errors than factually trained teams?

RQ3: Will strategic team training positively affect percentage of aircraft landed, such that strategically trained teams will land more aircraft than factually trained teams?

Method

Participants

Twenty participants from a large Midwestern University participated in the current study. All of the participants were undergraduates and received $10 per 1.5 hours (one session) of their time. Five, four-person teams were created by matching participants on the availability of their schedules. Twelve of the participants were male, eight were female, and 95 percent were Caucasian.

Materials

CTEAM, a real time air traffic control simulator, which has been used by the Federal Aviation Administration, was the computer program used. As in previous work using CTEAM, the objective of the participants was to use this dynamic microworld to command aircraft to reach their destination as quickly as possible while making the fewest operational errors.

Training Videos. Three training videos were viewed by strategically trained teams prior to the start of the measured trials. These videos were a combination of Microsoft Power Point presentations and various replay files, which depicted CTEAM scenarios of how to correctly and incorrectly complete various tasks (e.g., handing off aircraft). The first training video comprised the basics of CTEAM. Specifically, this video contained information on how to change the direction, speed, and altitude of an aircraft, how to activate an aircraft, and the definition of air traffic control; this was consistent with the initial information provided by Raacke (2003).

Unique to the current study, however, was the use of replay files that specifically demonstrated how to change the dynamics of the aircraft (i.e., changing direction, altitude, and so forth) and what separation errors look like. Participants were subsequently asked to demonstrate the previously viewed skill set.

The second video consisted of individual CTEAM strategies on how to maneuver aircraft from one gate or location to another successfully. These strategies were utilized by Friel et al. (2002), which included the diamond method (moving aircraft in a diamond shape instead of a box shape), string of pearls (lining up aircraft that are going in one direction), and the different altitude method. The replay files consisted of the correct way to control aircraft (diamond) and an incorrect way (square).

The final video contained information on how to succeed in CTEAM as a team, rather than at the individual level (video two). This video illustrated the correct and incorrect way to make handoffs and to change the altitude of aircraft during handoffs so that they do not crash into each other at a gate (handoff location). The replay files consisted of the correct and incorrect way to complete the actions.

Each video lasted approximately 30 minutes, and any questions asked by the participants were answered. In addition, after each video was viewed, the participants performed the CTEAM task either at an individual (after video 1 and 2) or a team (after video 3) level. Hence, these videos were interactive in the respect that after participants were shown how to perform a task in CTEAM, they were then asked to demonstrate compliance with each skill set.

Procedure
The overall design of the experiment followed the training protocol used by Raacke (2003), which used practice, low, medium, and high-density scenarios. The current training protocol utilized the same gradual increase in difficulty. After the three training videos and compliance trials, the teams engaged in two practice scenarios, then two low density, two medium density, and two high-density scenarios. The purpose of the two practice trials in the initial session was to familiarize the participants with working as a team on CTEAM. Each CTEAM session lasted 28 minutes, and each team completed two sessions every time they came into the lab. Each session consisted of two scenarios of the same density (e.g., Low-Low, Medium-Medium, or High-High).

Due to errors outside the control of the researcher, ie, participant cancellations, only the first eight trials will be analyzed. Identical to Raacke (2003), low, medium, and high-density scenarios differed from one another in the number of aircraft presented throughout the scenarios. After the second session was completed for each day, the participants were thanked for their time, paid, and told to return for the next scheduled time. Upon completion of the entire experiment, the participants were thanked and fully debriefed.

Behavioral Measures
The total delay time, number of operational errors, and percentage of aircraft that reached their destination were computed for each of the strategically trained and factually trained teams for the first eight trials. The total delay time was calcu-

lated based on the amount of time it took the aircraft to reach their destination, with lower delay time indicative of better performance. The number of operational errors was computed by summing the number of times any given aircraft crashed into the barriers, other aircraft, airports, and/or the number of times the aircraft got too close to obstacles, with fewer errors being indicative of better performance. Finally, the percentage of aircraft to reach their destination was calculated based on the cumulative percentage of aircraft that landed in their respective airports within the time constraints of the scenarios, with a higher percentage being indicative of better performance.

Results

In order to address the research questions, mean delay time, mean number of operational errors, and mean percentage of aircraft landed for the first eight trials were computed for the strategically trained and factually trained teams. As seen in Figure 2, mean delay times were consistently less for strategic training across all three densities. Indeed, a point-by-point comparison reveals nearly a 500 msec reduction in delay times for strategic training.
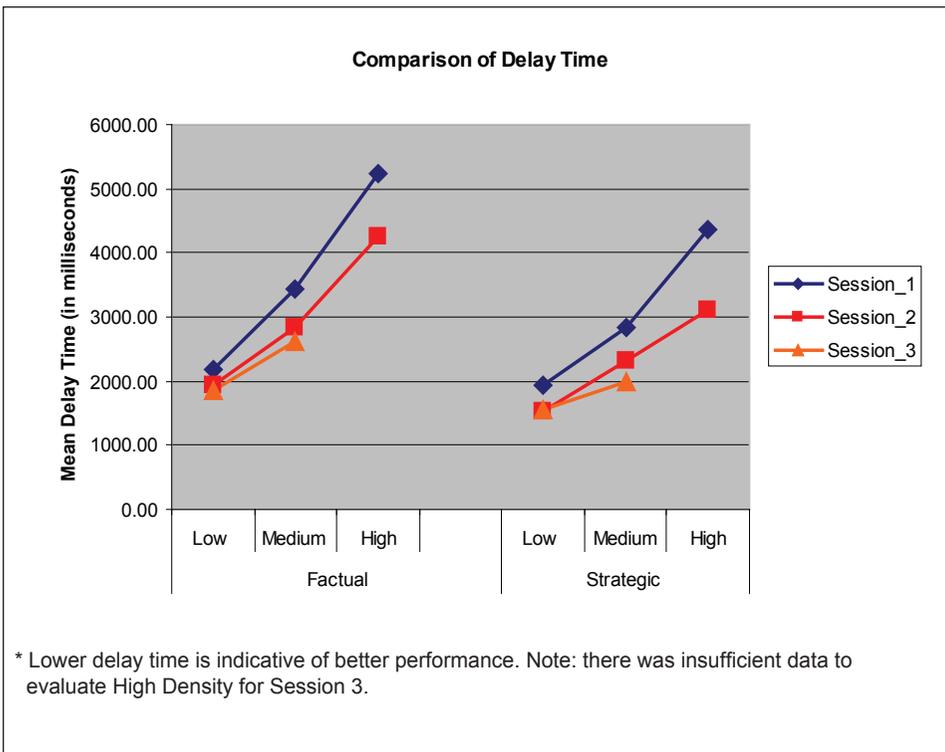
Figure 2. Mean delay time for the strategically trained and factually trained teams for all sessions.

In order to determine if training was beneficial for reducing time, an 8 (time) x 2 (study) mixed analysis of variance (ANOVA) was conducted with delay time as the within-subjects factor. Results showed that there was a significant main effect for time, $F(7,56) = 14.26$, $p < .0001$, *partial η^2* = .64, power = 1.00. There was not a significant main effect for study, nor a significant time x study interaction. Although not statistically significant, the results indicate that the strategically trained teams outperformed the factually trained teams on time delay for all experimental sessions.

The results for the number of operational errors appear in Figure 3. In the first session, those who were strategically trained had more errors than the factually trained group. However, there was little difference between the two training conditions for sessions 2 and 3. A mixed Anova showed that there was a significant main effect for time, $F(7,56) = 11.72$, $p < .0001$, *partial η^2* = .59, power = 1.00. There was not a significant time x study interaction or a main effect for study. This result suggests that training method had little effect on number of errors.

* Fewer number of errors is indicative of better performance. Note: there was insufficient data to evaluate High Density for Session 3.
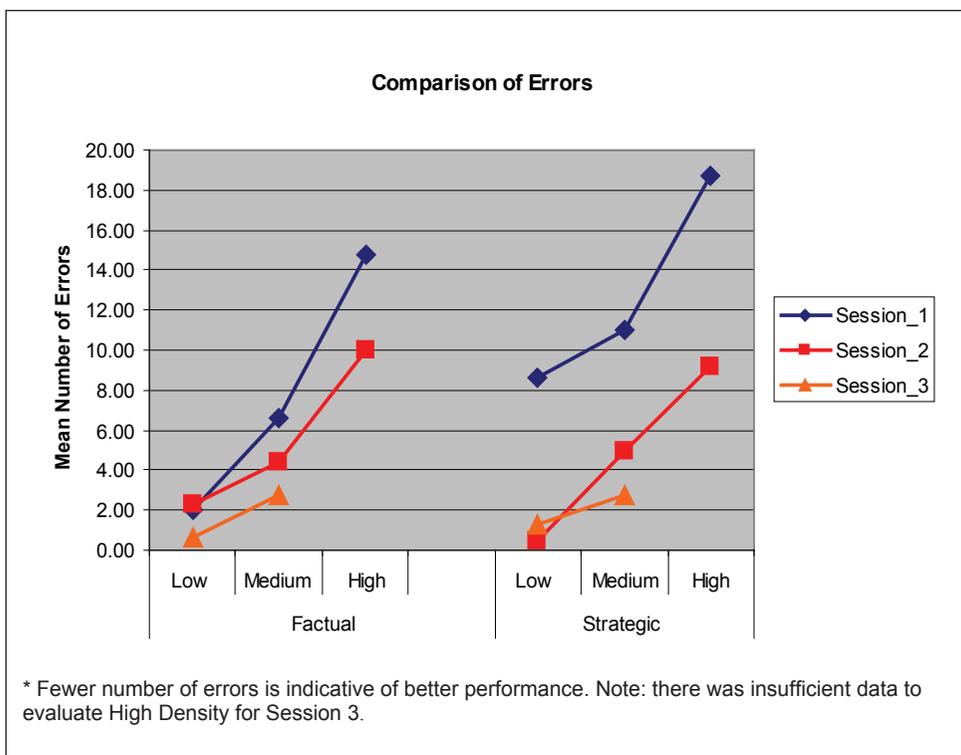
Figure 3. Mean number of operational errors for the strategically trained and factually trained teams across all sessions.

As shown in Figure 4, at least 79% of all aircraft reached their destination within the allotted time. However, there is little evidence of systematic differences between training conditions. A mixed Anova showed a significant main effect for time, $F(7,56) = 2.97$, $p < .01$, *partial η^2* = .27, power = .90. There was no significant

main effect for study, nor a significant time x study interaction. These results suggest the training did not affect the percentage of aircraft landed on time.
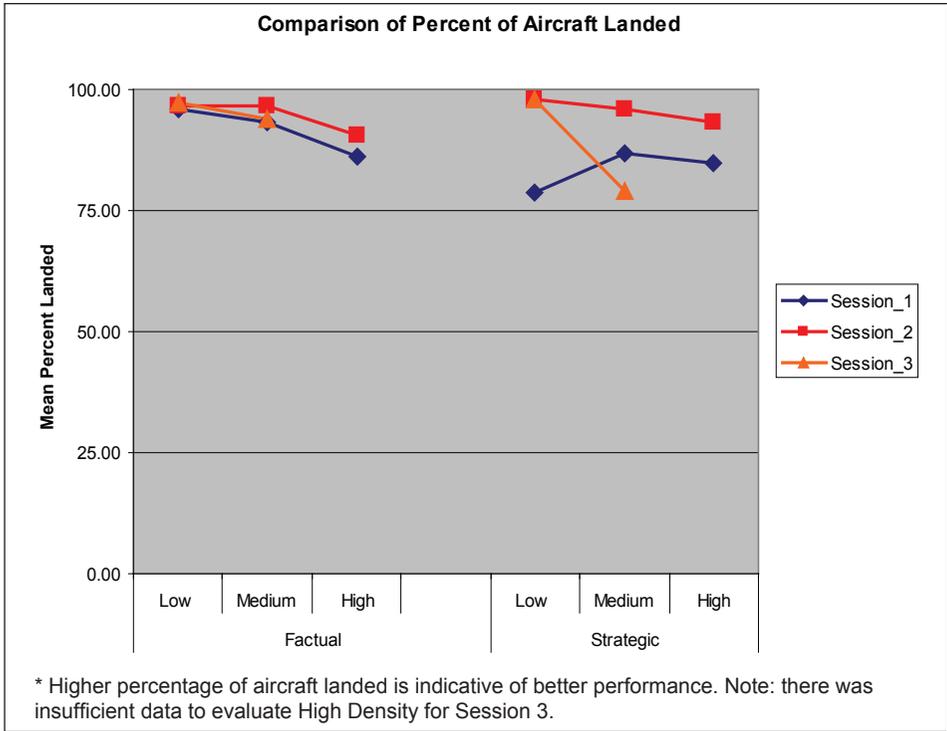
Figure 4. Mean percentage of aircraft that reached their destination for the strategically trained and factually trained teams for all sessions.

Discussion

The purpose of the current study was to compare two types of team training in air traffic control: strategic vs. factual training. The first research question asked whether strategic training would lead to lower delay times. The results showed that strategically trained teams consistently performed more quickly across all conditions in the experiment. For the second and third research questions, however, there were no significant differences of training on numbers of errors or on the percentage of aircraft landed.

Given the content of the training, it should not be surprising that strategically trained operators worked more quickly. That is, strategies such as "Diamond in the Sky" are intended to achieve the quickest route to destination. In contrast, these strategies had little direct relation to the other dependent variables – number of errors and percent reaching destination.

Conclusions

It is important to determine the effect that training has on a team, especially when the poor performance of that team has serious consequences, as in an air traffic control team. Past research on individuals (Friel et al., 2002) demonstrated that strategic training significantly increased performance compared to factual. The current study extended these findings to teams. That is, strategically trained teams were superior on delay times compared to factually trained teams.

Although the results from the current study show that strategically trained teams operated more quickly, there were no significant differences in the number of errors nor in the percentage of aircraft landed. Given the emphasis on efficient air traffic strategies in the training, this result is encouraging, as it suggests that team training can have a positive impact on aggregate behavior.

