

2010

*International Journal of Applied Aviation Studies*

Volume 10, Number 1

A publication of the FAA Academy



Federal Aviation  
Administration

# ***International Journal of Applied Aviation Studies***

---

A Publication of the FAA Academy Oklahoma City, Oklahoma



Volume 10, Number 1 2010



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](mailto: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

International Journal of Applied Aviation Studies  
Volume 10, Number 1  
ISSN Number: 1546-3214  
Copyright © 2010, FAA Academy  
1st Printing August, 2010

## 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.

## PUBLISHER

Academy Superintendent  
Deputy Superintendent

Gary Condley  
Sunny Lee-Fanning

## PUBLICATION COMMITTEE

Manager, Regulatory Standards Division  
Manager, Technical Operations Training Division  
Manager, Air Traffic Division  
Manager, Airports & International Training Division

David Long  
Jessie McMullen  
William J. Mumper  
Sherry Reese

## EDITORIAL STAFF

Managing Editor  
Associate Editor  
IJAAS Founding Editor and Research Consultant

Kay Chisholm  
Deann King  
Todd P. Hubbard

## CONSULTING EDITORS

### *International*

Ruwantissa I.R. Abeyratne	ICAO, Montreal, Canada
Pierre R. Band	Institut du Cancer de Montréal, Québec, Canada
James Barker	Dalhousie University, Halifax, Nova Scotia, Canada
Simon Bennett	Institute of Lifelong Learning, Leicester, UK
Henk Blom	National Aerospace Laboratory NLR, Amsterdam, The Netherlands
Robert Bor	Royal Free Hospital, London, UK
Peter Brooker	Cranfield University, Cranfield, UK
Bryan Burke	Macquaire University, Sydney, Australia
Ing. F.J.L. Bussink	National Aerospace Laboratory, Amsterdam, The Netherlands
Petr Cášek	Honeywell International Advanced Technology Center Czech Republic
Michael B. Charles	Southern Cross University, Tweed Heads, NSW Australia
Klaus Christoffersen	Acuite, Hamilton, Ontario, Canada
Desmond M. Connolly	QinetiQ, Farnborough, Hampshire, UK
Finian Connolly	Executive & Professional Training Inst., Drogheda, Co Louth Ireland
Hugh David	Eurocontrol, East Sussex, UK
Sidney Dekker	Linköping Institute of Technology, Linköping, Sweden
Teresa C. D'Oliveira	ISPA, Lisbon, Portugal
Robert van Doorn	Universiteit Maastricht, The Netherlands
Åsa Ek	Lund University, Lund, Sweden
Alois Farthofer	Human Factors Research, St. Georgen, Austria
Massimo Felici	University of Edinburgh, Edinburgh, UK
Rodney Fewings	Cranfield University, Cranfield, Bedfordshire UK
Gerard J. Fogarty	University of Southern Queensland, Toowoomba, Australia
Jarle Gimmestad	Braathens Airlines, Fornebu, Norway
Eric L. Groen	TNO Defence, Soesterberg, The Netherlands

Go to Table of Contents

Gael P. Hammer	Johannes Gutenberg-University, Mainz, Germany
Don Harris	Cranfield University, Cranfield, Bedfordshire, UK
Irene Henley	University of Western Sidney, Sidney, Australia
Ivonne A. Herrera	SINTEF Technology & Society Safety Research, Trondheim, Norway
Toshio Hirose	Sendai Nishikicho Clinic & Occupational Health Center, Sendai Japan
Alexandra Holmes	Clockwork Research Ltd., London, UK
Hans-Jürgen Hörmann	Deutsches Zentrum für Luftund Raumfahrt (DLR), Hamburg, Germany
Graham Hunt	Massey University, Auckland, New Zealand
Magnus Jacobssen	Linköping Institute of Technology, Linköping, Sweden
Lynn M. Jeffrey	Massey University Auckland, Auckland, New Zealand
Chris W. Johnson	University of Glasgow, Glasgow, Scotland
K. Wolfgang Kallus	Karl Franzens University, Graz, Austria
Michael John Kay	RMIT University, Melbourne, Victoria, Australia
Jörg Leonhardt	Deutsche Flugsicherung GmbH, Langen, Germany
Dirk Lehmkuhl	University of St. Gallenm, St.Gallen, Switzerland
George Leloudas	Gates and Partners Solicitors, London, UK
Wen-Chin Li	National Defense University, Taipei, Taiwan, R.O.C.
Monica Martinussen	University of Tromsø, Tromsø, Norway
Elizabeth Mathews	Aviation English Services, Wellington, New Zealand
Jeremy Mell	French Embassy, Washington, DC
Jim Mitchell	University of Western Sydney, Penrith South, Australia
Kjell Mjøs	Norwegian University of Science and Technology,Oslo, Norway
Brett Molesworth	University of New South Wales, Sydney, Australia
David G. Newman	Flight Medicine Systems PTY LTD, Victoria, Australia
Jan Noyes	University of Bristol, Bristol, UK
Paul O'Connor	National University of Ireland, Galway, Ireland
David O'Hare	University of Otago, Dunedin, New Zealand
Esther Oprins	Air Traffic Control the Netherlands, Schiphol Airport, Netherlands
Christine Owen	University of Tasmania, Tasmania, Australia
Michal Pěchouček	Czech Technical University, Prague, Czech Republic
Stefanie Petrie	Luftwaffe Institute of Aviation Medicine, Fürstenfeldbruck Germany
Andrew Potter	Cardiff University, Cardiff, UK
Teemu Reiman	VTT Technical Research Centre of Finland, Otaniemi, Finland
Alfred Roelen	National Aerospace Laboratory NLR, Amsterdam, The Netherlands
Jan Joris Roessingh	National Aerospace Laboratory, Amsterdam, The Netherlands
Jens Rolfsen	Det Norske Veritas, Høvik, Norway
Paul Roosens	University of Antwerp, Antwerpen, Belgium
Michael Siegrist	Institute for Environmental Decisions, Zurich, Switzerland
T. Leigh Signal	Massey University, Wellington, New Zealand
Adrian Smith	Aeromedical Centre, Dhahran, Kingdom of Saudi Arabia
Richard H. Y. So	University of Essex, Colchester, UK
Manuel Soler	Universidad Rey Juan Carlos, Madrid, Spain
Sybert Stroeve	NLR Air Transport Safety Institute, Amsterdam, The Netherlands
Shinji Suzuki	University of Tokyo, Tokyo, Japan

Hakan Tarakci	Melbourne Business School, Melbourne, Australia
Steve Thatcher	University of South Australia, Mawson Lakes, Australia
M. "Mattie" Tops	Leiden University, Leiden, The Netherlands
Claudia van de Wal	VALK Foundation, Netherlands
Lucas van Gerwen	VALK Foundation, Netherlands
Leo Vermeulen	University of Pretoria, Pretoria, South Africa
Joachim Vogt	Technische Universität Darmstadt, Darmstadt, Germany
Anthony S. Wagstaff	Institute of Aviation Medicine, Oslo, Norway
Stephen Walsh	Interaction Training Associates, Netherlands
Mark Wiggins	Macquarie University, North Ryde, NSW, Australia
Damien J. Williams	University of Bristol, Bristol, UK
Robert Wolfger	Austrian Airlines, Strasshof, Austria
Rogier Woltjer	Linköping Institute of Technology, Linköping, Sweden
Jin-Ru Yen	National Taiwan Ocean University, Keelung, Taiwan, ROC
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
Steve Anderson	St. Cloud State University, St. Cloud, MN
Rick Anglemeyer	SCSI, Colorado Springs, CO
Anthony R. Artino, Jr.	Uniformed Services University, Bethesda, MD
Francis "Frank" Ayers	Embry-Riddle Aeronautical University, Prescott, AZ
Larry Bailey	FAA Academy, Oklahoma City, OK
Jack Barker	Mach One Leadership, Miami, FL
R. Kurt Barnhart	Kansas State University, Salina, KS
Robert Baron	The Aviation Consulting Group, Myrtle Beach, SC
Paul T. Bartone	National Defense University, Ft. McNair, Washington, DC
Ellen Bass	University of Virginia, Charlottesville, VA
J. Matthew Beaubien	American Institutes for Research, Washington, DC
Beth M. Beaudin-Seiler	Western Michigan University, Battle Creek, MI
Wendy Beckman	Middle Tennessee State University, Murfreesboro, TN
Sean Belcher	NASA Ames Research Center, Moffett Field, CA
Theodore Beneigh	Embry-Riddle Aeronautical University, Daytona Beach, FL
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
Erin E. Block-Bowen	Purdue University, West Lafayette, IN
Deborah Boehm-Davis	George Mason University, Fairfax, VA
Philip Bos	University of Central Missouri, Warrensburg, MO
Brent Bowen	Purdue University, W. Lafayette, IN
Tim Brady	Embry-Riddle Aeronautical University, Daytona Beach, FL
Jonathan Bricker	University of Washington, Seattle, WA
Christopher R. Brinton	Mosaic ATM, INC., Leesburg, VA



Dana Broach	Civil Aerospace Medical Institute, Oklahoma City, OK
Philip Budd	Graduate Studies in Counseling, Bethany, OK
Judith Bürki-Cohen	Volpe National Transportation Systems Center, Cambridge, MA
Ernesto A. Bustamante	University of Idaho, Moscow, ID
James N Butcher	University of Minnesota, Minneapolis, MN
Paul W. Buza	Wuesthoff Medical Center, Melbourne, FL
Evan Byrne	National Transportation Safety Board, Washington, DC
John H. 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
William Caldwell	Southern Illinois University, Carbondale, IL
Kim Cardosi	Volpe National Transportation Systems Center, Cambridge, MA
Thomas Q. Carney	Purdue University, West Lafayette, IN
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
Anna T. Cianciolo	Command Performance Research, Inc., Champaign, IL
Daniel W. Clark	Critical Concepts Consulting, Olympia, WA
Maria Consiglio	NASA Langley Research Center, Hampton, VA
David Conway	Southeastern Oklahoma State University, Durant, OK
Paul A. Craig	Middle Tennessee State University, Murfreesboro, TN
Jerry Crutchfield	Civil Aerospace Medical Institute, Oklahoma City, OK
Thom Curtis	University of Hawaii, Hilo, HI
John Deaton	Florida Institute of Technology, Melbourne, FL
Stephen Dedmon	Embry-Riddle Aeronautical University, Daytona Beach, FL
Asaf Degani	NASA Ames Research Center, Moffett Field, CA
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
Theresa A. Domagalski	Western Carolina University, Cullowhee, NC
Nancy Dorigi	NASA Ames Research Center, Moffett Field, CA
Wayne Dornan	Middle Tennessee State University, Murfreesboro, TN
Boyd Falconer	Russell Reynolds Associates, Chicago, IL
Richard Fanjoy	Purdue University, West Lafayette, IN
Douglas R. Farrow	Federal Aviation Administration, Washington, DC
Jeffrey Forrest	Metropolitan State College of Denver, Denver, CO
Paul M. Foster	FAA Western-Pacific Regional Office, Lawndale, CA
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
Devra L. Golbe	Hunter College/CUNY, New York, NY
A. F. Grandt	Purdue University, West Lafayette, IN
Mavis Green	Utah Valley University, Provo, UT
Regina A.Greenwood	Nova Southeastern University, Fort Lauderdale-Davie, FL



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
P. A. "Peter" Hancock	University of Central Florida, Orlando, FL
Jeffrey T. Hansberger	George Mason University, Fairfax, VA
Frederick D. Hansen	Oklahoma State University - Tulsa, Tulsa, OK
Ron A. Hess	University of California, Davis, CA
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
C.M. Holloway,	NASA Langley Research Center, Hampton VA
Kent Holtorf	Holtorf Medical Group, Torrance, CA
Willem Homan	Western Michigan University, Battle Creek, MI
Amy Hoover	Central Washington University, Ellensburg, WA
Bruce Hoover	Parks College, St. Louis University, St. Louis, MO
Todd P. Hubbard	Oklahoma State University, Stillwater, OK
David R. Hunter	Artis, LLC, Reston, VA
Peter Hwoschinsky	FAA National Headquarters, Washington, DC
Kathi Ishimaru	FAA, Northwest Mountain Regional Office, Renton, WA
Florian Jentsch	University of Central Florida, Orlando, FL
Alan W. Johnson	Air Force Institute of Technology, Wright-Patterson AFB, OH
Jeffrey Johnson	St. Cloud State University, St. Cloud, MN
Robert Kaps	Southern Illinois University, Carbondale, IL
Merrill R. Karp	Arizona State University, Mesa, AZ
Kim Kenville	University of North Dakota, Grand Forks, ND
David King	Northwestern State University, Natchitoches, LA
Deann King	FAA Academy, Oklahoma City, OK
Raymond E. King	Civil Aerospace Medical Institute, Oklahoma City, OK
Mike D. Kinney	U.S. Dept of Energy, Natl Nuclear Security Admin., Las Vegas, NV
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
Steven J. Landry	Purdue University, West Lafayette, IN
Jeff Lancaster	Honeywell Aerospace Advanced Technology, Golden Valley, MN
Victor LaSaxon	The Boeing Company, Midwest City, OK
Jeffrey M. Lating	Loyola University, Baltimore, MD
David Lecraw	Volpe Center, Cambridge, MA
Suresh K. Lodha	University of California, Santa Cruz, CA
Loukia Loukopoulos	NASA Ames Research Center, Moffett Field, CA
Chien-tsung Lu	Purdue University, W. Lafayette, IN
Gary Luckenbaugh	Lockheed Martin, Rockville, MD
Rebecca Lutte	University of Nebraska at Omaha, Omaha, NE
Nickolas D. Macchiarella	Embry-Riddle Aeronautical University, Daytona Beach, FL
Richard L. Mangrum	Kent State University, Kent, OH
Carol Manning	Civil Aerospace Medical Institute, Oklahoma City, OK