The results from the current studies have implications for the Federal Aviation Administration, especially within the context of team training for air traffic controllers. The current study provides evidence that strategic team training is effective under certain conditions. First, the strategies should correspond directly to the behavioral measure; in this study, minimum time strategies lead to reductions in delay time. Second, operators need to be shown how the strategies work; videos were used here. Third, the training needs to be phased in, so that the emphasis is first on individual performance and then on team performance.

However, it should be noted that the present training protocol did not incorporate traditional team training skills, such as teamwork building. Of course, it is possible to combine the present strategy-based approach with teamwork building or team cohesion exercise.

Despite the positive training results from this study, the degree of generalizability is not guaranteed. In other words, future research is needed to determine if the present training approach is effective in operational air traffic control.

References

Bailey, L. L., Broach, D. M., Thompson, R. C., & Enos, R. J. (1999). *Controller teamwork evaluation and assessment methodology: (CTEAM): A scenario calibration study.* (DOT/FAA/AAM-99/24). Washington, DC: Federal Aviation Administration Office of Aviation Medicine. Available from: National Technical Information Service, Springfield, VA 22161.

Barlett, C. P., & Vowels, C. L. (2006). *External factors and air traffic control: Examination of expert performance under differing situations*. Oral Presentation at the 15th Annual Oklahoma-Kansas Judgment Decision Making Conference, Oklahoma City, OK

Cooke, N. J., Salas E., Cannon-Bowers, J. A., & Stout, R. (2000). Measuring team knowledge. *Human Factors, 42*(1), 151-173.

Dwyer, J. L. (1984). Team research and training: A state of the art review. In F. A. Muckler (Ed.), *Human Factors Review: 1984*. Santa Monica, CA: Human Factors Society.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review, 69*, 109-135.

Friel, B. M., Thomas, R. P., Raacke, J. D., & Shanteau, J. (2002). *Assessing training efficacy in an air traffic control microworld*. Paper presented at the Oklahoma-Kansas Judgment and Decision Making Meeting. Oklahoma City, OK.

Gaeth, G. J., & Shanteau, J. (1984). Reducing the influence of irrelevant information on experienced decision makers. *Organizational Behavior and Human Performance, 33,* 263-282.

Raacke, J. (2003). Assessment of longitudinal team development in decision-making. Unpublished masters thesis, Kansas State University.

Thomas, R. P., Willems, B., Shanteau, J., Raacke, J., & Friel, B. (2001). CWS applied to controllers in high fidelity simulation of air traffic control. *Proceedings of the 2001 International Symposium on Aviation Psychology*.

The Significance of Demographic Characteristics in Airport Driver Training Programs

William B. Rankin, II, AAE

*The University of Central Missouri
TR Gaines 210
Warrensburg, MO 64093
954-806-9840*

Abstract

This is the second in a series that examines the problem of runway incursions at Operational Evolution Plan-35 (OEP-35) U.S. towered airports. According to the FAA Runway Safety Report (2004), vehicle deviations accounted for 20% (291 events) of all runway incursions during the period of 2000 through 2003. The focus of this quantitative correlational study examined if demographic characteristics are a significant factor in the airport driver training that employees receive at OEP-35 airports. Airport driver training officials were surveyed using a five-point Likert-type survey. The data suggested that demographic characteristics are significant factors in airport driver training and vary by geographic region. The data may assist airport operators in identifying significant demographic characteristics that affect the outcomes and the potential improvements that may enhance airport driver training programs in various geographic regions.

This is the second in a series that examines the problem of runway incursions at the Operational Evolution Plan-35 (OEP-35) U.S. towered airports. The data from the first study suggested that a relationship existed between the methods used for airport driver training and the number of runway incursions at the largest U.S. towered airports. The American Association of Airport Executives' (AAAE) interactive computer-based training program was found to be the most effective. The study supported the finding that both traditional and AAAE interactive computer based training were effective in reducing the overall number of runway incursions for all classes of incursions. This study examines the demographic data collected but not analyzed in the first study to determine what demographic data were significant in driver training programs, if any.

Since the mid-1920s, commercial aviation in the United States has achieved a remarkable safety record. Within the National Airspace System (NAS), millions of operations are completed safely every year (Federal Aviation Administration, 2005a). The combination of the pressure to reduce system delays, the complexity of airport operations, and the requirement for precise timing, make the airport movement areas unforgiving of errors by pilots, air traffic controllers, and vehicle drivers (FAA, 2002a).

According to Clarke (2002), the Federal Aviation Administration (FAA) has developed several training programs for pilots and air traffic controllers to make each group more aware of runway incursion problems. In addition, the FAA instituted Standardized Taxi Routes (STRs) by FAA Order 7110.116, to assist pilots and air traffic controllers with surface movement of aircraft. Finally, air traffic controllers are required to maintain a high level of runway incursion awareness through a monthly computer-based recurrent training program titled *Preventing Runway Incursions.*

Rankin (1994) identified training of ground vehicle operators as the most effective FAA initiative to reduce runway incursions, however, ground vehicle operator training is conspicuously absent from mention in most literature; even though vehicle operators traverse airport movement areas on a daily basis.

On June 21, 2002, FAA issued Advisory Circular (AC) 150/5210-20 to provide guidance to airport operators in developing training programs for vehicle ground operations. This was the first advisory circular providing airport operators with a list of training topics to include in a ground vehicle operator-training curriculum (FAA, 2002a).

The FAA (2004b) defined runway incursions as, "Any occurrence at an airport involving an aircraft, vehicle, person, or object on the ground that creates a collision hazard or results in a loss of required separation with an aircraft taking off, landing, or intending to land" (p. 9). The NAS continues to experience approximately one runway incursion per week, which is classified as significant or a barely avoided collision (FAA 2004b).

Runway incursions are divided into three classification types. These types include pilot deviations, operational deviations, and vehicle deviations. In the United States, pilot deviations account for approximately 57% of the total runway incursions, operational deviations account for 23%, and vehicle deviations account for 20% (FAA, 2004). After type, runway incursions are further stratified into four distinct categories by increasing severity, ranging from category D, the least severe, to category A, the most severe. Figure 1 illustrates the runway incursion categories by severity.

*Low*_____*High*

Category D	Category C	Category B	Category A
Little or no chance of collision but meets the definition of a runway incursion	Separation decreases but there is ample time and distance to avoid a potential collision	Separation decreases and there is a significant potential for collision	Separation decreases and participants take extreme action to narrowly avoid a collision, or there is a collision

Figure1. Runway incursion categories by increasing severity. (FAA, 2004)

Statement of Problem and Purpose

The first study addressed the problem of runway incursions at the largest U.S. OEP-35 towered airports. The focus of this quantitative correlational study examined if demographics are a significant factor in the airport driver training that employees receive. Although vehicle deviations represent a smaller portion of the total U.S. runway incursions, the potential risk in the terms of loss of life is significant. The most serious runway incursion to date (a pilot deviation) occurred in Tenerife, Canary Island on March 27, 1977, killing 583 people, and ranking as the worst disaster in aviation history (Clarke, 2002).

The purpose of this study was to identify those demographic characteristics, if any, which may be a significant factor in airport driver training.

This study was significant in that no previous study has examined if demographic characteristics are a significant factor in the airport driver training that employees receive at OEP-35 towered airports.

Review of Literature

The FAA Runway Safety Blueprint 2002-2004
The goals of the FAA *Runway Safety Blueprint* (2002a) are consistent with those identified by Rankin (1994) for all airports and included the following:

1. Develop and distribute runway safety education and training materials to controllers, pilots, and all other airport users.
2. Increase surface safety awareness throughout the aviation community.
3. Assess and modify procedures to enhance runway safety.
4. Improve runway safety data collection, analysis, and dissemination.
5. Identify and implement enhancements to improve surface communications.
6. Increase situational awareness on the airport surface.
7. Support and deploy new technologies that reduce the potential for collision.
8. Implement site-specific runway safety solutions in coordination with local aviation communities. (p. 4)

A Brief Current History on Training

Beldsoe (1999) pointed out that it was not until the 1950s that training literature began to mention the need to involve top management in training decisions. Training directors (that first appeared in the 1940s) became mid-to-high salaried positions in organizations.

In 1959, Kirkpatrick developed the first evaluation model that focused on training aspects or levels (Kirkpatrick, 1996). A popular topic at this time was organizational development, which led to the term human resource development by Leonard Nadler. Programmed training led to the popularity of teaching machines to deliver training (Bledsoe, 1999).

The 1970s and 1980s saw an increased emphasis on social issues in training. Popular topics included self-esteem training, quality circles for management, and diversity training (Bledsoe, 1999).

Bledsoe (1999) stated that with the election of President Bill Clinton, the 1990s saw the establishment of the Office of Work Based Learning and a national endorsement of public sector training. Current training initiatives include diverse areas from global organizations, performance support teams, information systems, and interactive computer-based programs to help workers and employers achieve a balance among training, work, and family.

Since training has been historically linked to education, it is not surprising to see computer-based interactive training linked to distance education. Nanney (n.d.) stated distance learning could be interactive or non-interactive:

Interactive learning can be synchronic or asynchronic, or a combination of the two. Synchronic learning is where the teacher and student perform interactively at the same time on the same subject and in every learning action they perform as in the traditional classroom. Non-interactive learning is mainly represented by the World Wide Web where the media transfers the knowledge to the learner. Distance education in its many forms is organized so that it is at some point between totally interactive and completely non-interactive learning. (Nanny, n.d., p. 1)

According to Filipczak (1996), computer-based interactive training can be the most effective way to train. Almost every example of effective computer-based training is based on multimedia simulation or a simulated environment. Filipczak stated:

The most powerful learning environments are simulators, but only when learning is designed into the environment. The point of multimedia training is not to create an interactive environment for its own sake, but to create a dynamic educational construct where people can learn. (p. 1)

According to Rohland (1996), "Computerization is the trend in training. More companies are seeing that taking people into the classroom is not the answer" (p. 1).

Past studies by Beckman (2000) and Ortiz (1993) concluded that computer-based training devices were found to have positive skill transfer capabilities. In the study by Ortiz, the time to learn a new skill was reduced if computer-based training was used prior to traditional training. On the same subject, a study was conducted by Baharestanl (2005) to determine whether computer-based learners would do as well as rote learners if both received the same instruction and method. The study concluded that computer-based training, reaching diverse groups such as air traffic controllers, pilots, and physics students, must engage the learner at the rote levels to ensure that all levels of learning are achieved.

Training and Airports

According to Ragan (1997), the risk of misunderstanding ATC instructions communicated via the radio is high and can have deadly consequences. Correctly understanding ATC information provided by the controller is essential for safe airport surface operations and can only be learned through a comprehensive driver-training program. For example, a paper by Ragan (1997) concluded with an anecdote about a student pilot with limited English proficiency who was asking the tower for permission to enter the traffic pattern to make a landing. The tower could not fully understand what he wanted, so the air traffic controller asked the student to state his intentions. The student responded by saying, "I intend to become a private pilot" (p. 34).

Finney (2000) pointed out that in a recent AAAE Airport Training Survey and Needs Assessment, over half of the U.S. airports still lack a specifically designated training manager, training remains primarily a decentralized function, funding continues to be a pressing concern, and there is a lack of consistency and comprehensiveness.

The FAA (2004) determined that the 35 busiest U.S. airports have twice the average number of reported runway incursions, and the problem continues to persist despite the best efforts of the industry. This is evidenced by the news articles in Table 1.

Table 1
Recent Newspaper Articles on Runway Incursions

"Airplane Fender-Bender At PHL Airport" CBS 3 (Philadelphia) – Nov. 30, 2004 12:05 PM (ET) There were some tense moments at Philadelphia International Airport on Tuesday morning when a small plane brushed a tractor towing an empty Delta passenger jet
"3 Hurt When Jet Brushes Airport Service Vehicle" Chicago Tribune – Nov. 19, 2004 09:44 AM (ET) Three people in an airport service van were injured Thursday when an airplane nearing a gate at O'Hara International Airport clipped the vehicle.
"2nd Runway Mix-Up AT LAX IS Investigated" Los Angeles Times – Nov. 16, 2004 07:00 AM (ET) A series of mistakes last week by a pilot and a controller caused a corporate jet to land on a runway at Los Angeles International Airport that two other aircraft had already been cleared to use, federal aviation authorities confirmed Monday.

"Near-Miss At Airport Was Fifth This Year" Cincinnati Enquirer – Nov. 11, 2004 07:23 AM (ET) The most recent close call on the runways at the Cincinnati/Northern Kentucky International Airport was the fifth in 2004, setting a record for such "incursions" at the airport.
"Feds See Aviation Reporting Errors" USA Today – Nov. 10, 2004 06:04 AM (ET) The dramatic near collision of two jets carrying about 300 people in Los Angles last August highlights the failure of federal regulators to keep track of dangerous runway incidents, federal investigators charged Tuesday.
"Plane Collides With Airport Vehicle" KYW-TV (Philadelphia) – Oct. 29, 2004 09:25 AM (ET) A commercial airliner apparently collided Wednesday with an airport support vehicle while taxiing to a runway at Philadelphia International Airport.
"FAA Probes 2 Close Calls At O'Hara" Chicago Tribune – Jul. 08, 2004 08:49 PM (ET) Two city trucks strayed onto an active runway at O'Hara International Airport within minutes of each other last month, forcing two planes to abort their landings and prompting federal investigation.
"Airport Tells Its Tenants To Slow Down" Denver Business Journal – May 18, 2004 12:30 PM (ET) Complaints about speeding and failing to yield to airplanes prompted Denver International Airport security to send a warning letter last month to all airport tenants. Lori Beckman, director of security, reminded tenants that violations of the driving rules within the airport property could result in suspension.

Note. From: AAAE (2005)

Kirkpatrick's Model

Kirkpatrick addressed the four *aspects of training* evaluation in Kirkpatrick's model with respect to training effectiveness (p. 44). According to Tidler (1999), Kirkpatrick's model is widely accepted by the American Association for Training Development, and there are four *aspects of training* in Kirkpatrick's model with respect to training effectiveness (p. 44). These aspects include:

1. Reactions -- What trainees' say about the value of the training.
2. Learning -- Objectives met, knowledge, and skills learned.
3. Behavior -- The skills acquired are implemented on-the-job.
4. Results -- Impacts on job performance. (Kirkpatrick, 1983, ¶ 6)

In this study, Kirkpatrick's model was selected as most appropriate for the development of a model for the study of airport driver training methods at the largest U.S. towered airports.

A review of data from the U.S. Census Bureau data (2007) revealed the following U.S. population demographic characteristics for the US population:

1. Race – white 80.2%; black 12.8%; American Indian and native persons 1%; Asian persons 4.3%; native Hawaiian and other Pacific islander 0.2%; persons reporting two or more races 1.5%; persons of Hispanic or Latino origin 14.4%; white persons not Hispanic 66.9%. Language other than English spoken at home 17.9%
2. Age – persons under 18 years old 24.8%; persons 65 years old and older 12.4%.
3. Education – high school graduates 80.4%; bachelor's degree or higher 24.4%.
4. Income – median household income $44,334; per capita money income $21,587; persons below poverty 12.7%.
5. Marital – N/A (U.S. Census Bureau, 2007)

A review of data from the U.S. Census Bureau data (2007) revealed the following U.S. population demographic characteristics for South Florida:

1. Race – white 66.6%; black 22.3%; American Indian and native persons 0.2%; Asian persons 0.7%; native Hawaiian and other Pacific islander 0%; persons reporting two or more races 4.7%; persons of Hispanic or Latino origin 65.8%. Language other than English spoken at home 74.6%
2. Age – persons under 18 years old 21.7%; persons 65 years old and older 17%.
3. Education – high school graduates 52.7%; bachelor's degree or higher 16.2%.
4. Income – median household income $23,483; per capita money income $15,128; persons below poverty 28.5%.
5. Martial – N/A (U.S. Census Bureau, 2007)

Highlights of Methodology

For the purposes of this study, a quantitative and limited qualitative methodology was used. A five point Likert-type survey instrument was used to collect the necessary data and qualitative responses to five end-of-survey questions (see Appendix A). The first study identified one independent variable *method of training.* The independent variable consisted of two airport driver-training methods – AAAE interactive computer-based airport driver training and traditional airport driver training. The study also identified four dependent variables *runway incursion categories A through D (Figure 1).*

In this study, five intermediary independent variables or *demographics* were analyzed. The demographic variables include (a) race, (b) age, (c) education, (d) income, (e) and marital status. The statistical analysis used in this study was multivariate analysis of variance (MANOVA).

Limitations of Study

According to Wells and Rodriques (2004), the National Plan of Integrated Airport Systems (NPIAS) contains a listing of more than 3,334 public funded airports in the United States. Of the 3,334 plus public funded airports listed in the NPIAS, only 490 airports have operating control towers, and only towered airports report

runway incursions (FAA, 2004). With respect to the towered airports, FAA (2004) stated commercial aviation aircraft operations at the OEP-35 airports are predominantly commercial aircraft and account for "the majority (87 percent) of category A and category B runway incursions" (p. 36).

This study was limited to an analysis of vehicle deviations that cause runway incursions at the OEP-35 airports. According to the FAA (2004), "vehicle/pedestrian deviations represented 18 percent of the runway incursions at the OEP-35 airports, which is in proportion to their national representation (20 percent)" (p. 36).

There were several potential limitations with respect to the survey instrument. These included (a) the effective sample size of participants, (b) the accuracy of the data provided by the participants, and (c) the pitfall of correlation versus causation for forming conclusions.

Statement of Hypotheses

In concert with the stated research question: Are demographics characteristics significant factors in the airport driver training employees receive at OEP-35 airports? There was one null and alternative hypotheses framed for this study. The hypotheses were formulated as follows:

1. H_{01} : Demographics characteristics are not significant factors in the airport driver training that employees receive at OEP-35 airports.
2. H_{11} : Demographics characteristics are significant factors in the airport driver training that employees receive at OEP-35 airports.

Selection of Participants

The population sampled for this study was comprised of employees that completed airport driver training from 18 of the OEP-35 airports responding to the survey. Targeted participants included 390 randomly selected employees who have successfully completed airport driver training and who are authorized to drive vehicles onto and within the airport movement areas. A scaled survey instrument was used to gather the data on the demographic (see Appendix A).

Discussion of Data Processing

Power analysis software obtained from the UCLA Department of Statistics was used to estimate the required number of *completed* surveys. The calculations showed that at least 194 completed surveys needed to be collected from participants at the OEP-35 airports to estimate the mean response values for questions 1 through 26 within a desired precision of .10 (University of California, 2005).

FAA (2004) considers runway incursions rare events relative to total aircraft flights over finite periods of time (5.6 incursions per half million aircraft flights per year). According to Aczel and Sounderpandian (2002) "if we count the number of times a rare event occurs during a fixed interval, then that number would follow a

Poisson distribution" (p. 151). Using software obtained from the UCLA Department of Statistics webpage (University of California, 2005), a Poisson power analysis was used to estimate the number of years of runway incursion data needed from the 2004 FAA *Runway Safety Report*.

The statistical analysis used in the study included multivariate analysis of variance (MANOVA). For MANOVA the independent variable was specified as *method of training* and the covariates were specified as *demographics*. The dependent variables were specified as *runway incursions categories A through D.*

Since more than one dependent variable was specified, the MANOVA using Pillai's trace, Wilks' lambda, Hotelling's trace, and Roy's largest root criterion with approximate *F* statistic was provided as well as any subsequently needed univariate analysis of variance for each dependent variable (Norusis, 2003).

Reliability of Survey Instrument
The first step was to determine the reliability of the survey instrument. According to Norusis (2003):

In classical theory, a subjects' response to a particular item is the sum of two components: the true score and the error. The true score is the value of the underlying construct that is being measured; the error is the part of the response that is due to question-specific factors. The index most often used to quantify reliability is Cronbach's Alpha. Good scales have values larger than 0.8. (pp. 437-438)

In the case of this study, SPSS© software was used to calculate the Cronbach's Alpha value of 0.864 shown in Table 2 for the 26 survey questions used to study driver training methods.

Table 2
Cronbach's Alpha Reliability Statistics

Cronbach's Alpha	N of items
.864	26

Distribution of the Dependent Variables

The second step was to determine what distribution the dependent variables (runway incursion categories A through D) followed. As previously stated, FAA (2004) considers runway incursions rare events relative to total aircraft flights over finite periods of time (5.6 incursions per half million aircraft flights per year). According to Aczel and Sounderpandian (2002) "if we count the number of times a rare event occurs during a fixed interval, then that number would follow a Poisson distribution" (p. 151).

Multivariate Analysis of Variance (MANOVA) Results

The third step was to examine if demographic characteristics are a significant factor in the airport driver training that employees receive at the Operational Evolution Plan (OEP-35) US towered airports.

MANOVA of the Effects of Demographic Characteristics for OEP-35 Airports.
The MANOVA analysis identified the effects of the covariates in the model for all *OEP-35* airports. The variable *race* was not found to be statistically significant at the 0.228 level. This variable measured the participants' races in the categories of white, African-American, Hispanic, Asia-Pacific Islander, and Native American. The variable *age* was found to be statistically significant at the 0.000 level. This variable measured the participants' age in the categories of 18-25 years, 25-35 years, 35-45 years, 45-55 years, and 55+ years. The variable *education* was found to be statistically significant at the 0.000 level. This variable measured the participants' education in the categories of no high school, high school/GED, some college, two year college, four year college, Master Degree, Doctoral Degree, professional degree JD, MD. The variable *income* was found to be statistically significant at the 0.037 level. This variable measured the participants' income in the categories of 20k or less, 20k – 30k, 30k – 40k, 40k – 50k, 50k +. Finally, the variable *marital* was not found to be statistically significant at the 0.316 level. This variable measured the participants' marital status in the categories of single, married, separated, divorced, widowed. See Appendix B for the results of the multivariate (MANOVA) analysis.

MANOVA Analysis of the Effects of Demographic Characteristics for South Florida OEP-35 Airports.
The MANOVA analysis identified the effects of the covariates in the model for South Florida OEP-35 airports. The variable *race* was the only variable found to be statistically significant at the 0.002 level (see Appendix C).

Analysis and Evaluation of Findings

MANOVA analyses supported the alternative hypothesis that demographics characteristics are significant factors in the airport driver training that employees receive at OEP-35 airports.

Qualitative Comments from Survey

Qualitative comments were grouped by common threads as follows:

1. The typical comments with regard to the most favorable aspects of training centered on computer-based technology and the ability of the training to show the types of signs used in the airfield environment as follows: "The AAAE system is an excellent computer based training system that is user friendly." and "The use of PowerPoint to show signs locations at MIA provides a clear understanding of what they mean" and "overheads used where great."

2. Typical comments with regard to the least valuable aspects of the training centered on the lack of staffing, standardization, funding, and technology as follows: "We need full time trainers" and "Driver training programs are different for airports statewide including resources to maintain or increase operator knowledge. Staffing, training and technology are limited based on fiscal funding." Finally, another common response was "We need more money allocated to driver training programs. Training programs have been historically cut in favor of construction programs."

3. Typical comments with regard to what improvements could be made centered mostly on technology and regulation as follows: "More use of computer interactive software would make it much easier to provide a better environmental awareness than when listening to a trainer" and "movement area licenses should be regulated and monitored by the FAA to uniformly address incursions and compliance from a common FAA standard."

4. Typical comments with regard to obstacles that stand in the way of knowledge and skills learned centered on diversity of the workforce issues at the South Florida airports as follows: "We have a variety of different languages and English is the only language used in aviation; we should use Spanish as well." and "English is my second language." and "There is a vast diversity of the people that take the training, and for many English is their second language." Finally, one common problem was "it is difficult to understand air traffic control instructions."

Conclusions

The MANOVA analysis supported the alternative hypothesis that demographic characteristics are significant factors in the airport movement area driver training that employee receive at OEP-35 airports.

FAA efforts to date have not focused on the study of demographic characteristics associated with airport driver training. All the airports responding to the survey offer primary and recurrent training on a yearly basis. There are two fundamental precepts that are essential for a successful approach to airport driver training; (a) training is not a one-time event, and (b) the most effective way to teach vehicle drivers safety on movement areas is to simulate the actual environment they work in on a reoccurring basis. Training at the OEP-35 airports satisfies these two fundamental precepts.

There was no surprise that education, age, and incomes were found to be statistically significant demographic characteristics at all OEP-35 US airports. Although many non-movement area employees change jobs frequently due to low wage rates, this is not typical for employees licensed to drive on the movement areas. Airport employees that operate ground vehicles in and onto the movement areas of airports typically have a professional certification or a college degree (i.e. electricians, firefighters, airport operations personnel, etc.). One explanation for these demographics characteristics being significant for movement area drivers is that the longer employees work in a given job the more annual recurrent driver training they receive. As the years past, the employee's age and earnings increase. As a result, the three demographics of education, age, and income appear to be

interrelated and may support the conclusion that many of those employees authorized to drive onto movement areas of airports have stable employment records and receive annual recurrent training on a regular basis. As expected, highly educated employees with movement area driving privileges are likely to hold higher positions within their respective company, are older, and earn more income.

With regard to *race* being identified as the only significant variable at OEP-35 airports in South Florida, lack of understanding ATC communications may prove to be the primary issue related to driver training deficiencies in this geographic region. This is highlighted by US Census Bureau data that 74.6% of the language spoken at home in South Florida is other than English, while the US norm is 17.9% (U.S. Census Bureau, 2007). From the airport driver training perspective, communication or the use of consistent terminology is a primary concern.

This conclusion is supported by the qualitative responses, which indicated that English is a second language for many South Florida employees with airport movement area driving privileges. Since English is the industry adopted language for aviation operations worldwide, this finding is problematic. This is evidenced by Clarke (2002) who stated, "The use of consistent terminology (in ATC communication) is recommended for all involved" (p. 14).

Recommendations

Since the data suggested that there is potential to increase knowledge through annual recurrent training, all U. S. air carrier airports should be required to provide recurrent airfield movement driver training on an annual basis. As a minimum, FAA advisory Circular AC 150/5340-20 should be amended. Concerning airport driver-training programs, the circular states "This curriculum should include initial, and may include recurrent and/or remedial, instruction of employees, tenants, contractors, and users who have access to the airside of the airport" (FAA, 2002a, ¶ 6). This sentence should be amended as follows: *This curriculum should include initial and recurrent instruction of employees, tenants, contractors, and users who have access to the airside of the airport.*

Additionally, it is recommended that the Federal Aviation Administration should mandate that all U.S. air carrier airports be required to provide annual recurrent driver training for those employees licensed to operate vehicles into and onto movement areas. This initiative should also be added to the *FAA Runway Safety Blueprint 2002-2004* and implemented through an amendment of Federal Aviation Regulation Part 139. Subsequent to the later action, an amendment to Federal Aviation Regulation Part 139 should require airport operators to address recurrent airport driver training in their Airport Certification Manual.

Certain regions (like South Florida) have racial and ethnic differences, which leads to communications barriers that are not experienced in other regions of the United States. This may mean that educational materials need to be translated

into other languages in order for the materials to be delivered effectively, or that English competency test should be required by FAA regulation before any employee may seek airport movement area driving privileges.

References

Aczel, A. D., & Sounderpandian, J. (2002). *Business statistics* (5th ed.). New York: McGraw-Hill Irwin.