Roderick "Rod" D. Margo	Condon & Forsyth LLP, Los Angeles, CA
Steven Marks	Oklahoma State University, Stillwater, OK
Lynne Martin	NASA Ames Research Center, Moffett Field, CA
Royce Ann Martin	Bowling Green State University, Bowling Green, OH
Luigi Martinelli	Princeton University, Princeton, NJ
Patrick R. Mattson	St. Cloud State University, St. Cloud, MN
Celeste Mayer	North Carolina State University, Raleigh, NC
Raymon M. McAdaragh	FAA, Hampton, VA
Michael E. McCauley	Naval Postgraduate School, Monterey, CA
Kathleen L. McFadden	Northern Illinois University, DeKalb, IL
Ole J. Mengshoel	NASA Ames Research Center, Moffett Field, CA
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
Edward F. Murphy, Jr	Embry Riddle Aeronautical University, Vacaville, CA
Tom Nesthus	Civil Aerospace Medical Institute, Oklahoma City, OK
Mary Niemczyk	Arizona State University, Mesa, AZ
Mark I. Nikolic	The Boeing Company, Seattle, WA
Michael S. Nolan	Purdue University, West Lafayette, IN
Gary J. Northam	Embry-Riddle Aeronautical University, Prescott, AZ
Dale B. Oderman	Purdue University, West Lafayette, IN
Robert P. Olislagers,	Arapahoe County Public Airport Authority, Englewood, CO
Jim Oppermann	Ohio State University, Columbus, OH
Korhan Oyman	Florida Institute of Technology, Melbourne, FL
Raja Parasuraman	George Mason University, Fairfax, VA
Bonny Parke	NASA Ames Research Center, Moffett Field, CA
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
Thomas Prevot	NASA Ames Research Center, Moffett Field, CA
Roni Prinzo	Civil Aerospace Medical Institute, Oklahoma City, OK
Michael D. Proctor	University of Central Florida, Orlando, FL
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	University of Central Missouri, Warrensburg, MO
Esa M. Rantanen	Rochester Institute of Technology, Rochester, NY
Dawna L. Rhoades	Embry Riddle Aeronautical University, Daytona Beach, FL
Stephen Rice	New Mexico State University, Las Cruces, NM
Robert Ripley	Auburn University, Auburn, AL
Vladimir Risukhin	Western Michigan University, Kalamazoo, MI

Charles L. Robertson	University of North Dakota, Grand Forks, ND
Bonnie L. Rogers	Wichita State University, Wichita, KS
Ericka Rovira	U. S. Military Academy, West Point, NY
Jason J. Saleem	Indiana Univ.-Purdue University, Indianapolis IN
Christian M. Salmon	The George Washington University, Ashburn VA
Steven M. Samuels	US Air Force Academy, Colorado Springs, CO
Nadine Sarter	University of Michigan, Ann Arbor, MI
Ian Savage	Northwestern University, Evanston, IL
Stewart W. Schreckengast	Purdue University, West Lafayette, IN
David Schroeder	Civil Aerospace Medical Institute, Oklahoma City, OK
Pete Schumacher	University of North Dakota, Grand Forks, ND
Gregory L. Schwab	Indiana State University, Terre Haute, IN
Scott Shappell	Clemson University, Clemson, SC
J. Anthony Sharp	Elizabeth City State University, Elizabeth City, NC
John W. Sheremeta, Jr.	American Airlines, Lake Ronkonkoma, NY
Mark A. Sherman	Farmingdale State College, Farmingdale, NY
Lance Sherry	George Mason University, Fairfax, VA
L. James Smart	Miami University, Oxford, OH
Kenneth Sperry	The Boeing Company, Midwest City, OK
John E. Stewart II	U.S. Army Research Institute, Fort Rucker, AL
Earl S. Stein	William J. Hughes Technical Center, Atlantic City, NJ
Thomas A. Stoffregen,	University of Minnesota, Minneapolis, MN
Alan J. Stolzer	Embry-Riddle Aeronautical University, Daytona Beach, FL
Allan Storm	US Air Force, Arlington, VA
Judith B. Strother	Florida Institute of Technology, Melbourne, FL
Craig Gerald Stroup	American West Airlines, Phoenix, AZ
Joy L. Taylor	Stanford/VA Aging Clinical Research Center, Palo Alto, CA
Terry von Thaden	University of Illinois, Savoy, IL
Jeffrey M Thompson	University of Minnesota, Minneapolis, MN
Pamela S. Tsang	Wright State University, Dayton, OH
Judith Foss Van Zante	NASA Glenn Research Center, Cleveland, OH
Stephen Véronneau	CAMI, Oklahoma City, OK
Ken Wallston	Vanderbilt University, Nashville, TN
Ron Ward	FAA Academy, Oklahoma City, OK
Thomas Weitzel	Embry-Riddle Aeronautical University, Daytona Beach, FL
Victor Welzant	International Critical Incident Stress Foundation, Inc. Ellicott City, MD
Michael Wetmore	Central Missouri State University, Great Falls, MT
Christopher Wickens	University of Illinois Urbana-Champaign, Urbana, IL
Michael Wiggins	Embry-Riddle Aeronautical University, Daytona Beach, FL
Craig Williams	Fort Wayne-Allen County Airport Authority, Fort Wayne, IN
Kevin Williams	Civil Aerospace Medical Institute, Oklahoma City, OK
Dale Wilson	Central Washington University, Ellensburg, WA
Donna Forsyth Wilt	Florida Institute of Technology, Melbourne, FL
Peter Wolfe	Professional Aviation Board of Certification, Washington, DC
Shawn R. Wolfe	NASA Ames Research Center, Moffett Field, CA
Richard D. Wright	Volpe National Transportation Systems Center, Weston, MA
John Young	Purdue University, West Lafayette, IN

## PHILOSOPHY STATEMENT

<sup>1</sup>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).<sup>1</sup>

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.

<sup>1</sup>Lanczos, C. (1988). Applied Analysis. Mineola, NY: Dover Publications, Inc.

## **Papers**

What are the causes of pilot burnout? In Fanjoy, Harriman, and DeMik's study, *Individual and Environmental Predictors of Burnout among Regional Airline Pilots*, findings suggest shortened rest periods, adverse weather, aircraft maintenance issues, and pressures to meet on time performance goals are potential contributors to pilot burnout and safety concerns.

In *Pilots, Controllers and Mechanics on Trial: Cases, Concerns, and Countermeasures*, Sidney Dekker examines criminal prosecutions of pilots, air traffic controllers, and maintenance technicians in the wake of aviation incidents and accidents worldwide. Dekker assesses the possibility of mitigating the criminalization trend, examines the concerns surrounding criminalization, and reviews the diversity of countermeasures that are currently being developed in aviation.

In *Differential Effects of Likelihood Alarm Technology, Type of Automation, and Type of Task on Decision Making as Applied to Aviation and UAS Operations*, Bustamante and Clark examine the differential effects of likelihood alarm technology on decision-making accuracy. Their findings reveal theoretical implications for decision-making models and practical applications to the design of decision support tools for aviation and UAS operations.

Smith, Bjerke, NewMyer, Niemczyk, and Hamilton present the findings from the 2010 Pilot Source Study in *Pilot Source Study: An Analysis of Pilot Backgrounds and Subsequent Success in US Regional Airline Training Programs*. University researchers independently analyzed the data and integrated their results. Of the 2,156 pilots in the study, more than half of them had a baccalaureate degree, had an aviation degree, were flight instructors, had 1,000 or fewer hours of flight time, and had no prior airline pilot or corporate pilot experience.

A well-trained CISM team can help high risk/high reliability organizations become high resiliency organizations. In *Critical Incident Stress Management (CISM): An Effective Peer Support Program for Aviation Industries*, Mitchell and Leonhardt provide an overview of the basic principles and practices of CISM within the aviation industry. This article details the core services provided by such teams and suggests the types and level of training necessary for CISM teams.

FAA/Industry Training Standards (FITS) stands out as an example of how industry, academia, and the regulating body can work together to create workable and useful solutions to increase safety in general aviation. The FITS program is centered on three key elements: scenario-based training, single pilot resource management, and learner-centered grading. In *Changing General Aviation Flight Training by Implementing FAA Industry Training Standards*, Halleran and Wiggins examines the development, components, application, and outcomes of the FITS program.

In the *Reciprocal Development of Expertise in Air Traffic Control*, Owen and Page investigate learning in the air traffic control workplace, where experienced controllers act as instructors of trainee controllers. The results show how engaging in on-the-job-training enriches the reflective learning process for instructors.

[Go to Table of Contents](#)

Organizational strategies offer wide-ranging benefits in the high reliability domain where reflective practices are limited by the intensity and immediacy of the work.

Rogers, Boquet, Howell, and DeJohn report on a research experiment that evaluated simulator-based upset recovery training transfer in *A Two-Group Experiment to Measure Simulator-Based Upset Recovery Training Transfer*. Statistical analysis of data collected during flight testing suggests that simulator-based training combined with classroom instruction improves a pilot's ability to recover an airplane from an upset.

Knecht and Ball investigate whether brief video weather training products can significantly affect pilot weather knowledge and flight behavior in the face of potential instrument meteorological conditions in their paper, *Effects of Video Weather Training Products on General Aviation Pilot Weather Knowledge and Flight Behavior into Adverse Weather*. They conclude that weather training requires systematic, lengthy study and practice.

In *MD-11 Landing Incidents/Accidents 1992 – 2009: A Textual Analysis of NTSB and FAA Reports*, John T. Cocklin provides the results of a textual analysis of MD-11 incident/accident public reports from 1992-2009. Three distinct types of landing events were discovered. Cocklin discusses the MD-11's relaxed stability design and related pitch sensitivity, and reviews possible improvements in incident/accident reporting.

Two senior level Aeronautical Engineering Technology (AET) classes at Purdue University were chartered to produce a component for actual use within aircraft maintenance laboratories. Dubikovsky, Ropp, and Lesczynski's, *Developing Next Generation Research Competencies Through Collaborative Student Design and Advanced Manufacturing Projects* examines the AET program, which incorporates additional outcomes of forward thinking design and innovative research considerations into a traditional design-test-build project.

Based on the results of an international survey of flight attendants and pilots, Brown and Rantz find a critical need for improved communication between crewmembers. In *The Efficacy of Flight Attendant/Pilot Communication in a Post-9/11 Environment: Viewed from Both Sides of the Fortress Door*, the authors note the causes and provide recommendations for improving preflight briefings and CRM training.

---

### **Previously Published Technical Reports**

*The article within this section of the IJAAS is a reprint in whole or in part of government technical report. Technical Reports are peer reviewed **as is** and are reprinted on recommendation of the peer review panel. The IJAAS should not be cited. The original publication citation is provided on the title page of the report.*

---

Hewett, Curry, and Gaydos assess the impact of low to moderate levels of hypoxia on the cognitive performance of aircrew in the *Subtle Cognitive Effects of Moderate Hypoxia*. The results indicate that healthy individuals do not experience significant cognitive deficit, as measured by the CogScreen®-HE, when exposed to moderate levels of hypoxia at or below 14,000 ft.

KC

# **International Journal of Applied Aviation Studies**

Volume 10, Number 1

2010

Review Process	1
Policy and Disclaimer	2
Production Staff	3
Philosophy Statement	10
Editor's Notes	11

## **Papers**

<b>Individual and Environmental Predictors of Burnout among Regional Airline Pilots</b>	<b>15</b>
---	-----------

*Richard O. Fanjoy and Stanley L. Harriman*

<b>Pilots, Controllers and Mechanics on Trial: Cases, Concerns and Countermeasures</b>	<b>31</b>
--	-----------

*Sidney Dekker*

<b>Differential Effects of Likelihood Alarm Technology, Type of Automation, and Type of Task on Decision Making as Applied to Aviation and UAS Operations</b>	<b>51</b>
---	-----------

*Ernesto A. Bustamante and Rylan M. Clark*

<b>Pilot Source Study: An Analysis of Pilot Backgrounds and Subsequent Success in US Regional Airline Training Programs</b>	<b>73</b>
---	-----------

*Guy M. Smith, Elizabeth Bjerke, David A. NewMyer,  
Mary Niemczyk, and Raymond A. Hamilton*

<b>Critical Incident Stress Management (CISM): An Effective Peer Support Program For Aviation Industries</b>	<b>97</b>
--	-----------

*Jeffrey T. Mitchell and Joerg Leonhardt*

<b>Changing General Aviation Flight Training by Implementing FAA Industry Training Standards</b>	<b>117</b>
--	------------

*Michele Summers Halleran and Michael E. Wiggins*

<b>The Reciprocal Development of Expertise in Air Traffic Control</b>	<b>131</b>
---	------------

*Christine Owen and Wendy Page*



<b><i>A Two-group Experiment to Measure Simulator-based Upset Recovery Training Transfer</i></b>	<b>153</b>
<i>Rodney O. Rogers, Albert Boquet, Cass Howell, and DeJohn</i>	
<b><i>Effects of Video Weather Training Products on General Aviation Pilot Weather Knowledge and Flight Behavior into Adverse Weather</i></b>	<b>169</b>
<i>William Knecht and Jerry Ball</i>	
<b><i>MD-11 Landing Incidents/Accidents 1992 – 2009: A Textual Analysis of NTSB and FAA Reports</i></b>	<b>203</b>
<i>John T. Cocklin</i>	
<b><i>Developing Next Generation Research Competencies Through Collaborative Student Design and Advanced Manufacturing Projects</i></b>	<b>219</b>
<i>Sergey Dubikovsky, Timothy D. Ropp, and Thomas Lesczynski</i>	
<b><i>The Efficacy of Flight Attendant/Pilot Communication in a Post-9/11 Environment: Viewed from Both Sides of the Fortress Door</i></b>	<b>227</b>
<i>Lori J. Brown and William G. Rantz</i>	

### ***Previously Published Technical Report***

The article within this section of the IJAAS is a reprint in whole or in part of a government technical report. The technical report was peer reviewed **as is** and is reprinted on recommendation of the peer review panel. The IJAAS should not be cited. The original publication citation is provided on the title page of the report.

<b><i>Subtle Cognitive Effects of Moderate Hypoxia</i></b>	<b>249</b>
<i>Kate J. Hewett, Ian P. Curry, and Steven J. Gaydos</i>	

## ***Individual and Environmental Predictors of Burnout among Regional Airline Pilots***

Richard O. Fanjoy and Stanley L. Harriman

*Purdue University  
1401 Aviation Drive  
West Lafayette, IN 47907-2015  
765-494-9964  
[rofanjoy@purdue.edu](mailto:rofanjoy@purdue.edu)  
[sharrima@purdue.edu](mailto:sharrima@purdue.edu)*

Randal J. DeMik  
*Lewis University  
One University Parkway, Unit #282  
Romeoville, IL 60446-2200  
818-836-5559  
[demikra@lewisu.edu](mailto:demikra@lewisu.edu)*

### ***Abstract***

*Recent accidents and incidents in regional airline operations suggest that pilot burnout and fatigue may be contributing factors that warrant further investigation and remediation. In this exploratory study, the Maslach Burnout Inventory (MBI-GS) and an environmental factors survey were administered to a sample of 248 regional airline pilots. The MBI-GS is a 22-question, Likert-scaled instrument that measures three subscales of burnout: exhaustion, cynicism, and professional efficacy. Findings from the survey instruments suggest that moderate levels of burnout currently exist within the sample population. In addition, findings suggest a relationship between elevated regional airline pilot exhaustion levels and a perception of inappropriate management pressure to make on time goals. These findings have implications for regional airline pilot job satisfaction as well as training curricula. As pilots' levels of stress and burnout increase, safe and efficient air travel may be compromised. It is therefore important to study these detractors further and develop appropriate interventions.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

## **Individual and Environmental Predictors of Burnout among Regional Airline Pilots**

Regional airlines constitute a substantial segment of the US civil air fleet. In 2008, 68 regional airlines transported 159.32 million passengers or one out of every four air travelers in the United States (RAA, 2009). With the growth of regional airlines has come a shortfall of experienced pilots that has led to the hiring of pilots with significantly less flight experience than their predecessors. At the same time, economic pressures have driven airlines to employ utilization strategies that may lead to a gradual deterioration of the quality and functionality for both equipment and personnel. A resulting concern within the commercial flight environment is the troubling phenomena of pilot fatigue/job burnout among regional airline pilots that has drawn the attention of aviation industry officials, congressional leadership, and the traveling public.

Robert Sumwalt, vice president of the National Transportation Safety Board, notes that pilot fatigue has been a key factor in commercial aircraft accidents and incidents over the last 15 years, resulting in 250 fatalities (Chen, 2009). Several recent pilot fatigue-related accidents and incidents have been prominently featured in the media, including: a Pinnacle Airlines runway overrun in Michigan (2007), a Go! Airlines overflight of Hilo, Hawaii attributed to pilots who had fallen asleep in the cockpit (2007), and the Colgan Air crash in Buffalo (2009). The media has suggested that pilot fatigue and burnout are undesirable byproducts of economic pressures to improve airline performance. Although airline officials have assured the traveling public that pilots are well trained and encouraged to keep safety first, when making decisions regarding the safe conduct of flight, it appears that subtle company pressures associated with continued employment frequently override common sense decision making that has been the hallmark of industry pilots. In an earlier survey of 1,424 regional airline pilots, 89 percent of the respondents identified fatigue associated with flight operations as a moderate to serious concern (Co, Gregory, Johnson, & Sosekind, 2000). Regional pilots who fly for airlines from other nations echo safety concerns seen in the U.S. The president of the Air Canada Pilots Association notes "pilots are flying more hours a month, more days a month," leading to the increased likelihood of fatigue (Butler, 2008). A Belgian pilot source notes subtle management pressure to accept aircraft with minor mechanical malfunctions reflecting European Union laws recently introduced that require compensation for passengers whose flight is delayed or cancelled (Pilots 'under pressure to take risks,' 2005). In addition to pressure for on-time takeoffs and landings, there is pressure to keep the airplanes in the air as much as possible and thereby enhance an airline's profit margin. Finally, pilots are frequently pressured to accept aircraft with reduced levels of fuel on board in order to accommodate additional cargo or to reduce fuel burn rates with correspondingly lighter aircraft. All of these pressures produce stress and mental fatigue in pilot crews.

A recent study of pilots who work for major airlines suggested that work overload, organizational politics, and an undervalued reward system were significantly correlated with job burnout factors (Kearney, 2008). According to Bennett (2003), pilots interviewed from a UK-registered low-cost carrier reported that they felt stressed and fatigued and that these issues required close attention in the industry. In a later article, Bennett (2006) concluded that pilots flying at a low-cost carrier working in a complex and challenging environment were continuing to ex-

perience stress and fatigue. An ongoing study of job burnout within the global civil aviation industry is currently being conducted by International Transport Workers Federation researchers to evaluate the impact of long working hours and precarious employment practices on fatigue (ITWF, 2009). Dr. Michael Bagshaw, head of aviation medical services for British Airways, suggested that automated flight systems are an additional cause of pilot fatigue and stress. He believes technology associated with evolving automated flight systems may be approaching the mental processing limits of assigned flight crewmembers (McConnell, 1997). Each of the several environmental stresses within the airline industry has potential for contributing to a significant level of job burnout. The present exploratory study was developed to investigate burnout levels within the regional airline pilot ranks.

## **Job Burnout**

Burnout has been described as a syndrome of physical and emotional exhaustion resulting from the development of negative self-concept, negative job attitudes, and reduced personal accomplishment (Maslach, 1976; Maslach & Jackson, 1981; Truchot, Keirsebilick, & Meyer, 2000). More specifically, burnout is defined as “a prolonged response to chronic emotional and interpersonal stressors on the job” (Maslach, 2003, p.189). The subscales of burnout have been identified as exhaustion, cynicism, and professional efficacy. According to Maslach, Jackson, and Leiter (1996), an elevated exhaustion level prompts actions to distance oneself emotionally and cognitively from one’s work. The elevated cynicism aspect is reflected as an indifference towards work. Finally, a depressed professional efficacy level reflects feelings of incompetence and a lack of achievement and productivity at work. The development of burnout that occurs from a combination of these factors may result in physical, psychological, and behavioral symptoms.

Several studies (Kestnbaum, 1984; Farber, 1990; Piercy & Wetchler, 1987; Pines & Aronson, 1981) have described the physical symptoms of burnout as headaches, difficulty sleeping, gastrointestinal problems, hypertension and/or a poor appetite. Psychological symptoms may include increased negative self-talk, depression, boredom, stress, cynicism, anxiety, irritability and/or difficulty in interpersonal relationships. Behavioral symptoms may include diminished attention, increased absenteeism, and attrition. Consequently, increased levels of stress and burnout can lead to ineffective job performance, exhaustion, physical complaints, anxiety, depression, and substance abuse (Huebner & Huberty, 1984; Pines & Aronson, 1981). The concern of this study is that as burnout is identified, there may be implications for regional airline pilot job satisfaction, flight safety, and expanded professional training needs.

## **Burnout in Other Occupations**

Previous research into high stress occupational populations such as policemen, firemen, and medical staff has identified common stressor factors that include a hierarchical organizational structure, high environmental risk, a 24-hour job rotation system, and responsibilities in emergency/life-or-death situations (Winick, Rothacker, & Norman, 2002). Studies of the nursing career field have found job stress issues associated with limited resources, conflicts with colleagues or physicians, working in life-or-death situations, excessive workload, professional knowledge and technical requirements, handling patient’s emotional needs, and an inability to succeed in a professional role (Dewe, 1987; Guppy & Gutteridge,

1991; Harris, 1989). Related studies have found that nurses experiencing higher levels of burnout were judged independently by their patients to be providing a lower level of patient care (Leiter, Harvie, & Frizzell, 1998; Vahey, Aiken, Sloane, Clarke, & Vargas, 2004). Researchers found that job stress for teachers resulted from external or school demands, pressure from students, the degree of school support, school policies, and personal factors, such as the lack of professional development (McCormick & Solman, 1992). In many work environments, coupled with the difficulty of attaining promotion, employees who lack confidence in their future may develop a mindset of failure, which in turn may lead to job burnout (Hall, 1986).

### **Lack of Professional Development/Autonomy**

In research related to job burnout, occupations with direct public contact have attracted much more attention than flight crews. The primary duty of flight crews is to conduct routine flight services, which frequently lacks challenge, complicity, and autonomy (Liang & Hsieh, 2005). If individuals perceive their job as lacking challenge and autonomy, their commitment to objectives and efforts can potentially be affected. Moreover, a job lacking feedback could directly influence the psychological sensation of success, which could, in turn, alter a person's work attitude. Pines and Yanai (2001) adopted the psychodynamic-existential perspective in explaining that job burnout stems from people's inability to secure the value of survival on their job or career as well as the meaning and importance of working. With respect to the examination of job stress from the perspective of occupations with job characteristics somewhat similar to those of flight crews, it was discovered in a study by Duffy and McGoldrick (1990) that the major stressors for bus drivers are route timing, traffic problems, passenger and company conflicts, work protection, health, and family problems. In addition to common work pressures or demands, there may be specific job characteristics that will affect workers' level of burnout, such as shift work (Jamal, 2004).

### **Work Overload/Lack of Resources**

A critical aspect of burnout occurs when people are unable to recover from daily work demands. When this kind of stress is a chronic job condition, not an occasional emergency, there is little opportunity to recover and restore balance (Maslach & Leiter, 2008). A condition of work overload or lack of resources can be frequently attributed to management practices of the organization. A study of work-related stress in American police officers indicated that organizational factors such as bad management or work conditions were more frequently identified as negative stressors than potential violence or exposure to human misery (Storch & Panzarella, 1996). In a study of Scottish police officers, the highest levels of stress were related to organizational factors such as perceived staff shortage and inadequate resources (Biggam, Power, MacDonald, Carcary, & Moodie, 1997). The high workload of regional airline pilots, that results from the large number of takeoffs and landings associated with several legs per day and irregular work hours, may lead to fatigue symptoms (Neider, Vejvoda, & Mass, 2008). Weitzel and Hampton (1999) suggest that current work schedules for regional airline pilots and flight crews are not compatible with human limitations and existing Federal Aviation Regulations have not evolved to meet the demands of a 24-hour industry.

## **Burnout Crossover Phenomenon**

If there is no appropriate intervention for employee burnout within the organization, burnout crossover may occur and further reduce productivity or safety practices. Crossover occurs when psychological strain experienced by one person affects the level of strain for another person who shares the same social environment (Bolger, DeLongis, Kessler, & Wethington, 1989; Westman & Etzion, 1995). Previous findings from field research by Bakker and Schaufeli (2000) suggest that individuals suffering from burnout may communicate symptoms to their colleagues and that social comparison processes may play a role in the transmission of burnout symptoms among employees. Groenestijn, Buunk, & Schaufeli (1992) found that nurses who were aware of burnout complaints from their colleagues and felt a strong need for social comparison were more susceptible to burnout compared to those who had a low need for social comparison. Bakker, Westman, & Schaufeli (2007) investigated burnout crossover in high school teachers and found that when teachers were exposed to an interview with a colleague who talked negatively about their students and teaching, they reported significantly higher levels of emotional exhaustion and depersonalization than teachers who were exposed to an interview that was negative in tone, but unrelated to work. Evidence for burnout crossover was also found among general practitioners and constabulary officers (Bakker, Schaufeli, Sixma, & Bosveld, 2001; Bakker, van Emmerik, & Euwema, 2006). While on duty days, regional pilots spend a considerable amount of time within confined spaces in team settings both on the job and during layovers. The result may be a working environment that perpetuates this burnout crossover phenomenon.

## **Perceived Job Risks and the Effects of ‘Pilot Pushing’**

In addition to actual exposure to stressors, a perception of risk may also increase burnout. According to risk perception theory, perceptions of risk are made up of a cognitive component such as the probability of being injured and an affective component such as an emotional reaction to the risk that is manifested as worry (Rundmo, 2002). Baugher and Roberts (1999) assessed the extent to which petrochemical plant employees worried about two types of hazards: risk of fire/explosions (i.e., acute events) and exposure to chemicals (i.e., a chronic hazard for this occupation). They found that despite the infrequency of acute events and despite a substantially low degree of perceived risk of these events, petrochemical workers still had a high degree of worry about them. ‘Pilot pushing’ is the pressure that pilots face from management to keep airplanes in the air as much as possible by agreeing to fly legs with critical equipment problems, in severe weather, with reduced fuel requirements, or in a state of fatigue. Pilot pushing may result in chronic stress that has a significant impact on many health outcomes (Day, Therrien, & Carroll, 2005). In addition to these questionable management pressures, the emphasis of worrying about the safe outcome of flights and hazards is important because some of the risks faced by pilots have the potential to be severe and even fatal. Based on the literature of affective risk perceptions, continuing pressure to operate with a perceived unreasonable level of risk may lead to elevated levels of burnout (Day, Sibley, Scott, Tallon, & Ackroyd-Stolarz, 2009).