American Association of Airport Executives (AAAE). (2005, February 17). *Interactive employee training (IET)*. [Brochure]. Retrieved March 26, 2005, from http://www.aaae.org

Baharestani, H. H. (2005). Relationships among Computer-Based Instruction and Reasoning Ability on Science Students. *International Journal of Applied Aviation Studies 5*(2), 336-347

Beckman, W. (2000). A comparison of the effectiveness of PC-based aviation training devices and conventional flight training devices for instrument flight training. *The Journal of Aviation/Aerospace Education and Research, 9*(2), 27-33.

Bledsoe, M. D. (1999). Correlations in Kirkpatrick's training evaluation model. *Dissertation Abstracts International*, 60/01, 54. (UMI No. 9937553)

Clarke, B. (2002). *Runway incursions*. New York: McGraw-Hill.

Federal Aviation Administration (FAA). (2004, August). *FAA runway safety report FY 2000 - FY 2003*. Retrieved March 28, 2005, from http://www.faa.gov

Federal Aviation Administration (FAA). (2002a, July). *Runway safety blueprint 2002-2004*. Retrieved March 28, 2005, from http://www.faa.gov

Federal Aviation Administration (FAA). (2002b, June 21). *AC No: 150/5210-20 - Ground Vehicle Operations on Airports*. Retrieved April 12, 2005, from http://www.faa.gov

Filipczak, B. (1996). Endangered! The nature of computer interactivity. *Training, 33*(7), 52-58. Retrieved March 30, 2005, from ProQuest database.

Finney, M. I. (2000, March/April). Getting employee training up to speed. *Airport Magazine*, 14-21.

Kirkpatrick, D. (1983). Four Steps to Measuring Training Effectiveness. *Personnel Administrator, 28*(11) 19-25.

Kirkpatrick, D. (1996). Revisiting Kirkpatrick's four-level model. *Training and Development, 50*(1). 54-59.

Nanney, B. A. (n.d.). Distance education. Retrieved March 27, 2005, from http://www2.gsu.edu

Norusis, M. J. (2003). *SPSS--Statistical Procedures Companion* (12.0 ed.). Upper Saddle River, NJ: Prentice Hall.

Ragan, P. H. (1997). Aviation english: An introduction. *The Journal of Aviation/Aerospace Education and Research, 7*(2), 25-34.

Rankin, W. B. (1994). *Runway incursions: A censorious examination of runway incursions and the federal aviation administration's runway incursion program.* Unpublished master's thesis, Embry-Riddle Aeronautical University, Daytona Beach, Florida.

Rohland, P. (1996). New techniques spice up corporate training. *Eastern Pennsylvania Business Journal 7*(13), 6-19. Retrieved March 30, 2005, from ProQuest database.

Tidler, K. L. (1999). Evaluation of continuing medical education using Kirkpatrick's four levels of evaluation. *Dissertation Abstracts International*, 61/01, 59. (UMI No. 9958848)

University of California, Los Angeles (2005). *UCLA calculators.* Retrieved January 26, 2006, from http://calculators.stat.ucla.edu

U.S. Census Bureau (2007, May 7). U*S Census Bureau.* Retrieved July 25, 2007, from http://quickfacts.census.gov

Wells, A. T., & Rodriques, C. C. (2004). *Commercial aviation safety* (4th ed.). New York: McGraw-Hill.

Driver Training Effectiveness Survey

Name of Airport_____**Date**_____

Guidelines: Read each of the following questions and circle the number that most appropriately represents your opinion. For the qualitative questions on the final page, please clearly print your responses in the space provided. N/A is not applicable.

LEARNING OBJECTIVES MET	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	N/A
When you completed the Driver training, you were able to:						
1. Understand the meaning of runway/taxiway signs, surface markings and lighting	1	2	3	4	5	N/A
2. Understand ATC language, clearances, instructions, and light signals	1	2	3	4	5	N/A
3. Understand Notice to Airmen (NOTAMS)	1	2	3	4	5	N/A
4. Understand and use airport diagrams to navigate on the surface	1	2	3	4	5	N/A
5. Use of new surface navigation technologies	1	2	3	4	5	N/A
6. Understand FAR Part 139 and Advisory Circular Safety Regulations	1	2	3	4	5	N/A
KNOWLEDGE INCREASE	None	Basic	Good	Sound	Expert	
7. My average level of knowledge and skill on movement area driving before completing driver training was:	1	2	3	4	5	
8. My average level of knowledge and skill on movement area driving after completing driver training was:	1	2	3	4	5	
ON-THE-JOB CONFIDENCE	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	N/A
After I completed the driver training, I was confident I would be able to:						
9. Understand the meaning of runway/taxiway signs, surface markings and lighting	1	2	3	4	5	N/A

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	N/A
10. Understand ATC language, clearances, instructions, and light signals	1	2	3	4	5	N/A
11. Understand Notice to Airmen (NOTAMS)	1	2	3	4	5	N/A
12. Understand and use airport diagrams to navigate on the surface	1	2	3	4	5	N/A
13. Use of new surface navigation technologies	1	2	3	4	5	N/A
14. Relate FAR Part 139 and Advisory Circular Safety Regulations to driving on the airport surface	1	2	3	4	5	N/A

EFFECTIVENESS OF MATERIALS & METHODS	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	N/A
15. The driver-training program was clear and easy to follow.	1	2	3	4	5	N/A
16. The materials used provided enough information to facilitate the learning objectives	1	2	3	4	5	N/A
17. The overheads used are clear and easy to follow	1	2	3	4	5	N/A
18. The driver training video used is successful in illustrating the learning objectives	1	2	3	4	5	N/A
19. The exercises effectively cover the learning objectives	1	2	3	4	5	N/A

Type of Training

20. This is traditional training (not computer interactive based training) Yes No (Circle)
21. This is AAAE interactive computer-based training? Yes No (Circle)

Demographic Information
22. Race   White – 1   African-American – 2   Hispanic – 3   Asian-Pacific Islander – 4
Native American – 5   (Circle One)
23. Age   18yrs -25yrs – 1   25yrs – 35yrs – 2   35yrs – 45yrs – 3   45yrs – 55yrs – 4
55+yrs – 5   (Circle One)
24. Education   No High School – 1   High School/GED – 2   Some College – 3
        2yr College Degree – 4   4yr College Degree – 5   Master Degree – 6
        Doctoral Degree – 7   Professional Degree JD, MD – 8   (Circle One)
25. Income   20k or less - 1   20k – 30k – 2   30k – 40k – 3   40k – 50k – 4
        50k+ - 5   (Circle One)
26. Marital   Single – 1   Married – 2   Separated – 3   Divorced – 4   Widowed – 5
                      (Circle One)

Comments
1. What did you find **most** valuable about your driver training program? Please indicate why.
2. What did you find **least** valuable about your driver program? Please indicate why.
3. What improvements could be made to make your driver training program more effective?
4. The obstacles that stand in the way of the successful application of the knowledge and skills
   learned in this program are:
5. Overall Comments:

Appendix B

Multivariate Analysis of Variance Results for Significance of Demographic Variables for all OEP-35 US Towered Airports

Effect		Value	F	Hypothesis df	Error df	Sig.
Race	Pillai's Trace	.024	1.421(a)	4.000	227.000	.228
	Wilks' Lambda	.976	1.421(a)	4.000	227.000	.228
	Hotelling's Trace	.025	1.421(a)	4.000	227.000	.228
	Roy's Largest Root	.025	1.421(a)	4.000	227.000	.228
Age	Pillai's Trace	.095	5.965(a)	4.000	227.000	.000
	Wilks' Lambda	.905	5.965(a)	4.000	227.000	.000
	Hotelling's Trace	.105	5.965(a)	4.000	227.000	.000
	Roy's Largest Root	.105	5.965(a)	4.000	227.000	.000
Education	Pillai's Trace	.130	8.503(a)	4.000	227.000	.000
	Wilks' Lambda	.870	8.503(a)	4.000	227.000	.000
	Hotelling's Trace	.150	8.503(a)	4.000	227.000	.000
	Roy's Largest Root	.150	8.503(a)	4.000	227.000	.000
Income	Pillai's Trace	.044	2.594(a)	4.000	227.000	.037
	Wilks' Lambda	.956	2.594(a)	4.000	227.000	.037
	Hotelling's Trace	.046	2.594(a)	4.000	227.000	.037
	Roy's Largest Root	.046	2.594(a)	4.000	227.000	.037
Marital	Pillai's Trace	.021	1.191(a)	4.000	227.000	.316
	Wilks' Lambda	.979	1.191(a)	4.000	227.000	.316
	Hotelling's Trace	.021	1.191(a)	4.000	227.000	.316
	Roy's Largest Root	.021	1.191(a)	4.000	227.000	.316
Method	Pillai's Trace	.243	18.206(a)	4.000	227.000	.000
	Wilks' Lambda	.757	18.206(a)	4.000	227.000	.000
	Hotelling's Trace	.321	18.206(a)	4.000	227.000	.000
	Roy's Largest Root	.321	18.206(a)	4.000	227.000	.000

a Exact statistic
b Design: Intercept+Race+Age+Education+Income+Martial+Method

Multivariate Analysis of Variance Results for Significance of Demographic Variables for South Florida OEP-35 US Towered Airports

Effect		Value	F	Hypothesis df	Error df	Sig.
Race	Pillai's Trace	.159	10.436(a)	1.000	55.000	.002
	Wilks' Lambda	.841	10.436(a)	1.000	55.000	.002
	Hotelling's Trace	.190	10.436(a)	1.000	55.000	.002
	Roy's Largest Root	.190	10.436(a)	1.000	55.000	.002
Age	Pillai's Trace	.024	1.356(a)	1.000	55.000	.249
	Wilks' Lambda	.976	1.356(a)	1.000	55.000	.249
	Hotelling's Trace	.025	1.356(a)	1.000	55.000	.249
	Roy's Largest Root	.025	1.356(a)	1.000	55.000	.249
Education	Pillai's Trace	.015	.836(a)	1.000	55.000	.365
	Wilks' Lambda	.985	.836(a)	1.000	55.000	.365
	Hotelling's Trace	.015	.836(a)	1.000	55.000	.365
	Roy's Largest Root	.015	.836(a)	1.000	55.000	.365
Income	Pillai's Trace	.027	1.518(a)	1.000	55.000	.223
	Wilks' Lambda	.973	1.518(a)	1.000	55.000	.223
	Hotelling's Trace	.028	1.518(a)	1.000	55.000	.223
	Roy's Largest Root	.028	1.518(a)	1.000	55.000	.223
Marital	Pillai's Trace	.011	.608(a)	1.000	55.000	.439
	Wilks' Lambda	.989	.608(a)	1.000	55.000	.439
	Hotelling's Trace	.011	.608(a)	1.000	55.000	.439
	Roy's Largest Root	.011	.608(a)	1.000	55.000	.439

Pilot Perceptions
On Using a Ballistic Parachute System

Arlynn McMahon

Aero-Tech, Inc.

Abstract

This research studied pilots' perceptions about the use of a ballistic parachute with 1,003 valid survey respondents. Overall, pilots felt that an aircraft equipped with a parachute was safer than one without; however, flight experience affected pilot opinions. Flight instructors certificated for less then 2 years imagined the parachute would lower them gently to the ground. Research determined that new pilots might favor flight schools providing parachutes on training aircraft. However, without training-scenarios to teach aeronautical decision-making regarding deployment, flight instructors may use the parachute indiscriminately. The study revealed that pilots made decisions differently when considering flights in a parachute equipped aircraft during four scenarios. Research included factors that may influence pilot opinions, the Risk Homeostasis Theory (RHT), and recommendations.

In 1998, the Cirrus Design Corporation added a whole-airplane, ballistic parachute system to their SR20 airplane (Cirrus, 2007). Then, in 2004, the FAA approved the *Light Sport Aircraft Rule* (FAA, 2004),[1] inaugurating the new Light Sport Aircraft Category.[2] Many Light Sport Aircraft (LSA) manufacturers offer a ballistic parachute system as an option on new aircraft. The parachute concept continues to gain popularity and traditional aircraft such as the Cessna 152/172/182 now offer retrofit options (BRS FAQs).

Currently there is little information on pilots' opinions of the safety and usefulness of the ballistic parachute. Do pilots perceive an aircraft equipped with a parachute as being safer than one without? Are these pilots willing to invest time and money to learn to use the equipment? Will pilots make decisions differently when flying an aircraft with a ballistic parachute?

As with any new technology, understanding how the ballistic parachute system works is instrumental to understanding how the pilot interfaces with it and its limitations. The parachute system uses a three-point harness and Kevlar webbing, which connects the parachute to the airplane (POH, 2000 p7-68). The pilot

deploys the system by pulling a red activation handle, launching a rocket. In about two seconds, the parachute inflates with a canopy of about 2,400 SF (Goyer, 2004). The effect of deceleration on the pilot and passengers is about 3 G's. Descent under the parachute is about 1,800 feet per minute (POH, 2000). Once the parachute is deployed, the pilot has no control and no steering. The pilot is required to shut down the engine, even if it is operational. "Once the red handle is pulled, everybody onboard is along for the ride" (Goyer, 2004, p. 44). Touchdown is not particularly gentle – as much as 15Gs, nearly all of which are vertical (Wallace 1994).

The system has two important limitations for successful deployment: (1) altitude/attitude/airspeed and (2) the pilot's decision-making process. The recommended deployment altitude is 2,000' AGL or higher with wings level and a minimum possible airspeed. The maximum demonstrated deployment speed is 133 KIAS (POH, 2000, p. 10-3). Pilot judgment is important. The pilot must decide when to activate and under what conditions to deploy the system (POH, 2000 p10-3).

Literature Review

A literature review revealed little data on the use of the ballistic parachute system. Information was not found in Federal Aviation Administration (FAA) Advisory Circulars, Aviation Regulations, or in the Aeronautical Information Manual (AIM). Primary sources of information are the aircraft Pilot Operating Handbook or the Approved Flight Manual (POH) of those airplanes having a parachute system installed and the information provided from parachute system manufacturers. NTSB reports may be the best available source of information on when pilots chose to deploy the ballistic parachute during emergencies and the results of those deployments.