Control of one’s job environment is associated with many organizational and individual outcomes (de Lange, Taris, Kompier, Houtman, & Bongers, 2003). Studies suggest that such control has resulted in lower burnout levels for nurses and



physicians (Aiken, Clarke, Sloane, Sochalski, & Silber, 2002; Keeton, Fenner, Johnson, & Hayward, 2007). Job control may have an even greater role in high risk industries, such as medical flight transport, where workplace hazards are prevalent. Increasing employee control over their work functions in these industries is important because it may allow employees to reduce uncertainty levels associated with work-related hazards and risks (Leiter & Robichaud, 1997). Therefore, increased employee job control may promote decreased levels of burnout, and may also buffer the negative consequences of safety incidents and concerns of burnout (Day et al., 2009). If airline management provides pilots with more work autonomy, pilots may be able to exert more influence over the stress-provoking areas of their work life (Dwyer & Ganster, 1991).

## Research Questions

The following research questions were addressed by the present study:

*What are the perceived levels of exhaustion, cynicism, and burnout among regional airline pilots as measured by the Maslach Burnout Inventory-General Survey (MBI-GS)?*

*Is there a relationship between Individual characteristics (flight time, age, education level, duty schedule, and crewmember position) and levels of burnout among regional airline pilots?*

*Is there a relationship between burnout factors and regional airline pilot perceptions of management pressure to meet on-time goals?*

## Methodology

### Participants

A survey instrument was administered to a sample of U.S.-based regional airline pilots ( $N = 248$ ) to assess levels of occupational burnout. Participants did not receive payment for participating in the study and were advised that information was collected anonymously. This human subjects study was approved by the sponsoring university's Institutional Review Board.

### Measures

The Maslach Burnout Inventory-General Studies (MBI-GS), consisting of 22 questions that are used to measure the symptoms of burnout, was administered. The MBI-GS is designed to assess three aspects of burnout (Exhaustion [EX], Cynicism [CY], and Professional Efficacy [PE]) using a Likert-type response scale to indicate how often the participants experienced a given thought or feeling. Responses range from 0 (never) to 6 (daily). The construct validity of the MBI-GS has been supported through validity studies (Kitaoka-Higashiguchi, Ogino, & Masuda, 2004; Langballe, Falkum, Innstrand, & Aasland, 2006) that suggest the MBI-GS provides a suitable measurement to assess burnout across a wide range of professions. For the purposes of this study, minor modification of wording to match the regional airline pilot work environment was necessary. For example, "work" and "organization" as referred to in the original questions, were changed to "flight duties" and "airlines" respectively.



The last section of the questionnaire was designed to measure the perception of pressure from airline management to complete a flight with questionable safety risks or hazards. This included: accepting aircraft with critical equipment problems, starting or continuing flight into severe weather or icing conditions, accepting critically reduced fuel requirements to accommodate revenue, and overall pressure to make on-time goals. Constructed in a format similar to the MBI-GS items, answers to these items indicated how often the respondents experienced this type of pressure, ranging from 0 (never) to 6 (daily).

## Procedures

Study participants were recruited through written requests for research participation in the *Flight Safety Information Newsletter* with a distribution of over 25,000 aviation subscribers/readers. A description of the study included a website address, which upon accessing, allowed the participants to view an electronic questionnaire, represented written consent, and confirmed their volunteer status. Participants completed a 36-question survey which first asked respondents to describe individual and environmental variables, including: age, highest educational degree obtained, current flight status, years as a regional airline pilot, total flight time, type of aircraft flown, work schedule, and salary.

## Analysis

### Biographical data

A total of 248 surveys were submitted of which 55 were incomplete and omitted from the analysis. Descriptive data was compiled for the remaining 193 surveys. The largest percentage of respondents (38.5 %) was aged 27-32, followed by the next highest age group, 33-38 (21 %). The majority of respondents had completed at least a four year degree college program (83.4 %). The majority of the respondents were currently operating a regional jet variant (85.9 %). Current flight status of respondents was somewhat split between captain (51.2 %) and first officer (43.1 %). The remaining small percentage of pilots was employed primarily in management. Employment time with regional carriers, for the most part, fell into three primary year groups: 1 to 4 years (33.8 %), 5 to 8 years (34.9 %), and 9-12 years (12.3 %). Total flight time of respondents varied widely, but the largest number of respondents (51.3 %) indicated that they had accumulated more than 5,000 hours. Monthly work schedules for respondents varied widely, but most (84.9 %) indicated they had 11 to 15 days off per month. Hours flown per month also varied widely among respondents, but the largest group (69.8 %) indicated that they averaged 78-89 flight hours per month. Salaries varied widely but the two largest groups of respondents were paid \$30K to \$40K per year (22.6 %) and more than \$70K per year (34.4 %). Descriptive data are presented in the Appendix.

### MBI-GS data

Using the MBI-GS analysis rubric, levels of exhaustion (EX), cynicism (CY), and professional efficacy (PE) were calculated by determining the average Likert response for corresponding MBI-GS questions. Survey questions pertaining to *exhaustion* address "feeling drained," "feeling used up," "feeling tired in the morning," and "feeling burned out." Questions pertaining to *cynicism* address "feeling less interested," "feeling less enthusiastic," "wanting to be not bothered," "doubting

significance to the employing airline,” and “being cynical about work.” Questions pertaining to *professional efficacy* address “effective at problem solving,” “making a contribution,” “doing a good job,” “feeling exhilarated,” and “having a sense of accomplishment.” According to the literature, individuals who exhibit high levels of both EX and CY as well as low levels of PE are considered strong candidates for burnout designation. Levels that contribute to burnout, according to the validated MBI-GS rubric were average EX of 4.2 and higher, average CY of 2.2 and higher, and average PE of 4.0 and lower (Maslach, Jackson, & Leiter, 1996). Upon analysis, 32.6 percent (63) of the sample population were identified as high burnout candidates. In addition, 51.8 percent of the sample was identified as exhibiting high exhaustion levels and 72.5 percent exhibited high cynicism levels. An additional 53.8 % of the sample exhibited low professional efficacy levels.

A bivariate correlation procedure was completed between the three MBI-GS factors (EX, CY, and PE) and the various biographical data responses using Predictive Analytics Software (PASW). Significant ( $p < .01$ ) correlations with level of *exhaustion* were found for age, years of work, hours flown per month, and current salary. A significant correlation ( $p < .05$ ) was found for level of cynicism and professional efficacy for hours flown per month. No other significant ( $p < .05$ ) correlations with biographical data were found for *exhaustion*, *cynicism*, and *professional efficacy* levels. See Table 1.

Table 1

*Bivariate Correlation – Burnout Factors vs. Biographical Data (n=193)*

Effect	EX	CY	PE
Age	-.256*	-.140	.096
Years worked	-.194**	-.054	.099
Total flight time	-.161*	.046	.000
Days off/month	.002	-.115	.083
Flight hours/month	.200**	.174*	-.149*
Current salary	-.220**	-.046	.040

Note: \*  $p < .05$ , \*\*  $p < .01$

Upon completion of the bivariate correlation, one-way ANOVAs with Tukey’s post-hoc pair wise comparisons were completed to further evaluate the relationships between previously identified biographical groups and burnout factors. As a result of this analysis, a significant difference was noted between mean *exhaustion* levels of pilots over 50 and pilots aged 21 to 26 ( $p < .000$ ), 27-32 ( $p < .001$ ), or 33-38 ( $p < .000$ ). No other significant differences in means were noted for burnout factors with other biographical data groupings.

Next, a bivariate correlation procedure was completed with the burnout factors and questions regarding management pressure to make on-time goals. Significant correlations ( $p < .01$ ) were identified for *exhaustion* and *cynicism* with management pressure to make on time goals through: shortened rest periods, accepting aircraft with critical equipment problems, and starting or continuing flight into severe weather conditions. Similar correlations at the  $p < .05$  level were noted for profes-

sional efficacy. See Table 2 for these results. No correlations were noted between the burnout factors and management pressure to accept reduced fuel loads.

Table 2

Bivariate Correlation –Burnout Factors vs. Management Pressure to meet On-time Goals (n=193)

Effect	EX	CY	PE
<b>Management Pressure:</b>			
Shortened Rest	.584**	.431**	-.172*
Equipment problems	.367**	.331**	-.166*
Severe Weather	.496**	.404**	-.146*
Reduced fuel loads	-.009	.074	-.142

Note: \* $p < .05$ , \*\* $p < .01$

Finally, a GLM univariate procedure was completed to model biographical data and management pressure variables identified in the preceding analysis as predictive factors for levels of exhaustion and cynicism. The resulting model to predict exhaustion had an adjusted  $r^2$  of .443. A similar procedure was conducted to model cynicism resulted in an adjusted  $r^2$  of .285. Removing additional non-significant variables from these models resulted in a decreased  $r^2$ . The analysis did not suggest a useful model for professional efficacy. These results are depicted in Tables 3 and 4.

Table 3

Univariate Model for Exhaustion (n=193, adjusted  $r^2 = .443$ )

Factor	df	F	$p$
Flight w/shortened rest	1	46.226	.000**
Flight w/adverse weather	1	13.224	.000**
Flight position, Capt-F/O	1	8.045	.005**
Age	1	6.557	.011
Years worked for airline	1	2.370	.125

Note: \*\* $p < .01$

Table 4

*Univariate Model for Cynicism (n=193, adjusted r<sup>2</sup>=.285)*

Factor	df	F	p
Flight w/shortened rest	1	14.945	.000**
Flight w/adverse weather	1	6.405	.012
Days off/month	1	4.843	.029
Age	1	4.379	.038
Overall Mgt Pressure	1	3.001	.085
Years worked for airline	1	3.000	.085
Flight position, Capt-F/O	1	2.8	.097
Flight w/reduced fuel	1	2.380	.125
Level of Education	1	1.764	.186
Total flight time	1	1.614	.206

Note: \*\* $p < .01$

## Discussion

This study was conducted to investigate current levels of burnout among regional airline pilots. Although study limitations in the area of sample selection and instrumentation bias restrict generalization of the results, the analysis suggests relationships for further study with a larger sample. The current study should be considered a pilot for such studies. The analysis of current MBI-GS survey results, in particular, points to significantly elevated levels of exhaustion, cynicism, and burnout, as well as reduced levels of professional efficacy, that may be present among the general population of regional airline pilots. The analysis also suggests that mean levels of exhaustion are very different for pilots over 50 and their younger counterparts. The number of current survey participants in the 21 to 26 (21%), 27 to 32 (16%) and 33 to 38 (22%) age groups exhibiting high levels of exhaustion was much greater than their over 50 (0%) counterparts. Data also supports a higher average number of hours flown per month for the younger age groups. Whether this circumstance is a management prerogative or based on seniority of older pilots is unclear. However, there seems to be a threshold of flight hours per month that contributes to higher levels of exhaustion.

A relationship might be hypothesized between higher levels of cynicism and more senior pilots who have been with regional airlines longer (or who had accumulated more total flying hours). This conclusion was not supported by analysis of the data which suggests that more respondents from the 27 to 32 (77%) and 33 to 38 (80%) age groups exhibited higher levels of cynicism than respondents in the over 50 (38%) group. In addition, the analysis suggests that respondents from the 27 to 32 (60%) and 33 to 38 (61%) age groups exhibited lower levels of professional efficacy than their counterparts in the over 50 group (38%). Each of these factor levels are based on self-reported answers to survey questions.

Further analysis was completed to identify relationships between burnout factor levels and respondent answers to questions about management pressure to meet on time goals. The survey asked respondents to comment on their perception of management pressure to fly with excessively shortened rest periods, critical equipment problems, severe weather conditions, or reduced fuel requirements. The analysis shows a significant relationship between levels of exhaustion, cynicism, and burnout, and perceived management pressure to continue flight with shortened rest periods, critical equipment problems, and severe weather conditions. Although survey responses may have been influenced by a presumed relationship between continued employment and acceptance of the described flight conditions, the analysis suggests there is, in fact, a relationship between perceptions of management pressure and burnout factors. This finding seems to highlight a need for further investigation into pilot perceptions of acceptable flight practices and what reality exists regarding inappropriate management pressure to meeting on time flight goals.

### **Conclusion**

A large body of literature details investigations into the causes of burnout in various occupations and the effect of this condition on job performance. This study represents an initial look into environmental factors that may impact levels of exhaustion, cynicism, professional efficacy, and ultimately burnout, of regional airline pilots. Although findings of this study do not conclusively identify environmental factors responsible for burnout in the subject population, they do suggest areas for further investigation. Further investigation may include a larger sample size, ethnographic research as a means of understanding the lived reality of the regional pilots, or include the results of this study with other published studies to provide a meta-analysis. The researchers acknowledge that there may be sample bias due to the small sample size; nevertheless, this pilot study has been useful to develop, adapt, and check the feasibility of techniques. In particular, study findings suggest shortened rest periods, adverse weather, aircraft maintenance issues, and pressures to meet on time performance goals are potential contributors to pilot burnout and resulting safety concerns. This study highlights an opportunity to review subtle and not-so-subtle management pressure on flight crews to meet performance goals as well as pilot perspective on implied and actual authority for flight conduct. Further work in the area of management and pilot training is indicated. This level of attention is critical given the recent trend of commercial aircraft accidents and incidents as well as the eroding public perception of commercial air travel.