Research Objective

Due to the lack of practical information regarding the ballistic parachute system, the survey was an effort to obtain information from pilots about their opinions regarding the equipment. The primary objective of the survey was to address:

- How do pilots perceive the ballistic parachute system's use and utility?
- Given an opportunity to fly an aircraft with a ballistic parachute system, would pilots perceive that they would make in-flight decisions differently?"
- Does pilot experience influence their opinions about having a ballistic parachute system?

Method

The Survey

Survey questions assessed pilot opinions regarding the ballistic parachute, provided scenarios when the pilot might deploy the parachute and collected the respondent demographics. Each survey question was first administered to in-house flight instructors and customers, testing for clarity and understanding before survey distribution. Then the survey was published and made available via the

Internet. The survey contained 13, 5-point Likert-scale multiple-choice questions in determining pilot opinions and demographics, two yes/no questions, plus an area for respondents to provide comments (See Appendix A1). Survey data was extracted into categories according to pilot age, experience, and certificate level. Data was further classified based on whether the respondents owned a parachute-equipped aircraft. Data was cross-sectioned to look for overall trends and statistical differences among the groups. The responses were managed on the secure server of the online survey tool, which provided reports and statistical analysis. Survey results were interpreted using QuestionPro.com® statistical analysis software.

Survey Recruitment

Recruiting participants into the survey consisted of four main avenues. First, in an attempt to reach the average pilot presently flying at General Aviation flight schools, the researcher sent e-mails to flight school managers announcing the survey. These schools were identified from BE A PILOT™, a national effort, whose purpose is to match names and contact information of prospective pilots to participating flight schools. These schools encouraged their instructors and students to participate in the study. In addition to the flight schools at BE A PILOT™, the researcher also sent emails to prospective pilots from BE A PILOT™. It was hoped that this group would provide insight into the type of equipment that prospective flight training customers might expect on aircraft. Third, Ballistic Recovery Systems, Inc. (BRS[3]) a parachute system manufacturer, invited owners of aircraft equipped with their system to participate. This group was expected to provide viewpoint and insight into the technology and its use. The survey tool was able to segment this group from the general results. Fourth, the Aircraft Owners and Pilot's Association (AOPA) invited their general membership into the survey in an e-mail newsletter.

Results

Respondents

During the seven days that the online survey was available, 1,010 participants started the survey and 995 completed it. The analysis included data only from those respondents who completed all or better than 80% of the survey questions (n=1,003). The survey was discontinued after seven days because of the unexpected large amount of data collected during that period. The average time required for respondents to complete the survey was 6 minutes 38 seconds.

The participation rate from flight schools was nearly 100%, while participation from prospective pilots was less than 10%. Another 13% of the respondents (n=117) were owners of aircraft equipped with a parachute system (it is unknown how many owners were invited by BRS to participate). Approximately 46% of respondents originated from the AOPA general membership.

Survey respondents represented each of the 50 U.S. states, all providences of Canada, and 32 other countries. Table 1 identifies the respondents by age group.

Table 1
Respondents by Age Group

Age	%
Less than 35 years old	30%
35 - 50 years old	37%
51 - 65 years old	28%
66 years and older	4%

About 13% (n=117) of respondents owned an aircraft with a parachute system and about 24% (n=234) of respondents reported having a close acquaintance that had been injured while flying. Table 2 summarizes the number included in the research data by pilot certificate.

Table 2
Proportion of Respondents in each Category

	Number of Respondents	As a Percent of Total Respondents
Not yet a pilot, but plan to start pilot training soon. (Not-Yet)	37	3.71%
Student Pilot (Student)	202	19.02%
Sport Pilot (Sport)	13	1.39%
Private Pilot (Private)	438	41.10%
Commercial without Instructor (Com)	73	7.42%
Airline Transport Pilot (ATP)	22	5.57%
Certificated Flight Instructor (CFI)	229	21.43%
Passenger Only (PAX)	4	0.37%
Total	1062	100.00%

Note: The total 1,062 includes 67 responders having multiple certificates. Primarily these included Airline Transport Pilot with Certificated Flight Instructor and Private Pilot or Sport Pilot with Flight Instructor. (Under the new FAA Sport Pilot Rule, an instrument rating and commercial pilot certificate are not prerequisite for a flight instructor certificate.)

Perspectives on Parachute Safety
Overall, nearly 77% (n=698) of all respondents indicated their opinion that an airplane with a parachute is safer than an airplane without.

Across all pilot certification categories, the majority of respondents *strongly disagreed* that they were more likely to fly into marginal conditions with a para-chute system. Flight instructors certificated for less than two years indicated the highest responses at 80.8% (n=117). Additionally, the majority of respondents indicated they would not make in-flight aeronautical decisions differently if flying an aircraft equipped with a parachute system. However, when offered four specific, uncomfortable, but survivable scenarios, many pilots indicated they would deploy the parachute instead of using pilot skill: (See Appendix A2.)

In Scenario #1: Given an engine failure, 67% (n=672) of respondents *agreed* or *strongly agreed* to activate the parachute rather than to attempt an emergency

landing in hostile terrain. The survey did not attempt to define "hostile" terrain, choosing instead to allow each pilot to define it for him/herself.

In Scenario #2: Given a complete electrical failure at night, 50 miles away from an airport, 35% (n=353) of respondents *agreed* or *strongly agreed* to activate the parachute rather than to land at the nearest airport. In most general aviation airplanes, the time required to fly 50 miles is about 20 minutes.

In Scenario #3: Given the structural failure of an aileron, 62% (N=618) *agreed* or *strongly agreed* to activate the parachute rather than attempting a landing at the nearest airport. An aileron-failure limits the controllability of flying the airplane, specifically in making turns, but the airplane would normally be controllable to a suitable landing destination.

In Scenario #4: Given an unintentional spin at 1,500 AGL, 55% (n=547) indicated activating the parachute rather than attempting a recovery. This altitude represents the outer limit that allows time for the parachute to open fully. It also represents a marginal altitude, given average pilot-skill that would probably be successful in recovering from an unintentional spin.

In fact, some respondents *strongly agreed* to pull the parachute on all four scenarios. Those with flight instructor certificates offered an interesting polarity: 1% of flight instructors, being certificated for more than five years, (n=106) indicated pulling the parachute on all four scenarios. Alternately, 100% of the flight instructors certificated for less than two years (n=117) indicated pulling the parachute on all four scenarios. At the same time, the vast majority of flight instructors certificated for less than two years rated themselves as *normal* to *very confident* in routine flight skills.

Flight instructors offered enough disparity so that new-instructors, those with less than two years experience, were segregated from the more experienced and compared. The majority of new-instructors were from Florida (n=46), Kentucky, (n=27) Ohio (n=24), Oklahoma (n= 12) and California (n=6). Looking closer at this group, none indicated owning a parachute system. None reported a close acquaintance ever being hurt while flying an aircraft.

Delivered Gently to the Ground. Of particular importance to a pilot's perspective regarding the use of a parachute system is the pilot's expectation of parachute performance. The survey attempted to test the expectation of parachute performance by having respondents choose between four possible outcomes after the parachute was pulled. The survey choices suggested the parachute would (1) "Deliver me gently to the ground," (2) "Deliver me safely to the ground," (3) "Deliver me safely to the ground although the airplane may be destroyed," or (4) "Likely seriously injure occupants and destroy the airplane."   Of the 157 respondents who indicated that the parachute would "deliver me gently to the ground," 137 also chose to pull the parachute on all four scenarios. None of the respondents in the "gently" group owned a parachute system. Three reported

having an acquaintance hurt while flying an aircraft. In the "gently" group, 138 respondents had been certificated for less than two years while six were not a pilot and five experienced pilots had been certificated for more than ten years. One held an airline transport pilot certificate. The vast majority of pilots in the "gently" group were from Florida (n=122). All flight instructors certificated for less than two years (n=117) indicated that the parachute would deliver them "gently."

Interestingly, while pilots did not see themselves making decisions differently while flying with a parachute, the majority felt that *other* pilots would. Overall, 58.61% (n=585) of total respondents indicated that other pilot were more likely to fly into marginal conditions with a parachute system.

Will Pilots Pay To Have a Parachute? The research tested the willingness of respondents to pay an additional $5 per hour to rent a parachute-equipped airplane. The survey question was not to suggest that a flight school would or should increase rental rates by $5 to pay for the parachute. Rather, at the researcher's flight school, aircraft rental insurance is available to pilots at the rate of $5 per hour. The question tested if pilots would invest an equal amount to have a parachute, as they presently invest to have renter's insurance. The majority of the Not-Yet group indicated a willingness to pay the higher price; certificated sport pilots were evenly divided between willing to pay a higher price and not. However, overall, pilots *strongly disagreed* to pay a higher price for a parachute-equipped airplane. According to Respondent #439, "I don't think the parachute should cost $5 bucks more. Its safety gear " #552 commented, "I'm not paying $1 extra to have it and I'm not spending 1 hour to learn to use it." However, #816 disagreed, "I would pay extra to have one in a heartbeat. When you run out of options, you'll wish you had one."

The research tested if pilots would spend five hours learning to use the parachute system. Most flight schools require a "checkout" wherein pilots are asked to demonstrate or to learn the skills needed to operate the airplane and/or its installed equipment. The question was not to suggest that five hours would be required to learn the system. Rather it was to test if respondents were willing to invest the same amount of time learning this equipment as they would spend to learn other new aircraft technologies, as one example, the Garmin 1000 glass cockpit. A slight majority of respondents across all categories agreed to spend the time in training to use the parachute with the exception of new flight instructors. New flight instructors did not follow the trend with 94.78% considered training on the system as unnecessary.

Several owners of parachute-equipped aircraft commented on the need for standardized training and recommendations for deployment. Respondent #820 wrote, "I believe there should be formal FAA recommendations and maybe an Advisory Circular. This would encourage flight schools to utilize standardized procedures during training." Respondent #642 agreed, "I don't think you can train a pilot on the BRS by reading something, scenario training puts the pilot in a situation where he can actually imagine how scary it would be to use it."

The research examined customer appreciation for the parachute by preferring to rent at a flight school offering such equipment. The Not-Yet-A-Pilot category and CFI's-certificated- less-than-2-years were overwhelmingly in favor of schools offering this equipment. Those in the Student category only mildly favored these flight schools. While responders holding a Private pilot certificate did not.

Pertaining To Passengers

Research showed that passengers are an important consideration in choosing to fly an airplane with a parachute. Overall, 49% (n=593) of total respondents indicated that their first consideration for renting an airplane with a parachute was more concerning the safety of family or passengers than for the pilot's personal safety.

Survey Respondent #107 shares, "…My main interest would be the ability of my passengers to survive…" #246 agreed, "…The passenger confidence is a big part having a parachute on board, however when I asked my wife which would you give up weather on board or parachute her response was, " that's a tough one."

However, #414 disagreed, "I prefer to take my chances with my skill than to be a passenger along for the ride on a descent to hell. Why anyone would pay for this is beyond me. They should be outlawed on airplanes. I know they save lives on ultralights but those guys are crazy anyway." #538 said, "Why would I make decisions based on the safety of my passengers. If I am safe then my passengers will be safe. I make decisions based on what's safe for me and therefore my passengers will reap the rewards."

Hostile Terrain

The researcher considered the possibility that terrain may affect the opinion of respondents. The survey tool provided a physical location of each respondent, usually city, state, and nation. Analysis divided the U.S. into either hostile terrain or non-hostile terrain, to segregate the respondents and compare data.

Most states have a combination of terrain, and any terrain could be considered hostile in the deepfreeze of winter or during summer's high noon. Still, the research attempted to segregate those having the majority of terrain with limited, survivable, off-airport landing sites in the event of an emergency. States considered in the hostile terrain category were AK, AZ, CA, CO, HI, ID, MT, NM, NV, OR, SD, UT, WA, WV, and WY, resulting in 212 respondents. All other states were included in non-hostile terrain, and resulted in 680 respondents. Comparing hostile and non-hostile terrain respondents yielded no significant differences.

From Parachute Owners

Information pertaining to use of the parachute in this research stemmed largely from those respondents invited into the survey from BRS. Those who own a BRS system are primarily Cirrus aircraft owners. The Cirrus Owners and Pilot's

Association (COPA) employs a formal training program to educate its owners. A large segment of this training is dedicated to educating owners on the limitations of the parachute and proper scenarios for deployment. The opinions of BRS-owner respondents (n=107) differed from the general respondent-population in that 75.70% *strongly agreed* that an airplane equipped with a parachute is safer than one without. The majority, (59.8%) would give preference to a flight school offering an airplane with a parachute, and 58.8% would pay an additional $5 per hour to rent an aircraft with a parachute. The vast majority of BRS owners were *neutral* in deciding if other pilots are more likely to fly into marginal but legal conditions with a parachute system.

Survey comments from BRS owners reflected insights from their training and several pilots saw the need for a "parachute simulator" allowing pilots insight as to what to anticipate after pulling the parachute. Survey Responder #39 indicated, "Whole parachute systems are perceived by many as a safety panacea. The reality is that a majority of accidents occurs due to poor pilot decision-making. Training in aviation risk management is one of the keys to reducing GA accidents." Respondent #390 said, "The plane does not make a safe flight, the pilot's decisions do." According to Respondent #966, "… you don't need 5 hours time to learn how to operate the system. However, you could easily spend 1,005 hours studying critical decision making, related to the chute." Respondent #823 offers, "Don't feel as comfortable any more when flying airplanes or riding in airplanes that don't have a parachute system. You get spoiled pretty quickly. I just feel better knowing it's there." According to #417, "I have not spent good money just to take a parachute ride. I would rather be in control. I would rather control my destiny. If anyone tells me to pull the chord, I'll stop flying... Every photo should have a DON'T TRY THIS AT HOME disclaimer. What about the survivors with broken spinal chords and long-term foot and leg mangelments? ..."   #665 comments, "…Some have called the red handle of the parachute system a Transfer of Ownership Device. If I were to attempt an off airport landing I would almost for sure do at least $5,000 worth of damage. This is my insurance deductible. I would also quite possibly sustain some injury. For the same cost I can total the airplane in a parachute landing almost for sure receive no injury and still owe only $5,000!"