## References

- Aiken, L. H., Clarke, S. P., Sloane, D. M., Sochalski, J., & Silber, J. H. (2002). Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *Journal of American Medical Association*, 288, 1987-1993.
- Bakker, A. B., & Schaufeli, W. B. (2000). Burnout contagion processes among teachers. *Journal of Applied Social Psychology*, 30, 2289-2308.
- Bakker, A. B., Schaufeli, W. B., Sixma, H., & Bosveld, W. (2001). Burnout contagion among general practitioners. *Journal of Social and Clinical Psychology*, 20, 82-98.
- Bakker, A. B., Westman, M., & Schaufeli, W. B. (2007). Crossover of burnout: An experimental design. *European Journal of Work and Organizational Psychology*, 16(2), 220-239.
- Bakker, A. B., van Emmerick, I. J. H., & Euwema, M. C. (2006). Crossover of burnout and engagement in work teams. *Work and Occupations*, 33, 464-489.
- Baugher, J. E., & Roberts, J. T. (1999). Perceptions and worry about hazards at work: Unions, contract maintenance, and job control in the U.S. petrochemical industry. *Industrial Relations*, 38(4), 522-541.
- Bennett, S.A. (2003). Flight crew stress and fatigue in low-cost commercial air operations – an appraisal. *International Journal of Risk assessment and Management*. 4(2), 207-231.
- Bennett, S.A. (2006). A longitudinal ethnographic study of aircrews' lived experience of flying operations at a low-cost airline. *Risk Management: An International Journal*. 8(2), 92-117.
- Biggam, F. H., Power, K. G., MacDonald, R. R., Carcary, W. B., & Moodie, E. (1997). Self-perceived occupational stress and distress in a Scottish police force. *Work and Stress*, 11, 118-133.
- Bolger, N., DeLongis, A., Kessler, R., & Wethington, E. (1989). The contagion of stress across multiple roles. *Journal of Marriage and the Family*, 51, 175-183.
- Butler, D. (2008, January 31). More pressure than ever on pilots, union reps say. *The Ottawa Citizen*. Byline: Don Butler, Source: The Ottawa Citizen, Page: A3,
- Chen, S. (2009, May 15). Pilot fatigue is like 'having too much to drink'. Retrieved from: <http://www.cnn.com/2009/TRAVEL/05/15/pilot.fatigue.buffalo.crash/index.html>
- Co, E.L., Gregory, K.B., Johnson, J.M., & Sosekind, M. R. (2000). Crew factors in flight operations: a survey of fatigue factors in regional airline operations. *STAR*, vol. 38.
- Day, A. L., Therrien, D., & Carroll, S. A. (2005). Predicting psychological health: Assessing the incremental validity of emotional intelligence beyond personality, Type A behavior, and daily hassles. *European Journal of Psychology*, 19(6), 519-536.

- Day, A. L., Sibley, A., Scott, N., Tallon, J. M., & Ackroyd-Stolarz, S. (2009). Workplace risks and stressors as predictors of burnout: The moderating impact of job control and team efficacy. *Canadian Journal of Administrative Sciences*, 26, 7-22.
- de Lange, A. H., Taris, T. W., Kompier, M. A. J., Houtman, I. R. D., & Bongers, P. M. (2003). The very best of the millennium: Longitudinal research and the demand-control (support) model. *Journal of Occupational Health Psychology*, 8, 282-305.
- Dewe, P. J., (1987). Identifying the causes of nurse's stress: A survey of New Zealand nurses. *Work and Stress*, 1, 15-24.
- Duffy, C. A., & McGoldrick, A. E. (1990). Stress and the bus driver in the UK transport industry. *Work and Stress*, 4, 17-27.
- Dwyer, D. J., & Ganster, D. C. (1991). The effects of job demands and control on employee attendance and satisfaction. *Journal of Organizational Behavior*, 12, 595-608.
- Farber, B. A. (1990). Burnout in psychotherapists: Incidence, types, and trends. *Psychotherapy in Private Practice*, 8(1), 35-44.
- Groenestijn, E., Buunk, A. P., & Schaufeli, W. B. (1992). The danger of burnout contagion: About the role of social comparison processes. *Sociale Psychologie en Beïnvloeding van Intermenselijke en Gezondheidsproblemen* (pp. 88-103). The Hague, The Netherlands: VUGA.
- Guppy, A., & Gutteridge, T. (1991). Job satisfaction and occupational stress in UK general hospital staff. *Work and Stress*, 5, 315-323.
- Hall, D. T. (1986). Breaking career routines: Midcareer choice and identity development. *Frontiers of industrial and organizational psychology: Vol. 1. Career development in organizations* (1<sup>st</sup> ed., pp. 120 – 159). San Francisco: Jossey-Bass.
- Harris, P. E. (1989). The nurse stress index. *Work and Stress*, 3, 335-346.
- Huebner, E. S., & Huberty, T. J. (1984). Burnout among rural school psychologists. *Research in Rural Education*, 2, 95-99.
- International Transport Workers' Federation (2009). Stressed and fatigued on the ground and in the sky: Changes from 2000-2007 in civil aviation workers' conditions of work. Retrieved from: [http://www.ifatca.org/docs/itf\\_141109.pdf](http://www.ifatca.org/docs/itf_141109.pdf).
- Jamal, M. (2004). Burnout, stress, and health of employees on non-standard work schedules: A study of Canadian workers. *Stress and Health*, 20, 113-119.
- Kearney, L.. (2008). Examining the relationship between Part 121 air transport pilots and burnout. Ph.D. dissertation, Northcentral University, United States -- Arizona. Retrieved August 6, 2010, from Dissertations & Theses: Full Text. (Publication No. AAT 3337594).
- Keeton, K., Fenner, D. E., Johnson, T. R. B., & Hayward, R. A. (2007). Predictors of physician career satisfaction, work-life balance, and burnout. *Obstetrics and Gynecology*, 109, 949-955.



- Kestnbaum, J. D. (1984). Expectations for therapeutic growth: One factor in burnout. *Social Casework: The Journal of Contemporary Social Work*, 65, 374-377.
- Kitaoka-Higashiguchi, K., Ogino, K., & Masuda, S. (2004). Validation of a Japanese research version of the Maslach Burnout Inventory-General Survey. *Shinrigaku Kenkyu*, 75 (5), 415-419.
- Langballe, E.M., Falkum, E., Innstrand, S.T., & Aasland, O.G. (2006). The factorial validity of the Maslach burnout inventory-general survey in representative samples of eight occupational groups. *Journal of Career Assessment*, 14, 370-385.
- Leiter, M. P., & Robichaud, L. (1997). Relationships of occupational hazards with burnout: An assessment of measures and models. *Journal of Occupational Health Psychology*, 2(1), 35-44.
- Leiter, M. P., Harvie, P., & Frizzell, C. (1998). The correspondence of patient satisfaction and nurse burnout. *Social Science & Medicine*, 47, 1611-1617.
- Liang, S., & Hsieh, A. (2005). Individual's perception of career development and job burnout among flight attendants in Taiwan. *The International Journal of Aviation Psychology*, 15(2), 119-134.
- Maslach, C. (1976). Burned-out. *Human Behavior*, 5(9), 16-22.
- Maslach, C. (2003). Job burnout: New directions in research and innovation. *Current Directions in Psychological Science*, 12, 189-192.
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior*, 2, 99-113.
- Maslach, C. M., Jackson, S. E., & Leiter, M. P. (1996). *Maslach Burnout Inventory* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.
- Maslach, C., & Leiter, M. P. (2008). Early predictors of job burnout and engagement. *Journal of Applied Psychology*, 93(3), 498-512.
- McConnell, H. (1997, March 11). Hi-tech modern airliner can push pilot to crash, burnout. *Medical Post*, 33 (10), p.52.
- McCormick, J., & Solman, R. (1992). The externalized nature of teachers' occupational stress and its association with job satisfaction. *Work and Stress*, 6, 33-44.
- Neiderl, T., Vejoda, M., & Maass, H. (2008). Part 1: Investigation on cumulative psycho physiological effects of short-haul flight crew members (results). *Aviation, Space, and Environmental Medicine*, 79(3), 311.
- Piercy, F. P., & Wetchler, J. L. (1987). Family-work interfaces of psychotherapists. *Journal of Psychotherapy and the Family*, 3, 17-32.
- Pilots 'under pressure to take risks'. (2005, August 22). *Telegraph*. Retrieved from: <http://telegraph.co.uk/travel/733416/Pilots-under-pressure-to-take-risks.html>.
- Pines, A. M., & Aronson, E. (1981). *Burnout: From tedium to personal growth*. New York: Free Press.

- Pines, A. M., & Yanai, O. Y. (2001). Unconscious determinants of career choice and burnout: Theoretical model and counseling strategy. *Journal of Employment Counseling, 38*, 170-184.
- Regional Airline Association (2009). US regional airline fact sheet. Retrieved from [http://www.RAA.org/images/stories/fact\\_sheet\\_042709.pdf](http://www.RAA.org/images/stories/fact_sheet_042709.pdf)
- Rundmo, T. (2002). Associations between affect and risk perception. *Journal of Risk Research, 5*(2), 119-135.
- Storch, J. E., & Panzarella, R. (1996). Police stress: State-trait anxiety in relation to occupational and personal stressors. *Journal of Criminal Justice, 24*, 99-107.
- Truchot, D., Keirsebilck, L., & Meyer, S. (2000). Communal orientation may not buffer burnout. *Psychological Reports, 86*, 872-878.
- Vahey, D. C., Aiken, L. H., Sloane, D. M., Clarke, S. P., & Vargas, D. (2004). Nurse burnout and patient satisfaction. *Medical Care, 42*, 57-66.
- Weitzel, T., & Hampton, S. (1999). Fatigue – Investigation of a human factor for regional airline pilots. *International Symposium on Aviation Psychology, 10<sup>th</sup>, Columbus, OH; United States; 3-6 May 1999*. Pp. 873-878.
- Westman, M., & Etzion, D. (1995). Crossover of stress, strain, and resources from one spouse to another. *Journal of Organizational Behavior, 16*, 169-181.
- Winick, C., Rothacker, D. Q., & Norman, R. L. (2002). Four worksite weight loss programs with high-stress occupations using a meal replacement product. *Occupational Medicine, 52*, 25-30.

## APPENDIX

### Regional Airline Pilot Survey Sample Descriptive Data (n=193)

Descriptor	Group	% of sample
Age	27-32	38.9
	33-38	21.3
	Other	39.8
Education	4 yr. Degree min.	83.4
	Other	16.6
Current equipment	Regional jet variant	85.5
	Other	14.5
Current flight status	Capt	51.4
	F/O	44.0
	Other (mgt pilots)	4.6
Years employed	1-4	34.4
	5-8	35.4
	9-12	12.5
Total flight hours	less than 2000	7.8
	2000-4000	21.9
	4000-5000	18.2
	More than 5000	52.1
Days off/month	Less than 11 days	7.8
	11-15 days	84.9
	More than 15 days	7.3
Hours flown/month	less than 72 hours	18.2
	72-89 hours	70.9
	More than 89 hours	10.9
Current Salary	Less than \$30K	14.1
	\$30K-69K	57.0
	More than \$69K	34.9

## ***Pilots, Controllers and Mechanics on Trial: Cases, Concerns and Countermeasures***

Sidney Dekker

*Leonardo da Vinci Laboratory for Complexity and Systems Thinking  
Department of System Safety  
Lund University  
LTH HS 3  
P.O. Box 118  
221 00 Lund  
Sweden*

[sidney.dekker\[at\]lthhs.lu.se](mailto:sidney.dekker[at]lthhs.lu.se)

+46-435-445434

+46-435-45464 (fax)

### *Abstract*

*This paper examines criminal prosecutions of pilots, air traffic controllers, and maintenance technicians in the wake of aviation incidents and accidents worldwide, which points to an accelerating criminalization trend over the past fifteen years. It examines the concerns surrounding criminalization by considering its wisdom, fairness, and utility, taking into account the consequences for both the affected individual and the aviation industry as a whole. It concludes by reviewing the diversity of countermeasures that are currently being developed in aviation and assesses the possibility of mitigating the criminalization trend.*

## **Pilots, Controllers and Mechanics on Trial: Cases, Concerns, and Countermeasures**

There is increasing concern about pilots, controllers, and mechanics facing trial in the wake of incidents and accidents (Esler, 2009; Michaels, 2008; North, 2002; Ter Kulle, 2004; Thomas, 2007). Even though criminal prosecution has followed aviation accidents in the past, it now has become an automatic response to accidental death (or even just risk of death) in many countries (ICAO, 2007). Prosecution is often seen by those inside a profession as unfair, unnecessary, intrusive and “heavy handed” (Moran, 2008) as well as detrimental for safety initiatives aimed at increasing honest disclosure and the free flow of safety information (FSF, 2006; GAIN, 2004; ICAO, 2007).

The basis for responding to, and learning from, accidents in aviation is provided by Annex 13 to the ICAO convention. This represents an international treaty of all UN member countries, which establishes the purpose of investigations and protects those safety investigations for learning and system improvement only. Co-mingling safety investigations with criminal prosecutions is something that Annex 13 explicitly guards against, and ICAO member states are in principle obliged to enact its standards through their own regulatory and legal systems. There is, however, growing evidence of cracks and holes in the wall that putatively separates safety investigation from judicial probes. Data gathered by independent safety investigations has been appropriated by judicial action, and formal accident reports are used routinely either as evidence in court or as preparatory reading for prosecutors and judges. Despite the clear proscriptions in international treaties, professional arrangements and even national codes, Flight Data Recorder (FDR) and Cockpit Voice Recorder (CVR) records have been admitted as evidence in criminal prosecutions in several ICAO countries (Michaelides-Mateou & Mateou, 2010; North, 2002).