Discussion

Respondent # 932 reflects on CHUTE HAPPENS™. It's a playful colloquialism and marketing slogan, playing on the notion that bad-things do happen and that having a parachute would assure a happy landing. Reflecting on the number of respondents with an expectation of the parachute to deliver them safely or gently to the ground, it's not surprising that the majority would consider an aircraft equipped with a parachute to be safer than one without.

Media Influence on Pilot Opinions

Pilots with first hand knowledge of a ballistic parachute system may be educated on its performance; however, pilots without direct exposure may have their opinions affected by marketing and media. Mainstream information regarding the parachute system is at times misleading. This may give readers a false sense of parachute performance expectations that may carry over into the cockpit and the pilot's opinion on use of the parachute. As an example: From *Mechanical Engi-*

neering Magazine, "Every year about 1,000 people die in GA accidents and BRS claims that 60 percent of them could be saved using whole-aircraft parachute recovery systems" (Ashley, 1996). In *Omni*, "even when deployed from altitudes as low as 300 feet, the nose gear generally suffers the only damage" (Nobble, 1994). *Time Magazine* reported, "When the engine conks, a new system lets pilots and passengers stay aboard and float to safety... The plane drifts to earth for a safe, if still somewhat bumpy landing" (Gorman, 1993). Similarly, *New Scientist Magazine* reported, "This month five more people owe their lives to rocket-assisted rescue parachutes"(Byrne, 2004). This research considered "years" of pilot flying experience in lieu of pilot flight-hours because years may reflect the length of time that a pilot's opinion could be influenced by the general media.

Risk Homeostasis Theory

A subset of the survey questions tested pilots' possible decisions and actions while flying with a parachute system. The majority of respondents indicated that *they* would not fly into marginal conditions or make decisions differently when flying an aircraft equipped with a parachute system, but that *other* pilots would.

"Give me a ladder that is twice as stable, and I will climb it twice as high." This quote from Gerald Wilde (Wilde, 2001, p. 158) illustrates his Risk Homeostasis Theory (RHT). It is just a theory, but it portrays how people may change behavior in response to the addition of an added safety device (Wilde, 2001). Consider the possibility that the presence of a parachute system may encourage pilots to make decisions or to operate an airplane in a riskier manner than they would without having the parachute. The RHT theory says that people naturally have a level of risk they accept, tolerate, desire, or choose. Given additional safety equipment to make a task safer, people will engage in riskier behavior, so that, overall their level of risk remains constant.

Respondent #627 says, "I know pilots that have a false sense of security that lends the tendency to fly into conditions that they should not fly into. This 'security blanket' historically has shown that while they have saved lives, they have also gotten people out of situations they should have never gotten themselves into. Additionally, the insurance companies recognize this...compare rates on a parachute equipped Cirrus and a non-parachuted Columbia. You'll see, despite the parachute, the rates are higher on a Cirrus."

Respondent # 358 adds, "...the best safety system can be defeated by a motivated idiot. Pilots who fly into IMC with a parachute equipped plane are no better than drivers who speed through icy mountain passes, confident that antilock brakes will protect them from their foolishness."

Respondent #364 agrees, "On board parachutes would detract from normal necessary training time, add unnecessary expense to the airplane, and I think it would give a false sense or bravado to certain pilots. I have two friends with parachute-planes. They fly in the ice and other weather that no "good" pilot would. They depend on their parachute rather than train and prepare for the possible

aircraft emergency...Anyone worried about the uncontrolled falling of the airplane should use the 100% sure and much less expensive method of protection by simply staying on the ground and not flying at all. Get real, train realistic pilots and leave the others on the ground so we all can be safer."

If the RHT Theory is proven, then the fear of pilots operating aircraft haphazardly as a result of having a parachute could affect the parachute's utility and overall aviation safety.

About Flight Instructors
Of particular concern is the finding that 100% of flight instructor respondents less than 35 years old and having less than two years experience (n=117) expected an airplane equipped with a parachute to deliver the plane and its occupants gently to the ground. This group also indicated pulling the parachute on all four scenarios. This is alarming, given that the flight instructors have such a large influence on the general pilot populations' decisions and practices.

Four generalizations came to mind, which would require additional study: (1) Having invested in their aviation education, new-instructors now expect compensation to fly rather than paying to fly. Thus spending their time or money to learn the system was not their mindset. (2) New-instructors invested in an aviation education to become an airline pilot and do not have a desire to teach or to fly in aircraft normally used in training. (3) The instructor may have had a personal experience with flight training in an airplane in disrepair or questionable airworthiness. Thus, the instructor may feel a parachute would benefit that particular airplane. (4) A new instructor does not yet trust his or her skills and has not had the time to develop the experience needed to confidently handle an abnormal situation.

The following small sampling of comments is from the new-instructor group:
- "I'm a new CFI - it's comforting to know that either I or the student could have the option to pull the canopy if either of us gets afraid or uncomfortable."
- "I'm a new CFI and I don't mind saying that it scares me to fly with students. I don't like it. C-152s are death traps. They should all be fitted out with a parachute."
- "Lots of CFIs get killed while instructing. Parachutes would save lives."
- "What do we really know about how to react when the student makes a mistake?"
- "Students do stupid things. Give me a chute and I'll fly for you."
- "I am a small female and feel very confident in my abilities to fly the airplane. But what if I get a big man in the left seat and I can't get control of the plane back from him when he's screwed up a stall recovery? It would be nice to know that I have a parachute that I can pull and regain control."
- "If I had a parachute I would pull it in a heartbeat - when a student does something crazy, when the plane is doing something that I don't like, when a piece of equipment becomes inop (engine, vacuum, electric, anything!) Why should I take a chance on my skills if I have a chute? That's what it's there for."

- "If I MUST endure these little rust-buckets the least my flight school can do is to put a parachute on it so that I won't get hurt. … But I think it should be standard equipment and not something that I should have to pay extra for it – like seat belts. I also can't imagine spending 5 hours learning how to pull the handle. Duh. Just pull the handle – if you're not smart enough to do that then you shouldn't be flying anyway."
- "Six months ago my CFI was a student pilot. Now I am a CFI teaching students. Who knows how long this blind-leading-the-blind thing has been going on. But if it's to continue then stack the odds in everyone's best interest and spend the bucks to equip the airplane with safety equipment like a parachute. Why should someone get hurt while being paid very little, earning the privilege to work for the airlines? Spend the money. Buy the chute."
- "Flight instructors spend 10 to 12 hours in the plane every day. …Yes, have a chute and be ready to pull it!"
- "Some flight instructors are not very good with their new skills. Some have even fewer skills in supervising student actions and recovering from students upsets. It would be a good idea on flight training planes to help the instructor get control of the plane."
- "Flight training sucks. This is not why I got into aviation. Spending my days with people who just want to fly around is very different than spending my days with people who have goals, career aspirations and are motivated to get to the airlines. Down here at this level yes you positively need parachutes. Unfortunately, people at this level of aviation are so unintelligent that they probably won't pull the chute, anyway."

New-instructors indicated an unwillingness to invest five hours in training on the parachute system. Perhaps they did not feel five hours were necessary to learn the system and therefore indicated their lack of willingness to invest what they felt was an inordinate amount of time/money. According to Respondent #303, "A new flight instructor should not be without a parachute. But I don't think there is any training involved with it. Certainly not five hours. It's a simple system from what I understand." Respondent #770 agrees, "I currently rent a SR20. I think one or two hours training is enough."

Limitations

Asking pilots for their opinions through an Internet survey was a cost effective method that helped to address the research objectives. While a large sampling of opinions was gathered from pilots, the hope is that the respondent group is an accurate sample of the total pilot population. Most respondents did not have direct exposure to a parachute system, relying instead on magazine articles and advertising as a basis for forming their opinions.

Given the significant differences between experienced flight instructors and new-instructors, further research would be valuable in learning why the differences exist. What opportunities are in place to bring these groups together to

share information? Why are some pilots willing to give up their command to ride beneath the canopy on routine failures? Is additional training needed in specifically newly minted flight instructors? Does accelerated pilot training contribute to the pilot's opinions? One disturbing fact uncovered was that flight instructors were unwilling to spend five hours learning to properly use the system. Future studies should allow instructors to quantify precisely how much time they would invest.

An important but unanswered question concerns the effectiveness of the parachute in saving lives. The research was initially planned to include an analysis of NTSB accident/incident reports where the ballistic parachute system was involved. However, searching the NTSB database for keywords such as "BRS" or "parachute" revealed a small sampling of reports and nearly 1/3 of those omitted whether or not the parachute had been deployed. Thus, NTSB data was not included in the research as the sampling was not statistically relevant and the results inconclusive.

Recommendations

1. NTSB accident investigators should continue to report the status of the parachute activation lever and deployment on accident/incident reports so that future research can quantify the reliability of the system. In this way, pilots may form an opinion about the parachute system based on operational effectiveness.
2. FAA and industry safety initiatives should promote deployment scenarios training and materials to pilots flying with parachutes. (i.e., is pulling the parachute the best way to deal with an emergency when the plane is 2,000 feet above the ground, with wings level, flying at less than 133 knots?)
3. FAA policy and regulatory authors must not consider the parachute system as bon-a-fide safety gear until further study can prove it as such.
4. The design of training systems should consider the RHT theory so that future training includes pilot strategies in how to develop low-risk behaviors in addition to managing risks.
5. A solution for more accurate public expectations regarding the parachute must be communicated. Care must be used in marketing materials to more accurately reflect the typical result of the parachute pull. Perhaps a statement similar to that on weight loss advertisement, "Results not typical" should be included on parachute advertisements.
6. Owners of aircraft with ballistic parachutes should be on high alert during flight training, as it appears that owners and their flight instructors may not share like opinions in the pull/no-pull decisions. Owners should consider that the instructor hired to teach in their aircraft may not have been trained in the parachute and that the instructor may not make in-flight emergency decisions consistent with those of the owner.
7. A briefing in handling parachute-emergencies, similar to positive-transfer-of-controls, should be included in pre-takeoff briefings between pilots/instructors when flying parachute-equipped aircraft.
8. Flight schools choosing to train in parachute-equipped aircraft should invest in comprehensive training for staff instructors to redefine their expectations of the parachute and appropriate scenarios for its deployment. Following the lead of the Cirrus Owners and Pilot's Association, the syllabus should be steeped

in scenario training, aeronautical decision making, and risk management. It must also train instructors in how to pass this information to their students.

Conclusions

The research identified several key aspects of operating an aircraft equipped with a parachute. For example, airplanes equipped with ballistic parachutes will be operated differently if for no other reason than pilots now have a choice: pull the parachute or fly the plane. However, the parachute itself becomes an additional element for the pilot to consider while dealing with an already stressful situation. In addition, RHT suggests that pilots may engage in riskier flight conditions because of the sense of security that may accompany having a parachute.

Respondents in general view an aircraft equipped with a parachute as being safer than one without. While the pilot's perceived value of a parachute is high, some pilots will need convincing to pay higher rental rates that may be needed to pay for the system, and new-instructors will need convincing to invest time in training on the system.

Comments from respondents were passionate and at times lengthy, demonstrating the strong feelings – sometimes for and other times against – the parachute concept. In reading the comments, one can sense the quality of new flight instructors and their take on our industry - versus more seasoned pilots. Future research can be built around ways to bring new instructors and their seasoned counterparts closer together in viewpoint.

References

Ashley, S. *(1996, March). Dropping to safety,* Mechanical Engineering, 118, 160

Ballistic Recovery Systems, Inc. (n.d.). *BRS Frequently Asked Questions*. Retrieved on December 17, 2007 from http://www.brsparachutes.com/Default.aspx?TabId=26

Byrne, Gerry. (2004, April 24). Life-saving chutes to rescue jet planes. *New Scientist* Magazine, 2444.

Cirrus (2007). *Cirrus Design Corporation: History*. Retrieved on December 31, 2007 from http://www.cirrusdesign.com/about/history/default.aspx

FAA (2004, September). NPRM: Certification of Aircraft and Airmen for the Operation of Light-Sport Aircraft. NPRM. Federal Aviation Administration. Washington, DC

Gorman, C. (1993, September 20). *A Parachute – But No Jump.* Time Magazine*, 0040781X, Vol. 142, Issue 12.

Goyer, R*. (2004, December). Chute: What the whole-airplane recovery parachute system is doing for safety. *Flying Magazine, 131,* 44-50.

Nobble, G. (1994, October)*. Parachutes for Planes.* Omni Magazine*. 01498711, Vol. 17, Issue 1.

POH. (2000). *Pilot's Operating Handbook and FAA Approved Airplane Flight Manual for the Cirrus Design SR22*

Wallace, L.E. (1994, June). Ballistic recovery system: Lifesaving landings. *Aircraft Owners and Pilots Association (**AOPA**) Pilot, 73,* 87-91.

Wilde, G.J.S. (2001). *Target Risk*. Toronto, Ontario, Canada: PDE Publications,

[1]The FAA created a new rule for the manufacture, certification, operation, and maintenance of light-sport aircraft. This action recognized advances in sport and recreational aviation technology and mended the lack of appropriate regulations for existing aircraft. The effect of this rule was to promote the manufacture of safe and economical certificated aircraft and to allow operation of these aircraft by certificated pilots for sport and recreation. A sport pilot is required to receive at least 20 hours of flight training to be eligible for certification. A sport pilot is limited to flying a Light Sport Aircraft (LSA), during day, below 10,000 MSL, carrying a maximum of one passenger. A sport pilot does not fly in furtherance of business or for hire. A sport pilot is not required to hold a FAA medical certificate.

[2]A Light Sport Aircraft (LSA) is partly defined as having: a maximum takeoff weight of not more than 1,320 pounds for aircraft not intended for operation on water; a maximum airspeed in level flight with maximum continuous power (V_H) of not more than 120 KCAS; a maximum never-exceed speed (V_{NE}) of not more than 120 KCAS; a maximum stalling speed of not more than 45 KCAS; seating for the pilot plus a maximum of one passenger; fixed landing gear; and a single, reciprocating engine. A LSA may be an airplane, glider, balloon, powered parachute, weight-shift-control aircraft, or gyroplane.

[3]Many pilots refer to any aircraft ballistic parachute system as a "BRS." Since care must be taken in defining terms, in this research, BRS refers to the company, Ballistic Recovery Systems, Inc.

Appendix A1 – Overall Survey Results

1. An airplane equipped with a whole-airplane parachute is safer than an airplane without.

Strongly Agree	**384**	**38.48%**
Somewhat Agree	314	31.46%
Neutral	133	13.33%
Somewhat Disagree	99	9.92%
Strongly Disagree	68	6.81%

Mean	2.15
Standard Dev.	1.23
Variance	1.51
Mean Percentile	76.97%

2. When flying an airplane with a parachute system, training on the system is necessary.

Strongly Agree	**742**	**74.13%**
Somewhat Agree	89	8.89%
Neutral	13	1.30%
Somewhat Disagree	15	1.50%
Strongly Disagree	142	14.19%

Mean	1.73
Standard Dev.	1.42
Variance	2.02
Mean Percentile	85.45%

3. I would invest an additional 5 hours of training to learn how and under what conditions to operate the parachute.

Strongly Agree	**380**	**38.23%**
Somewhat Agree	227	22.84%
Neutral	93	9.36%
Somewhat Disagree	103	10.36%
Strongly Disagree	191	19.22%

Mean	2.49
Standard Dev.	1.54
Variance	2.38
Mean Percentile	70.10%

4. I would give preference to a flight school offering training / rental of an airplane with a parachute system.

Strongly Agree	**290**	**29.06%**
Somewhat Agree	209	20.94%
Neutral	209	20.94%
Somewhat Disagree	150	15.03%
Strongly Disagree	140	14.03%

Mean	2.64
Standard Dev.	1.40
Variance	1.96
Mean Percentile	67.19%

5. I would pay an additional $5 per hour to rent an aircraft with a parachute system.

Strongly Agree	234	23.49%
Somewhat Agree	183	18.37%
Neutral	147	14.76%
Somewhat Disagree	156	15.66%
Strongly Disagree	**276**	**27.71%**

Mean	3.06
Standard Dev.	1.55
Variance	2.39
Mean Percentile	58.86%

6. My first consideration for renting an airplane with a parachute is more concerning the safety of my family/passengers than for my personal safety.

Strongly Agree	212	21.20%
Somewhat Agree	**281**	**28.10%**
Neutral	261	26.10%
Somewhat Disagree	141	14.10%
Strongly Disagree	105	10.50%

Mean	2.65
Standard Dev.	1.25
Variance	1.57
Mean Percentile	67.08%

7. I would make in-flight aeronautical decisions differently if flying an aircraft with a parachute system.

Strongly Agree	51	5.12%
Somewhat Agree	105	10.53%
Neutral	116	11.63%
Somewhat Disagree	264	26.48%
Strongly Disagree	**461**	**46.24%**

Mean	3.98
Standard Dev.	1.21
Variance	1.46
Mean Percentile	40.36%

8. I would be more likely to fly into marginal, but legal conditions with a full airplane parachute system.

Strongly Agree	28	2.80%
Somewhat Agree	54	5.41%
Neutral	70	7.01%
Somewhat Disagree	181	18.12%
Strongly Disagree	**666**	**66.67%**

Mean	4.40
Standard Dev.	1.02
Variance	1.04
Mean Percentile	31.91%

9. Others are more likely to fly into marginal, but legal conditions with a full airplane parachute system.

Strongly Agree	272	27.25%
Somewhat Agree	**313**	**31.36%**
Neutral	223	22.34%
Somewhat Disagree	127	12.73%
Strongly Disagree	63	6.31%

Mean	2.39
Standard Dev.	1.19
Variance	1.42
Mean Percentile	72.10%

10. Given an engine failure, I would prefer to activate a full airplane parachute system than attempt an emergency landing in hostile terrain.

Strongly Agree	**436**	**43.60%**
Somewhat Agree	236	23.60%
Neutral	124	12.40%
Somewhat Disagree	108	10.80%
Strongly Disagree	96	9.60%

Mean	2.19
Standard Dev.	1.35
Variance	1.82
Mean Percentile	76.16%

11. Given a complete electrical failure on a night, cross-country flight, 50 miles from any airport, I would feel more comfortable activating a full airplane parachute system than attempting a landing at the nearest airport.

Strongly Agree	224	22.44%
Somewhat Agree	129	12.93%
Neutral	138	13.83%
Somewhat Disagree	223	22.34%
Strongly Disagree	**284**	**28.46%**

Mean	3.21
Standard Dev.	1.53
Variance	2.35
Mean Percentile	55.71%

12. Given the structural failure of an aileron, I would feel more comfortable activating a full airplane parachute system than attempting a landing at the nearest airport.

Strongly Agree	**352**	**35.27%**
Somewhat Agree	266	26.65%
Neutral	161	16.13%
Somewhat Disagree	140	14.03%
Strongly Disagree	79	7.92%

Mean	2.33
Standard Dev.	1.30
Variance	1.68
Mean Percentile	73.47%

13. Given an unintentional spin at 1,500 AGL, I would feel more comfortable activating a full airplane parachute system than attempting a recovery.

Strongly Agree	**370**	**37.07%**
Somewhat Agree	177	17.74%
Neutral	146	14.63%
Somewhat Disagree	157	15.73%
Strongly Disagree	148	14.83%

Mean	2.54
Standard Dev.	1.48
Variance	2.20
Mean Percentile	69.30%

14. In an emergency, a whole airplane parachute will:

Deliver me gently to the ground.	156	15.81%
Deliver me safely to the ground.	112	11.35%
Deliver me safely to the ground although the airplane may be destroyed	**588**	**59.57%**
Likely seriously injure occupants and destroy the airplane.	131	13.27%

Mean	2.70
Standard Dev.	0.89
Variance	0.79
Mean Percentile	57.42%

15. What is your age?

Younger than 35 years old	301	30.34%
35 to 50 years old	**369**	**37.20%**
51 to 65 years old	281	28.33%
66 or older	41	4.13%

16. How long have you been flying?

Less than 2 years	**389**	**39.25%**
2 to 5 years	173	17.46%

6 to 10 years	78	7.87%
More than 10 years	318	32.09%
Not a pilot; not a student-pilot	33	3.33%

17. Rate your confidence in your flying abilities:

Somewhat Confident	144	14.71%
Normally Confident	**475**	**48.52%**
Very Confident	314	32.07%
Exceedingly Confident	46	4.70%

18. Do you own a parachute system?

Yes	107	13.08%
No	**711**	**86.92%**

19. Have you or a close acquaintance ever been injured while flying?

Yes	234	23.68%
No	**754**	**76.32%**

Note: Questions adapted from Srkal, Milota K. (2005) The Marketability of New Rules and Technologies in General Aviation. Florida State University, D-Scholarship Repository.

Appendix A2: Comparing the percentage of pilots who chose to pull the parachute on each scenario based on age, certificate level, and years of flying experience.

Scenario	Years of Age		By Certificate Level						Years of Flying		
	Less than 35 yrs old N=301	More than 35 yrs old N=691	Sport N=13	Private N=250	COM¹ N=73	ATP² N=22	CFI less than 2 yrs N=117	CFI more than 5 yrs N=103	Less than 2 yrs N=389	More than 10 yrs N=317	Not a pilot N=37
Given an unintentional spin at 1,500 AGL, I would feel more comfortable activating a full airplane parachute system than attempting a recovery.											
strongly agree	**59.47**	27.33	**30.77**	29.30	24.66	22.73	**98.26**	16.50	**48.84**	22.33	**40.54**
agree	11.63	20.49	7.69	20.31	23.29	22.73	1.74	12.62	13.95	18.24	16.22
neutral	13.29	15.12	15.38	16.02	16.44	4.55	-	18.45	15.76	14.47	18.92
disagree	9.30	18.75	**30.77**	17.58	17.81	22.73	-	19.42	13.70	18.87	13.51
strongly disagree	6.31	18.31	15.38	14.06	17.81	**27.27**	-	**33.01**	7.75	**26.10**	10.81
Given the structural failure of an aileron, I would feel more comfortable activating a full airplane parachute system than attempting a landing at the nearest airport.											
strongly agree	**57.48**	25.76	**38.46**	20.48	**28.77**	31.82	100.00	15.53	**46.65**	23.66	**37.84**
agree	16.94	**30.86**	-	**36.14**	27.40	22.73	-	**27.18**	23.45	**29.34**	24.32
neutral	14.62	16.89	-	21.69	20.55	9.09	-	10.68	13.40	13.88	24.32
disagree	8.64	16.16	**38.46**	14.86	13.70	9.09	-	26.21	11.60	18.61	8.11
strongly disagree	2.33	10.33	23.08	6.83	9.59	27.27	-	20.39	4.90	14.51	5.41
Given a complete electrical failure on a night, cross country flight, 50 miles from any airport, I would feel more comfortable activating a full airplane parachute system than attempting a landing at the nearest airport.											
strongly agree	**45.33**	12.48	15.38	8.40	12.33	13.64	**99.12**	4.85	**39.18**	11.08	**27.03**
agree	10.33	14.08	**30.77**	14.80	9.59	13.64	0.88	4.85	13.40	9.18	18.92
neutral	9.00	16.11	15.38	14.80	5.48	9.09	-	14.56	12.89	12.03	21.62
disagree	17.67	24.38	7.69	30.80	28.77	9.09	-	23.30	17.78	23.73	24.32
strongly disagree	17.67	**32.95**	30.77	31.20	43.84	54.55	-	52.43	16.75	**43.99**	8.11
Given an engine failure, I would prefer to activate a full airplane parachute system than attempt an emergency landing in hostile terrain.											
strongly agree	**61.33**	36.18	15.38	**31.60**	**35.62**	40.91	99.13	28.16	**52.84**	35.53	43.24
agree	14.00	27.64	23.08	30.00	**35.62**	22.73	0.87	26.21	20.36	24.84	27.03
neutral	10.33	13.31	**38.46**	12.80	6.85	13.64	-	10.68	12.37	13.84	13.51
disagree	8.33	11.87	23.08	13.20	13.70	13.64	-	15.53	8.25	11.95	13.51
strongly disagree	6.00	11.00	-	12.40	8.22	9.09	-	19.42	6.19	13.84	2.70

Notes:

¹ *COM means commercial pilot without flight instructor certificate*

² *ATP means airline transport pilot*

Critical Incident Stress Management in Aviation
Edited by Jörg Leonhardt and Joachim Vogt

Book Review

Todd P. Hubbard

There are at least two ways to view a critical incident in aviation: from the first responder perspective; and from the perspective of those who care for victims of critical incident stress long after the event has occurred. This point is emphatically made by Mitchell in an interesting text on Critical Incident Stress Management (CISM), compiled by editors Jörg Leonhardt and Joachim Vogt. The primary goal of the text is to expose the reader to the nuts of bolts of CISM: how it is con-structed, made readied for use, and how it is deployed in the field by CISM teams. Although an endorsement by Richard Levenson Jr., on the book cover, suggested that researchers might be interested in this information, the contents are more narrowly focused and are more applicable to other groups: first responders who affect the victims of critical incidents, managers who have a responsibility for personnel who might be affected by critical incidents, persons who wish to join a CISM team, as well as newly licensed psychotherapists and physicians who have been assigned as staff psychologists or medical examiners with an air carrier. This rather short work, if used as a supplement to a course text, will do nicely in courses on Human Factors, Aviation Psychology, and Aviation Mental Health. It's a quick read and undergraduate or graduate students will find it helpfully pithy.

There are a few reasons why this book should be required reading for all those wishing to know more about CISM. Perhaps the chief reason for placing this book on your reading list is its attention to the details of what CISM is designed to be. Mitchell did a good job in Chapter 3 of setting the boundaries for the effec-tiveness of this type of stress management. Leonhardt and Vogt, in Chapter 4, provided additional neurobiological rationale for these boundaries, and Leonhardt (see Chapter 5) described the cognitive and affective aspects of each basic ele-ment of CISM by use of a U-shaped curve.

The book is designed for the skim reader. Within each chapter the reader will find references to other chapters in the text. Furthermore, much of the information is repeated in each chapter, so that if something were missed early on, it will be available in subsequent chapters. To accelerate the skimming process, the

authors have inserted subtitles, preceded by a list of those subtitles. These lists act as internal tables of contents, and they provide an advanced organizer of what is to come. As an added benefit, two appendices, just after the 170 pages of he 12-chapter narrative, provided a view of the core curriculum needed for CISM training and a listing of the certifications that are available as a result of this training.

Another reason to pay attention to this book is its cast of authors: almost entirely non-U.S. It is important to note that Ashgate Publishing, whose home office is in England, attracts authors from Europe; and, because they have an office in Vermont, they also attract authors from the United States, where most of their volumes are sold. Many of the European authors write in their native language and these papers are largely unavailable to U.S. readers because of the language barrier. However, Ashgate, under the guidance of a very clever and resourceful Guy Loft, produces texts in English that include material translated from other languages. Ashgate is a window into endeavors and fields often overlooked by scholars of limited reach who pursue works only in their native language. For this accommodation we should sincerely thank Ashgate. Where else would one find an interpreted review of the accident over Lake Constance (see Chapter 10), ably portrayed and presented by Leonhardt, Minder, Zimmermann, Mersmann, and Schultz?