Prosecution of pilots, controllers and mechanics is often based on general hazard statutes that have evolved from road traffic laws which criminalize the reckless endangerment of people or property (Esler, 2009; Tingvall & Lie, 2010). The leeway in such statutes for what can be considered sanctionable behavior is of course important for any open and democratic justice system. Nevertheless, it has led to very general risk statutes (such as Netherlands Aviation Act §5.3), which can, depending on prosecutorial ambition, criminalize anything that can be construed as dangerous in hindsight. US Federal Aviation Regulation (FAR) 91.13 holds this potential, for example, though it has not been extensively relied on for criminal prosecution of pilots or controllers:

- a) No person may operate an aircraft in a careless or reckless manner so as to endanger the life and property of others.
- b) No person may operate an aircraft, other than for the purpose of air navigation, on any part of the surface of an airport used by aircraft for air commerce (including areas used by those aircraft for receiving or discharging persons or cargo), in a careless or reckless manner so as to endanger the life and property of others.

The criminalization of error in aviation, particularly through criminal categories such as “causing air disaster” (RTE, 2009), may represent a jurisprudential evolution similar to that of “hate crime,” which went from a broad, amorphous social

concept to a determinate legal construct inside of a few decades through judicial rhetoric and successive jurisprudential meaning-making (Jacobs & Henry, 1996; Phillips & Grattet, 2000). Groups from inside the aviation industry suggest that this represents “overcriminalization” (Garland, 2002; Husak, 2008).

This paper presents data that supports the notion of a criminalization trend in aviation, and details the concerns about prosecuting practitioners (in part by borrowing from research in other fields). These are divided up as concerns about the wisdom, the fairness, and the utility of criminally prosecuting practitioners. It concludes with a review of current and possible countermeasures.

### Cases

In 1956, an Air France captain was convicted of involuntary manslaughter after 56 passengers were killed in a DC-6 visual approach accident at Cairo airport on a flight en route from Saigon to Paris (Esler, 2009). Since then, criminalization of pilots, controllers, and mechanics has occurred or is occurring in many other countries, including the U.S., the U.K., Japan, New Zealand, China, Libya, Korea, Yugoslavia, France, Argentina, Romania, Taiwan, Italy, Switzerland, Canada, Brazil, Indonesia, the Netherlands, Russia, Kenya, Turkey, Venezuela, Portugal, India, Spain, and Iran (Michaelides-Mateou & Mateou, 2010). Since criminal prosecution occurs almost only under state or national statutes, an exhaustive global corpus of cases is difficult to track and build. Yet, according to Michaelides-Mateou & Mateou (2010), almost half of these criminal cases have been brought since 2000 (see figure 1 for a graphical condensation of their case descriptions), attesting to a strong criminalization trend in aviation. The trend not only affects pilots, mechanics and air traffic controllers, but increasingly accountable managers and nominated post holders. A number of cases are presented in more detail in the appendix to this paper.

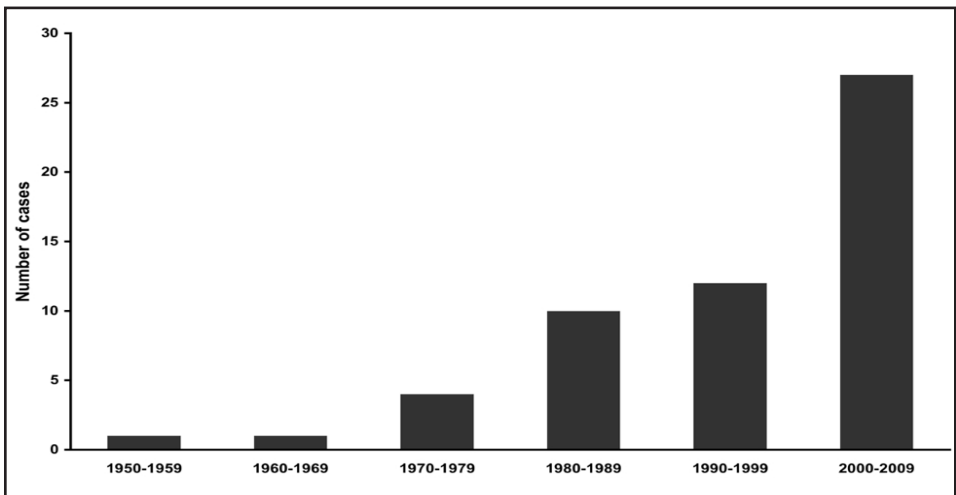


Figure 1: Number of worldwide cases of criminalizing human acts in aviation accidents and incidents per decade since 1950.

The criminalization trend in aviation mirrors developments (and concomitant concerns) in other fields, including shipping (Wallis, 2010), construction (ENR, 1997), chemical processing (Prakash, 1985), and health care (Grunsven, 1996; ISMP, 2007; Pandit, 2009; Skegg, 1998; Ukens, 2002), where “cases of doctors being subjected to criminal prosecution are on the increase” (Pandit, 2009, p. 379) and nurses’ errors are increasingly criminalized, also in the U.S. (Mee, 2007).

## Concerns

Concerns with the criminalization based on such cases and other statutes, can be divided up into those about the wisdom, fairness and utility of prosecution.

### Wisdom of prosecution

A focus of the industry has been on how judicial action in the aftermath of accidents and incidents interferes with independent safety investigations and destroys the willingness of people to voluntarily report errors and violations (Berlinger, 2005; Brous, 2008; Chapman, 2009; Dekker, 2007a, 2009; FSF, 2006; Thomas, 2007). Criminalization thus hampers the development of “safety cultures”: organizational cultures that encourage honest disclosure and open reflection on their practices with the aim to constantly improve quality and safety (Lauber, 1993).

A survey conducted by Michaelides-Mateou & Mateou (2010) confirms the preponderance of fear about prosecution and its detrimental effects on contributing to safety improvements. Says one of their respondents: “People cannot work with the fear of being prosecuted haunting them. Blaming and punishing someone will not help aviation safety. How can safety lessons be learnt if everyone is too scared to report an error or mishap?” (p. 282). Practicing under the threat of prosecution can only serve to hide errors, to condition people to get smarter at making evidence of possible criminalizable acts disappear, to discourage people from reporting their mistakes (Chapman, 2009; Michaelides-Mateou & Mateou, 2010).

Willingness to participate in independent safety investigations has also been found to go down when pilots or controllers have knowledge of previous criminal prosecutions: they do not want to incriminate themselves. Increasingly, pilots, mechanics and air traffic controllers refuse to participate in an independent safety investigation without the presence of a union representative or even a lawyer (Michaelides-Mateou & Mateou, 2010). This is testimony to the problem of co-mingling safety investigations and criminal probes, as it can stop people from cooperating with either of them (North, 2002).

Another effect is practicing more defensively, which may increase the unnecessary use of resources (Sharpe, 2004) or investments in paper trails that limit exposure and liability. Including managers in prosecution (particularly after pilots have died) may bring such adverse consequences. Organizational safety management can become an activity centered around reducing a company’s exposure and protecting management structures from criminal liability, which serves neither safety nor justice (Michaelides-Mateou & Mateou, 2010). Accountability demands that are seen as unreasonable and illegitimate (e.g. those imposed by the criminal justice system) can interfere with the conscientious execution of safety-critical work. There is experimental evidence suggesting that with unreasonable account-



ability demands, cognitive effort gets deflected into the management of liability risks—to the detriment of task-orientation (Lerner & Tetlock, 1999).

## Fairness of Prosecution

Over the last fifteen years, doubts have increasingly been voiced about the fairness of criminalizing errors that are made in the course of executing normal professional duties with no criminal intent in aviation and other fields (Mee, 2007; Merry & Peck, 1995; Moran, 2008; Reissner, 2009). There is also concern about the capriciousness of criminal prosecution: why some professionals, in some countries, get prosecuted for errors that have no such consequences elsewhere. Doubts also exist about the ability of a judiciary to make sense of the messy details of practice in a safety-critical domain (R. E. Anderson, 2005), let alone resist common biases of outcome knowledge and hindsight in adjudicating people's performance (J. C. Anderson, Jennings, Lowe, & Reckers, 1997; Arkes, Saville, Wortmann, & Harkness, 1981; Berlin, 2000; Dripps, 2003; Hawkins & Hastie, 1990; Hugh & Dekker, 2009; LaBine & LaBine, 1996; Laudan, 2006; Roese & Olson, 1996).

These doubts about a judiciary's ability to fairly adjudicate in the wake of professional mistake are amplified by a broad research consensus in safety research. Errors by pilots, controllers and mechanics, made in the normal pursuit of their duties, are heavily anchored and embedded in normal contexts in which they perform skilled work under conditions of resource constraints and outcome uncertainty (Woods, Dekker, Cook, Johannesen, & Sarter, 2010). This has raised significant skepticism about whether error can be punished or sanctioned away. Error is an inevitable part of the complex system in which it is generated (Amalberti, 2001, 2006; Clarke & Perrow, 1996; Leveson, 2002). Errors and other undesired outcomes are the inevitable product of the structural interactive complexity and tight coupling of the aviation system (Perrow, 1984). They occur not because unreliable people undermine otherwise smooth and well-functioning organizational processes. Rather, they emerge non-randomly as the side effects of well-organized processes (Pidgeon & O'Leary, 2000). Error in complex systems seems inevitable, no matter what sanction it might invite (Vaughan, 1996). Accidents that result in part from these inevitable errors are by definition unforeseeable and unintended. This makes it hard for accidents to meet the judicial principle of a *mens rea* (guilty mind) and thus puts them at odds with criminal prosecution (Michaelides-Mateou & Mateou, 2010).

For most professionals, a mistake that results in an incident, adverse event or inadvertent death is antithetical to their identities. It militates against their goals of delivering safe and efficient service (Berlinger, 2005; Sharpe, 2004; Wolf, 1994). Such errors, and their consequences, are experienced as a devastating failure to live up to the duty ethic inherent in the profession, and a betrayal of the trust that passengers or other users have put in it. The memory of mistake typically stays with professionals for many years (Serembus, Wolf, & Youngblood, 2001) and can cause excessive stress, depression, anxiety and other psychological ill-health (Berlinger, 2005; Lerner & Tetlock, 1999). Guilt and self-blame are also very common. Professionals can deny the role of the system or organization in spawning their mistake (Meurier, Vincent, & Parmar, 1998; Snook, 2000), despite the large research base to the contrary (Woods, et al., 2010).

When error gets criminalized, it can lead to sick leave, divorce, exit from the profession permanently or the committing of suicide (Chapman, 2009; Meszaros & Fischer-Danzinger, 2000; Moran, 2008; Tyler, 2003; Wolf, 1994). Another response to litigation, though rare, is anger and counter-attack, for example by filing a defamation lawsuit (R. E. Anderson, 2005; Sharpe, 2004). Criminalization can also have consequences for a person's livelihood (and his or her family), as licenses to practice may be revoked automatically which in turn can generate a whole new layer of anxiety and stress.

In the most constructive response, professionals try to process and learn from the error, discussing details of their error with their employer, contributing to its systematic investigation and helping with putting safety checks and improvements in place (Christensen, Levinson, & Dunn, 1992). The role of the organization in facilitating such coping (e.g. through peer and managerial support and appropriate structures and processes for learning from failure) is hugely important (Dekker & Laursen, 2007). It is crucial that employees do not get constructed as if they are the source of the problem and treated as somehow "troubled" as opposed to "normal" employees (Cooper & Payne, 1988; Dekker & Laursen, 2007). In aviation, and particularly in air traffic control, critical incident stress management (CISM) programs have been instituted in several countries. These voluntary peer programs have evolved from stress management interventions in particularly fire fighting and rescue services personnel and were first treated with suspicion by professionals because of the stigma of psychological infirmity its use might attract. It is now accepted and standard procedure in many organizations, however. Management has noticed that CISM helps professionals reenter productive operational life sooner after an incident, which benefits both organization and individual (Leonhardt & Vogt, 2006).

Neither CISM, nor people's progress through post-incident phases, has been investigated specifically for the influence of (criminal) prosecution. Prosecution probably affirms feelings of guilt and self-blame and exacerbates their effects, which are linked to poor outcomes in other criminological settings (Christensen, et al., 1992; Friel, White, & Alistair, 2008). At the same time, prosecution could destroy most opportunities for intervention by the employer or peers because it introduces new equations of mistrust, which can already be a problem after an adverse event (Scott, Hirschinger, & Cox, 2009). In addition, there could be organizational expediency and economy in not combating criminal prosecution of an employee as it publicly locates the source of the organization's safety problems in that single individual. Meaningful access could be cut off entirely when the professional is incarcerated (Learmount & Modola, 2004), and, not surprisingly, the prognosis for psychological health is never very good in that case (Friel, et al., 2008).

### **Utility of Prosecution**

There is no conclusive evidence about the extent to which the purposes of criminal justice (e.g. retribution, rehabilitation, prevention, and deterrence—specific or general) are served by the criminalization of professional mistake (Dekker, 2007c; Dekker & Hugh, 2009; A.F. Merry & McCall Smith, 2001). In fact, the prosecution of professionals can distort the allocation of scarce societal resources within the criminal justice system (Jacobs & Henry, 1996) when there are already bodies in place (e.g. accident investigation boards, medical discipline committees) that could be better positioned to deal effectively with the aftermath of failure in those

systems (FSF, 2006). In addition, broader, systemic interventions are known to have better safety effects than the prosecution of individuals. So how can a criminalization trend in aviation be explained?