The third reason why you should read this book is that it is one of a very few texts that expose the critical incident stress issues of persons in air traffic control: a career field that resists observation by outsiders. Much more is said about CISM and air traffic controllers than any other group, even more than airline pilots or flight attendants. We see here a description of stress related issues of air traffic controllers from EUROCONTROL, not controllers from the Federal Aviation Administration. These descriptions came from insiders, such as Riedle (see Chapter 2), Leonhardt (see Chapter 7), and Gaber and Drozd (see Chapter 9).

Reasons to buy this text outweigh reasons not to buy this text, but I must list at least two shortcomings in the book that cannot be overlooked. First, the editors have relied too heavily on Jeffrey Mitchell and the CISM program he developed. Those familiar with CISM and the literature that attacks Mitchell's claims that CISM really works and that it makes a difference will see the dependence on Mitchell to carry the book as a weakness. Second, the authors have unnecessarily confined their referential support to only those in the inner circle of CISM proponents, and in doing so have kept from the reader's eyes any counter opinion. One gets the feeling the all that is credible about CISM is adequately described by Mitchell and those who have co-authored with Mitchell. CISM is also prized more highly than other methods that have served to mitigate the affect of critical incident stress; because CISM, if one were to infer a meaning based on Mitchell's telling of the history of crisis intervention support services, is the outgrowth of earlier attempts to help victims of critical incident stress (see Chapter 6) and is therefore the enhancement of or the improvement of those earlier attempts. Rather than carry on a critique of the literature, I invite the discriminating reader to perform his or her own study of the literature in this text, and if the contents of this book show that CISM is a superior methodology to any other approach, then the contents of this book vindicate the authors.

In closing, I leave the reader with a word of caution. Some of the information in the book requires some understanding of stress related mental and emotional disorders. The authors do a fairly good job describing and defining these disorders, particularly Posttraumatic Stress Disorder (PTSD), but mental disorders are clearly beyond the scope of the book and any information about these disorders, their diagnosis, and the characteristics observable by trained psychologists or psychiatrists are more clearly described and handled in another book by Ashgate, Aviation Mental Health (2006). Aviation Mental Health gives the clinician's view of many of the same issues expressed in Critical Incident Stress Management in Aviation. Reading this text and Aviation Mental Health together will provide the reader with a more complete picture of mental health issues: those pertaining to psychological first aid; and, those pertaining to long-term psychological care.

Aviation English for ICAO Compliance.
By Henry Emery and Andy Roberts
with Ruth Goodman and Louis Harrison.

Cleared for Takeoff: English for Pilots, Books 1 & 2.
By Liz Mariner

English for Aviation.
By Sue Ellis and Terence Gerighty

Book Reviews

Graham Elliott and Theresa White
FAA Academy

For those responsible for arranging compliance with the new International Civil Aviation Organization (ICAO) English proficiency standards required as of March 2008, more help has arrived.

After a hiatus of more than two decades with no commercially produced aviation English course books widely available, no less than three new courses are reaching the market to address English language proficiency requirements. This timely trio joins three new on-line Aviation English courses and hopefully signals a full focus on the installation of successful training and the end of a 5-year period where the industry has grappled with what and how to test.

These aviation English courses appear now due to widening recognition that faulty communication from weak English proficiency has contributed to catastrophe in the past. This realization is thanks in large part to an ICAO study group that produced the key Document 9835: *Manual on the Implementation of ICAO Language Proficiency Requirements* in 2004 and groups of dedicated contributors before and since.

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-800, P.O. Box 25082, Oklahoma City, OK 73125. E-mail to kay.chisholm@faa.gov.

The new courses, two from the U.K. and one published in the U.S.A., all with audio files, commendable content, and many pedagogical qualities, are not aimed at the same audiences. Learner needs vary a good deal, and so users of these courses will need to be clear who they are aiming to serve: trainee controllers, *ab initio* pilot trainees, or experienced controllers or pilots. The courses vary in purpose, length, and approach. How to decide which of these courses, if any, to adopt?

Three aspects that combine to address and solve English language proficiency issues are training, testing, and administration. We are pleased to see these three new aviation English courses address mainly the training side of the triangle—those with responsibility for programming for compliance will have to manufacture their administrator support and find other sources for testing for ICAO compliance, most likely from supplementary providers.

Twenty-five years in aviation English has shown that cognition is embodied in action—pilots and controllers will learn their target English most readily if they work and interact in many job-related tasks. For this, the ICAO Doc 9835 asks for assessment of proficiency in six elements of language: Pronunciation, Structure, Vocabulary, Fluency, Comprehension, and Interactions over six levels of which Operational Level 4 (L 4) is the target. It seems logical then that these courses will (a) support learners in acquiring specific aviation vocabulary, (b) embrace use of radiotelephony between pilots and controllers so that learners improve their proficiency in the English used in ATC, and (c) allow them to build skills in plain English so they can interact in non-standard aviation situations.

These course books are intended to be part of the solution—training materials for English language development for civil authorities and airlines. What do they contain, and how good are they?

Aviation English for ICAO Compliance
By Henry Emery and Andy Roberts. Macmillan Publishers Limited.
One book of 12 units for 120 hours, with CDs for classroom use or self-study.
Stated target audience: Intermediate level English for Specific Purposes (ESP)

The principal authors are experienced aviation English teachers. The 127 pages hold 12 units, either as self-study or for classroom use. They claim to teach communication skills for pilots and air traffic controllers, helping them to achieve and maintain ICAO Operational Level 4.

Each colorfully illustrated unit begins with a reading on an aviation theme, and is followed by listening and speaking activities presented with professionally produced video and/or audio recordings. Following suggestions from ICAO Doc 9835, the units are organized by aviation topics such as aerodrome layout, runway incursions, fly by wire, electrical failure, engine failure, bird strike, hydraulic failure, on-board fire, gear/braking problems, decompression. Despite a lack of empirical evidence to support inclusion, this seems like a rich collection of industry topics that will be of interest and benefit to aviation personnel. The accompanying recordings, delivered for the classroom at an appropriately slower rate than in the real

world, use a wide range of accented Englishes, which gives *Aviation English* a solid international footprint.

Systematically organized among these topics are exercises that address features of pronunciation and English grammatical structure. These exercises also include a focus on word and sentence stress, intonation and pausing, diphthongs, question forms, and past tense endings, which are all issues probably beyond the knowledge of flight personnel to organize and tackle. Each unit contains a series of functions of language that include asking for clarification, explaining how something works, stating intentions, announcing decisions, expressing non-understanding, and so on. Each of four sections in the units contain vocabulary exercises, and the back of the book has a pair work section and listening scripts.

The accompanying Teacher's Book (which we did not see for this review) should help less well trained aviation instructors to operate the contents.

Because *Aviation English* is designed in the tradition of English teaching texts, its color and interesting graphic design may, perversely, cause it not to be taken as seriously as it warrants by the aviation community. Even in the hands of minimally qualified language instructors, this well-conceived English training product can play a central role in an English for Aviation program, which is targeting L 4.

Cleared for Takeoff: English for Pilots Books 1 & 2.
By Liz Mariner. AELink Publications Inc.
Two books each with 60 hours, with audio CDs. Includes a glossary and abbreviations.
Stated target audience: Intermediate level

The principal author is a flight instructor experienced in working with non-native English speakers in flight training. This two-book course also provides activities in an accompanying DVD. Publicity says it was designed for use in flight schools, international airline training, and English for aviation in language schools, either as an individual workbook or with classroom groups.

In essence, this is a simplified flight training manual made to be a starting point for flight training. As such, it will find favor with flight instructors who are working to build a language foundation with their non- English speaking trainees. The course targets standard procedures at the start of pilot training, with trainees at Pre-Operational L 3.

Developed before the ICAO proficiency standards were published, this course does not systematically address the ICAO requirements. There is no coverage of the functions of language recommended in ICAO Doc 9835, and no explicit inclusion of plain language elements, for example, for handling of emergencies, this course by itself cannot lead users to meet ICAO L 4.

The twelve units include many excellent descriptive graphics with suitably slower-paced audio recordings in U.S. accented English, often reaching above

delivery of a single word into appropriate chunks of language. The first six units focus on primary flight training principally in a U.S.-flavored setting. It continues with vocabulary associated with small aircraft, air traffic control communications; student and instructor communications; airport features and traffic pattern; aircraft features, ATIS; and the basics of flight. The units in Book 2 are pattern work; aircraft checklists; weather and weather reports; VFR navigation; and operating in controlled airspace.

As demanded by flight training, there is a heavy vocabulary load across both books, including, for example, thoughtful terms for instructors and students to work together, and navigational features on charts. However, there appears to be no explicit means of supporting learners to understand, acquire, and remember them. In Unit 2 alone, new vocabulary could total over a hundred items on airports, hazards, abbreviations, and surface features. While the useful 36-page glossary for both books holds over 500 items, and a list of useful abbreviations adds nearly 90 more, there are hundreds more items appearing in the texts which are not included in the glossary, such as *VOR* (Book 1, page 8) and *in the vicinity of* (Book 1, page 72) which will require explanations.

> *…a well structured vocabulary program needs a balanced approach that includes explicit teaching* (as demanded in this course) *together with activities providing appropriate contexts for incidental learning.*
>
> J.S.Decarrio

Everything is shown in a clear context, and there are useful collocations' such as "The ramp is sometimes called the *tarmac* or the *apron.*" However, presentations are generally well-contextualized lists that a tireless instructor will have to explain, one by one. While the recognition, understanding, and use of this vocabulary is clearly necessary, unless it is delivered in structured and logical patterns, and surrounded by grammatical context, learners will struggle to correctly recall, apply, and practice before applying the terms fluently.

A mistake common among those who are not aware of the need for comprehensible input for English learners occurs frequently—that of using teaching language above a learner's ability to comprehend:

> "...repeat the numbers after the speaker. Stressed syllables are shown with CAPITAL letters. zero, one, two, ..." (Book 1, page 2.)

If learners know what a *stressed syllable* is, we likely will not need to worry about how they pronounce *one*.

For English for Specific Purposes training to proceed efficiently and effectively, in prior pre-intermediate English as a Second Language (ESL) programming we first have to establish in learners a basic facility in manipulating English grammar, a basis of common vocabulary with intelligible pronunciation and some fluency, until the learner can handle the demands of English for Aviation.

This course is a shadow flight training manual with an excess of Q & A methodology for English learners. In the hands of a flight instructor who is sensitive to learner needs, it can work (and apparently has worked for the author), but the technical content is probably above the capacity of an uninformed English instructor ("What's a VOR?") to deliver.

The course needs fewer touch-and-go language learning activities such as *Choose the correct definition*…; *Complete the table*…; *Draw lines to match*…; *Fill in the missing words*… (Book 2, Unit 7). The course should provide more explicit opportunities for learners to repeatedly prepare and deliver the specific language they will need on the job, such as *Ask for permission to*…; *Conduct the checklist for*…; *Practice approaches to*…; *Describe when*…; *Explain why*…, so they will gain guided experience in creating and manipulating English to say what they want to say.

To solve this, the course will benefit from a teacher guide (currently in preparation) where more activities will lead to practice and fuller exploitation of the many excellent recordings, photos, diagrams, and tables.

English for Aviation for Pilots and Air Traffic Controllers
By Sue Ellis and Terence Gerighty. Oxford University Press (OUP), Express series
One book of 8 units for 30 hours self study, with audio CD, and CD-ROM with interactive material. *Stated target audience:* Intermediate level, short intensive course for pilots.

Written by two authors with extensive experience in aviation English and aligned to ICAO guidelines, this 69-page pre-intermediate course book leads sequentially from pre-flight checks, through the flight path, to switching off the engines. The gate-to-gate syllabus is closest of the three to work realities in aviation.

Part of the OUP Express series that offer "short courses in specialist English," it is approximately half the length of the other two courses; it contains up to 30 hours of work in eight units. We consider that with suitable instructor guidance, it could provide a basis for more than 60 classroom hours. It includes an audio CD, recorded with varied foreign English accents in interactive exercises for each unit. These provide interactions and listening comprehension of realistic pilot and air traffic controller interchanges, grammar uses, and applied vocabulary for both standard radiotelephony and non-standard plain English. It includes a list of all key words in the book.

The book is attractively designed and laid out and uses clear photographs, informative graphics, illustrations, and signs in support of unit topics. The first unit introduces air communications and aims to include learner experience of familiar routes and airports. It is followed by aerodrome information, leading to pre-flight checks, clearances, and handling delays and problems. Unit 3 includes ground

movements from start up and push back, to taxi and take off. And this is the tight sequence the course follows: unit 4 is departure, reports in climbing and cruising and includes traffic and weather; unit 5 contains en route events and leads to descent; unit 6 contains VFR and IFR contacts and approaches with radar vectoring; unit 7 deals with aspects of landing clearances, missed approaches, runway conditions, and ground movement; and finally, unit 8 includes exercises on dealing with authorities.

The experienced flight and controller personnel we showed this book to saw that it covered all phases, and they approved of the handling of the technical topics. This course will meet the approval of many other learners in aviation because its structure follows the world they know.

Because of its brevity, however, users might find that by itself the course does not go deep enough or is long enough to truly affect their English performance. This can be remedied in part by informed training design and tight benchmarking by each organizational user and by additional activities on a students' website being prepared by the publisher.

These courses each contribute solid English content at appropriate levels for various trainee targets, and those responsible for English compliance have three new tools from which to select.

It is clear that there will always be situations, usually non-routine, for which phraseologies do not exist; therefore, only continuous and sustained commitment from organizations across the industry will raise English standards and produce safer communications in English. We would like to see organizations recognize the need for their pilots and/or controllers to become familiar with the language required to be safe in the regions and airports in which they are training and working, unlike the generic language necessarily included in all these courses.