Over the last 30 years the societal interpretation of accidents has shifted dramatically. Failures such as the Three Mile Island nuclear accident and the collision of two 747's at Tenerife in the seventies made society more "risk conscious" (I. Wilkinson, 2001). Accidents today are not seen as meaningless coincidences but as evidence that a particular risk was not managed well. And behind such mismanagement there are people, single persons, or single acts of omission or commission by those persons (Bittle & Snider, 2006; Green, 2003). Accidents are failures of risk management, which opens the door for the search (judicial or otherwise) for someone who did not manage risk well. The accident can go, or even needs to go, on somebody's account (Douglas, 1992).

The end of the twentieth century has also seen an increase in the democratization and accessibility of knowledge, as well as consumer vocalism and activism. These can put the failings of complex systems (or alleged failings of individuals in them) on fuller display (Anon., 2005; Pandit, 2009) and animate societal responses to them. The media doubtlessly enjoys a strong role in celebrating certain accidents, while being able to ignore others (Dekker, 2007b; Ditton & Duffy, 1983; Ödegård, 2007; Palmer, Emanuel, & Woods, 2001). A recent study links cultural and political populism to the punitiveness of a country's criminal justice system (Miyazawa, 2008). Media coverage of an event has been shown to articulate and animate social reactions to the point of constructing anti-heroes (Elkin, 1955; McLean & Elkind, 2004) and their crimes (Dekker, 2007b; Ericson, 1995; Innes, 2004; Jacobs & Henry, 1996; Tuchman, 1978). There is a strong basis to believe that the coverage of, and discourse surrounding social issues (e.g. accidents and human error), can be linked at least in part to political populism, judicial responses and the criminalization of new categories of human action (Blackwelder, 1996; Engbersen & Van der Leun, 2001; Husak, 2008; Jacobs & Henry, 1996; Phillips & Grattet, 2000).

A gradual reduction in the acceptance of risk altogether (Beck, 1992) has accompanied these developments, and there are now societal expectations that some safety-critical activities are entirely accident-free, with a zero-tolerance of failure. Aviation may have its own success to thank for this in part. Its increasingly flawless performance may have sponsored a societal belief in its infallibility and a concomitant political intolerance of failure (Amalberti, 2001). This means that almost of necessity, explanations of residual failure in these systems get deflected toward individual culprits (Perrow, 1984). The prosecution of individuals may thus hold some utility both for society and, in the Perrowian argument, for its intent on preserving a particular economic and social order (Goode, 1994). As Perrow (1984) pointed out about "human error":

...if this attribution can be made, that is the end of serious inquiry. Finding that faulty designs were responsible would entail enormous shutdown and retrofitting costs; finding that management was responsible would threaten those in charge, but finding that operators were responsible preserves the system... (p. 146)

A letter sent by the Boeing Corporation to the independent safety investigation of two inexplicable 737 crashes in the 1990s was seen by Byrne as an example of this. Investigators had found no evidence that the crew had done anything wrong, but the manufacturer expressed its dismay about the “desire of certain participants in our group to revisit, reexamine, and theorize about airplane system failures that could have contributed to the accident...” (Byrne, 2002, p. 162).

This is where the utility of prosecution for some groups becomes the unfairness of prosecution for others (Menkel-Meadow, 2000). Even victims of the results of the pilot or controller error sometimes see this, which puts them in sharp contrast to the focus of criminal prosecution on the single acts of single people. After an air traffic controller was jailed in the wake of a 1976 accident over Zagreb that killed 176 people, the father of one of the victims led a campaign to prevent the controller’s jailing. His campaign was unsuccessful, but the father joined efforts to free the controller after he had served two years (Geoffrey Thomas, 2002). Jailing individuals after system failure can be seen as unfair and counterproductive even by the primary victims; it can be seen as scapegoating (Mellema, 2000), which gets the organization or other people off the hook and oversimplifies the complexity of contributory events. Most importantly, prosecution of an individual may not give primary victims confidence that a similar incident will be prevented in the future.

### **Countermeasures**

The criminalization trend over the last fifteen years has exposed the difficulty of how and where the line between honest professional mistake and criminally liable act should be drawn, and by whom. This makes coordinated global action very difficult (Esler, 2009). Professional bodies have proposed to increase their defensive posture in response to the criminalization trend, for example by being more careful with external liaisons, particularly when it comes to sharing safety-related information (ICAO, 2007). In Canada, for instance, some airlines have asked their regulator to sign a non-disclosure agreement before safety inspections are conducted. One aim could be to protect the identity of employees who might, by disclosing information about incidents or violations, offer evidence of what can later be construed as criminal activity (Schmidt, 2009).

Various industries and countries have moved to different solutions. Most initiatives remain local and contingent on national law (under which most criminal prosecution occurs). Some initiatives locate the power to draw the line between acceptable and sanctionable performance more strongly inside of professions, for example by a re-asserted role of ethics or similar committees. At least one country has installed a so-called judge of instruction, who functions as a go-between before a prosecutor can go ahead with a case against a professional by checking the prosecutor’s homework and ambitions and weighing other stakeholders’ interests (which can work as long as those are fairly and equitably represented) (Dekker, 2009).

Other initiatives, most of them local or industry specific, are being developed and range from raising awareness and rallying opinion (FSF, 2006; GAIN, 2004; ICAO, 2007); to alternative dispute resolution and mediation and the legal protection of certain statements by professionals in the wake of failure (e.g. “I’m sorry” laws) (Berlinger, 2005; Sharpe, 2003); to stonewalling, by keeping the independent safety investigation open until the period of limitation for criminal prosecution

has expired (this may be many years); or by refusing to cooperate with any inquiry at all and destroying safety-related data before any access can be gained from the outside (Dekker, 2007c). Jointly, these effects create an adversarial stance that severely reduces openness, and could be counterproductive to longer-term societal efforts to achieve a balance between learning and accountability in safety-critical systems (Anon., 2009; Dekker, 2007c; FSF, 2006; ISMP, 2007; Michaels, 2008; Pandit, 2009; Ter Kulle, 2004; G. Thomas, 2007).

The data presented in this paper shows that the current protections offered by ICAO Annex 13 and similar treaties are insufficiently anchored in national laws, safety regulations and legal practices—a lack from which very few countries seem exempt. Transnational initiatives, for example in the European Union, are currently being undertaken that try to address this (TTE, 2010). In the end, countermeasures should focus on the implementation of strong national legislation that fairly balances accountability and learning. Norwegian and Danish examples of establishing a compulsory, non-punitive, and strictly confidential reporting system for aviation incidents could represent one example, of both the difficulty and modest possible success. In Denmark, immunity against use of such a report in prosecution is guaranteed within 72 hours of the incident. This provision made that air traffic control reporting rates tripled from one year to the next (Norbjerg, 2003). Not long after it was implemented, the law was tested in court, though, importantly, not in a case that involved a loss of life. A pilot who was brought to court in 2002 on the basis of an incident report submitted by himself saw the evidence from his own report thrown out because of the new law. Yet he was found guilty of negligence and perhaps left wondering whether not submitting a report might have been a better idea after all. And of course, even in these laws there are always provisions that exclude deliberate negligence—a category that remains hard to define and is always open to judgment (Dekker, 2007c).

## Conclusion

Criminalization of errors ultimately raises the question of who—in a society or an organization or a profession—gets the power to draw the line between acceptable and unacceptable behavior, to draw a moral boundary, and who gets to enforce it (Dekker, 2009). Just as Foucault (1982) described about France 150 years ago, different professions, branches of government and institutions might be vying for power and influence over the moral and legal privilege of calling something a criminal or otherwise sanctionable act. From this point of view, the line is not a location but a judgment, influenced by politics, power, or even sensationalism and populism (Dekker, 2009; Foucault, 1982; Morrill, Snyderman, & Dawson, 1997; Osborne, Blais, & Hayes, 1999). In the meantime, however, criminal prosecution of professionals such as pilots, air traffic controllers, or mechanics is increasingly seen as a threat to safety. Its effect on willingness to report and disclose safety-related information is well documented. What is encouraging is that the field of aviation has also germinated a number of cross-industry initiatives aimed at mitigating the effects of criminal prosecution (FSF, 2006; ICAO, 2007), something that is not likely to abate in the near future (Esler, 2009; Michaelides-Mateou & Mateou, 2010).

## References

- Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems. *Safety Science*, 37(2-3), 109-126.
- Amalberti, R. (2006). Optimum system safety and optimum system resilience: Agonistic or antagonistic concepts. In E. Hollnagel, D. D. Woods & N. Leveson (Eds.), *Resilience Engineering: Concepts and Precepts* (pp. 253-274). Aldershot: Ashgate Publishing Co.
- Anderson, J. C., Jennings, M. M., Lowe, D. J., & Reckers, P. M. (1997). The mitigation of hindsight bias in judges' evaluation of auditor decisions. *Auditing*, 16(2), 20-39.
- Anderson, R. E. (2005). *Medical malpractice: A physician's sourcebook*. New York: Humana Press.
- Anon. (2005). Murder! Mayhem! Social Order! *Wilson Quarterly*, 29, 94-96.
- Anon. (2009). Criminalization in aviation: Are prosecutorial investigations relegating aviation safety to the back seat? *Brief*, 38, 10-22.
- Arkes, H. R., Saville, P. D., Wortmann, R. L., & Harkness, A. R. (1981). Hindsight Bias Among Physicians Weighing the Likelihood of Diagnoses. *Journal of Applied Psychology*, 66(2), 252-255.
- Beck, U. (1992). *Risk society: Towards a new modernity*. London: Sage Publications Ltd.
- Berlin, L. (2000). Malpractice issues in radiology: Hindsight bias. *American journal of roentgenology*.
- Berlinger, N. (2005). *After harm: medical error and the ethics of forgiveness*. Baltimore: Johns Hopkins University Press.
- Bittle, S., & Snider, L. (2006). From manslaughter to preventable accident: Shaping corporate criminal liability. *Law & Policy*, 28(4).
- Blackwelder, S. P. (1996). Fearless wives and frightened shrews: The construction of the witch in Early Modern Germany. *Contemporary Sociology*, 25(4), 525-563.
- Brothers, C., & Maynard, M. (2008). More than 150 die in Madrid plane crash. *New York Times*.
- Brous, E. (2008). Criminalization of Unintentional Error: Implications for TAANA. *Journal of Nursing Law*, 12(1), 5-12.
- Byrne, G. (2002). *Flight 427: Anatomy of an air disaster*. New York: Springer-Verlag.
- Chapman, C. (2009, 18 April). A criminal mistake? *Chemist & Druggist*, 8.
- Christensen, J. F., Levinson, W., & Dunn, P. M. (1992). The heart of darkness: The impact of perceived mistakes on physicians. *Journal of General Internal Medicine*, 7, 424-431.
- Clarke, L., & Perrow, C. (1996). Prosaic organizational failure. *American Behavioral Scientist*, 39(8), 1040-1057.
- Cooper, C. L., & Payne, R. (1988). *Causes, coping, and consequences of stress at work*. Chichester ; New York: Wiley.



- Dekker, S. (2006). *The field guide to understanding human error*. Aldershot, England ; Burlington, VT: Ashgate.
- Dekker, S. W. A. (2007a). Criminalization of medical error: Who draws the line? *ANZ Journal of Surgery*, 77(10), 831-837.
- Dekker, S. W. A. (2007b). Discontinuity and disaster: gaps and the negotiation of culpability in medication delivery. *Journal of law, medicine and ethics*, 35(3), 463-470.
- Dekker, S. W. A. (2007c). *Just culture: Balancing safety and accountability*. Aldershot, England: Ashgate Publishing Co.
- Dekker, S. W. A. (2009). Just culture: Who draws the line? *Cognition, Technology & Work*, 11(3), 177-185.
- Dekker, S. W. A., & Hugh, T. B. (2009). Balancing "No Blame" with accountability in patient safety. *New England Journal of Medicine*, 362(3), 275.
- Dekker, S. W. A., & Laursen, T. (2007). From punitive action to confidential reporting: A longitudinal study of organizational learning. *Patient Safety & Quality Healthcare*, 5, 50-56.
- Ditton, J., & Duffy, J. (1983). Bias in the newspaper reporting of crime news. *British Journal of Criminology*, 23(2), 159-165.
- Douglas, M. (1992). *Risk and blame: Essays in cultural theory*. London: Routledge.
- Dripps, D. A. (2003). Fundamental retribution error: Criminal justice and the social psychology of blame. *Vanderbilt Law Review*, 56(5), 1381-1438.
- Elkin, F. (1955). Hero symbols and audience gratifications. *Journal of Educational Sociology*, 29(3), 97-107.
- Engbersen, G., & Van der Leun, J. (2001). The social construction of illegality and criminality. *European Journal on Criminal Policy and Research*, 9, 51-70.
- ENR. (1997). Human error cited in bridge collapse. *Engineering News-Record*, 238, 18.
- Ericson, R. (Ed.). (1995). *Crime and the media*. Dartmouth: Aldershot.
- Esler, D. (2009). Flight risk: The threat of criminalization. *Aviation Week & Space Technology*, 2010, from [http://www.aviationnow.com/aw/jsp\\_includes/articlePrint.jsp?storyID=news/bca0309p1.xml&headline=Flight%20Risk:%20The%20Threat%20of%20Criminalization](http://www.aviationnow.com/aw/jsp_includes/articlePrint.jsp?storyID=news/bca0309p1.xml&headline=Flight%20Risk:%20The%20Threat%20of%20Criminalization).
- Foucault, M. (1982). *I, Pierre Rivière, having slaughtered my mother, my sister, and my brother-- : a case of parricide in the 19th century*. Lincoln Neb.: University of Nebraska Press.
- Friel, A., White, T., & Alistair, H. (2008). Posttraumatic stress disorder and criminal responsibility. *Journal of forensic psychiatry and psychology*, 19(1), 64-85.
- FSF. (2006). Aviation safety groups issue joint resolution condemning criminalization of accident investigations. In F. S. Foundation (Ed.). Washington, DC: Flight Safety Foundation.
- GAIN. (2004). *Roadmap to a just culture: Enhancing the safety environment*. Global Aviation Information Network (Group E: Flight Ops/ATC Ops Safety Information Sharing Working Group).



- Garland, D. (2002). *The culture of control: Crime and social order in contemporary society*. Chicago, IL: University of Chicago Press.
- Goode, E. (1994). Round up the usual suspects: Crime, deviance, and the limits of constructionism. *American Sociologist*.
- Green, J. (2003). The ultimate challenge for risk technologies: Controlling the accidental. In J. Summerton & B. Berner (Eds.), *Constructing risk and safety in technological practice*. London: Routledge.
- Grunsven, P. R. (1996). Criminal prosecution of health care providers for clinical mistakes and fatal errors: Is bad medicine a crime? *Journal of Health and Hospital Law*, 29, 107.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, 107(3), 311-311.
- Hugh, T. B., & Dekker, S. W. (2009). Hindsight bias and outcome bias in the social construction of medical negligence: a review. *Journal of Law and Medicine*, 16(5), 846-857.
- Husak, D. (2008). *Overcriminalization: The limits of the criminal law*. New York: Oxford University Press.
- ICAO. (2007). Working paper of the 36th session of the Assembly: Protection of certain accident and incident records and of safety data collection and processing systems in order to improve aviation safety. In I. C. A. Organization (Ed.) (Vol. A36-WP/71 TE/16, 17/8/07). Montréal: International Civil Aviation Organization.
- Innes, M. (2004). Crime as a signal, Crime as a memory. *Journal for Crime, Conflict and the Media*, 1(2), 15-22.
- ISMP. (2007). Criminal prosecution of human error will likely have dangerous long-term consequences. Retrieved 4 February, 2010, from <http://www.ismp.org/Newsletters/acutecare/articles/20070308.asp>.
- Jacobs, J. B., & Henry, J. S. (1996). The social construction of a hate crime epidemic. *The journal of criminal law and criminology*, 86(2), 366-391.
- LaBine, S. J., & LaBine, G. (1996). Determinations of Negligence and the Hindsight Bias. *Law and Human Behavior*, 20(5), 501-516.
- Langewiesche, W. (1998). *Inside the sky : a meditation on flight* (1st ed.). New York: Pantheon Books.
- Lauber, J. K. (1993). *A safety culture perspective*. Paper presented at the Flight Safety Foundation's 38th Corporate Aviation Safety Seminar.
- Laudan, L. (2006). *Truth, error, and criminal law : an essay in legal epistemology*. Cambridge ; New York: Cambridge University Press.
- Learmount, D., & Modola, P. (2004, April 4). ATC body blasts Linate verdicts. *Flight International*.
- Leonhardt, J., & Vogt, J. (2006). *Critical incident stress management in aviation*. Aldershot, Hampshire, England ; Burlington, VT: Ashgate.
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255-275.

- Leveson, N. (2002). *System safety engineering: Back to the future*. Boston: MIT Aeronautics and Astronautics.
- Mail, C. (2009). Helios trial examines technical aspects. Retrieved 5 February, 2010, from <http://www.cyprus-mail.com/helios/helios-trial-examines-technical-aspects>.
- McLean, B., & Elkind, P. (2004). *The smartest guys in the room : the amazing rise and scandalous fall of Enron*. New York: Portfolio.
- Mee, C. L. (2007). Should human error be a crime? *Nursing*, 37, 6.
- Mellema, G. (2000). Scapegoats. *Criminal Justice Ethics*, 3-9.
- Menkel-Meadow, C. (2000). Can they do that? Legal ethics in popular culture: Of characters and acts. *UCLA Law Review*, 48, 1305-1340.
- Merry, A. F., & McCall Smith, A. (2001). *Errors, medicine and the law*. Cambridge: Cambridge University Press.
- Merry, A. F., & Peck, D. J. (1995). Anaesthetists, errors in drug administration and the law. *New Zealand Medical Journal*, 108, 185-187.
- Meszáros, K., & Fischer-Danzinger, D. (2000). Extended Suicide Attempt: Psychopathology, Personality and Risk Factors. *Psychopathology*, 33(1), 5-10.
- Meurier, C. E., Vincent, C. A., & Parmar, D. G. (1998). Nurses' responses to severity dependent errors: A study of the causal attributions made by nurses following an error. *Journal of Advanced Nursing*, 27, 349-354.
- Michaelides-Mateou, S., & Mateou, A. (2010). *Flying in the face of criminalization*. Farnham, UK: Ashgate Publishing Co.
- Michaels, D. (2008, September 9). When a tragedy becomes a crime: Prosecutors probe air disasters. *The Wall Street Journal*, p. 17.
- Miyazawa, S. (2008). The politics of increasing punitiveness and the rising populism in Japanese criminal justice policy. *Punishment & Society*, 10(1), 47-77.
- Moran, D. (2008). I was treated like a criminal after a harmless drug error. *Nursing Standard*, 22, 33.
- Morrill, C., Snyderman, E., & Dawson, E. J. (1997). It's not what you do, but who you are: Informal social control, social status, and normative seriousness in organizations. *Sociological Forum*, 12(4), 519-543.
- Norbjerg, P. M. (2003). *The creation of an aviation safety reporting culture in Danish air traffic control*. Kastrup, Denmark: Naviair.
- North, D. M. (2000). Let Judicial System Run Its Course In Crash Cases. *Aviation Week & Space Technology*, 152(20), 66-67.
- North, D. M. (2002). Oil and water, cats and dogs. *Aviation Week & Space Technology*, 156(5), 70-70.
- Ödegård, S. (Ed.). (2007). *I rättvisans namn (In the name of justice)*. Stockholm: Liber.
- Osborne, J., Blais, K., & Hayes, J. S. (1999). Nurses' perceptions: When is it a medication error? *JONA*, 29(4), 33-38.

- Palmer, L. I., Emanuel, L. L., & Woods, D. D. (2001). *Managing systems of accountability for patient safety*. Washington, DC: National Health Care Safety Council of the National Patient Safety Foundation.
- Pandit, M. S. (2009). Medical negligence: Criminal prosecution of medical professionals, importance of medical evidence: Some guidelines for medical practitioners. *Indian Journal of Urology*, 25(3), 379-383.
- Perrow, C. (1984). *Normal accidents: living with high-risk technologies*. New York: Basic Books.
- Phillips, S., & Grattet, R. (2000). Judicial Rhetoric, Meaning-Making, and the Institutionalization of Hate Crime. *Law & Society Review*, 34(3), 567-606.
- Pidgeon, N., & O'Leary, M. (2000). Man-made disasters: why technology and organizations (sometimes) fail. *Safety Science*, 34(1-3), 15-30.
- Prakash, S. S. (1985). The inhuman error: Lessons from Bhopal. *New Management*, 3(1), 40-45.
- Reissner, D. (2009). A criminal mistake? *Chemist & Druggist*, 271(6693), 8-9.
- Roese, N. J., & Olson, J. M. (1996). Counterfactuals, Causal Attributions, and the Hindsight Bias: A Conceptual Integration. *Journal of Experimental Social Psychology*, 32(3), 197-227.
- RTE. (2009). Tunisair pilots sentenced over Sicily crash. Retrieved 4 February, 2010, from <http://www.rte.ie/news/2009/0325/italy.html>.
- Ruitenbergh, B. (2002). Court case against Dutch controllers. *The Controller*, 41(4), 22-25.
- Sarter, N. B., & Woods, D. D. (1995). "How in the world did we get into that mode?" Mode error and awareness in supervisory control. *Human Factors*, 37(1), 5-19.
- Sarter, N. B., & Woods, D. D. (1997). Teamplay with a powerful and independent agent: A corpus of operational experiences and automation surprises on the Airbus A320. *Human Factors*, 39, 553-569.
- Schmidt, S. (2009, 25 February). Some airlines require confidentiality deals before safety inspections. *Canwest News Service*.
- Scott, S. D., Hirschinger, L. E., & Cox, K. R. (2009). The natural history of recovery for the healthcare provider "second victim" after adverse patient events. *Quality Safety Health Care*, 18, 325-330.
- Serembus, J. F., Wolf, Z. R., & Youngblood, N. (2001). Consequences of fatal medication errors for healthcare providers: A secondary analysis study. *Med-Surg Nursing*, 10(4), 193-201.
- Sharpe, V. A. (2003). Promoting patient safety: An ethical basis for policy deliberation. *Hastings Center Report*, 33(5), S2-19.
- Sharpe, V. A. (2004). *Accountability: patient safety and policy reform*. Washington, D.C.: Georgetown University Press.
- Skegg, P. D. G. (1998). Criminal prosecutions of negligent health professionals: the New Zealand experience. *Medical Law Review*, 6(2), 220-246.

- Snook, S. A. (2000). *Friendly fire: The accidental shutdown of US Black Hawks over Northern Iraq*. Princeton, NJ: Princeton University Press.
- Ter Kulle, A. (2004). Safety versus justice. *Canso News*, 18, 1-2.
- Thomas, G. (2002). Aviation on trial. *Air Transport World*, 9, 31-33.
- Thomas, G. (2007). A crime against safety. *Air Transport World*, 44, 57-59.
- Tingvall, C., & Lie, A. (2010). *The concept of responsibility in road traffic (Ansvarsbegreppet i vägtrafiken)*. Paper presented at the Transportforum.
- TTE. (2010). *EU-27 to address air traffic accident investigations*. Brussels: European Union Transport, Telecommunications and Energy Council (TTE). .
- Tuchman, G. (1978). *Making news: a study in the construction of reality*. New York: Free Press.
- Tyler, K. (2003). Helping employees cope with grief. *HRMagazine*, 48(9), 54-58.
- Ukens, C. (2002). Ohio R.Ph. may avoid criminal trial. *Drug Topics*, 146, 26.
- Vaughan, D. (1996). *The Challenger launch decision : risky technology, culture, and deviance at NASA*. Chicago: University of Chicago Press.
- Venet, M. (1984, October 29-November 1). *Open Letter*. Paper presented at the Human factors in managing aviation safety: Proceedings of the 37th annual international air safety seminar, Zurich, Switzerland.
- Wallis. (2010). Masters and pilots jailed for Neftegaz-67 casualty. Retrieved 20 January, 2010, from <http://www.lloydslist.com/art/1263565647614>.
- Wilkinson, I. (2001). *Anxiety in a risk society*. New York: Routledge.
- Wilkinson, S. (1994). The Oscar November incident. *Air & Space*, 80-87.
- Wolf, Z. R. (1994). *Medication errors: The nursing experience*. Albany, NY: Delmar.
- Woods, D. D., Dekker, S. W. A., Cook, R. I., Johannesen, L. J., & Sarter, N. B. (2010). *Behind human error*. Aldershot, UK: Ashgate Publishing Co.

## Appendix — selected cases

Fourteen passengers died in a Swissair DC-8 runway overrun accident at Athens on a flight from Geneva en route to Mumbai and Beijing in October 1979. The airplane had been carrying sixteen tons of extra fuel due to uplift constraints at Athens, and had been at maximum landing weight. The runway was extremely slippery due to rain and rubber deposits and its profile made the end of the runway hard to see. A trial was held in April 1983 and the captain and first officer were convicted of manslaughter, criminal negligence and interruption of air traffic. They were sentenced to five years imprisonment. Swissair offered to post the twenty million drachma bail (then \$266,000), to allow the crew to leave Greece. The captain refused, however, wanting to make this a test-case for the tenability of criminalizing pilot error (Venet, 1984).

In November 1989, a British Airways 747 carried out a missed approach to Heathrow in thick fog, narrowly missing a hotel near the other end of the runway. Two years later, the captain was found guilty in a split verdict of negligently endangering the aircraft and its passengers, the first time in British aviation history. The problem had begun much earlier, with a dinner in Mauritius that incapacitated both the copilot and the flight engineer through gastroenteritis during the flight. The airline had routinely been giving dispensations to copilots unqualified to fly low-visibility approaches, and did so in this case as well, a practice that had been condoned by the Civil Aviation Authority (CAA). Interestingly, and quite uniquely, the aviation prosecutor in Britain is employed by the CAA. Having been convicted and demoted, the captain eventually committed suicide (S. Wilkinson, 1994).

In January 1992, an Air Inter Airbus A320 crashed into a mountain near Strasbourg, France, while executing a night approach, killing all 87 persons aboard. Although the flight crew performed the approach correctly, a contributing factor in the accident was an uncommanded descent by the A320 of 3,200 feet per minute instead of the required 700, only two nautical miles from the airport. The accident became a prime example of “mode error” where crews are led to believe that they are making inputs and giving the automation instructions in one mode (in this case: Flight Path Angle) whereas the aircraft is actually in a different mode (Vertical Speed) (Sarter & Woods, 1995, 1997). Criminalization didn't occur until a full fourteen years after the accident when five current and former Airbus executives, including the A320 chief designer; two retired Air Inter executives; the former head of the country's Direction Générale de l'Aviation Civile (DGAC) and the aviation authority's retired certification director, as well as an air traffic controller, were prosecuted in French criminal court for negligent homicide. Even though they were acquitted, Airbus and Air France (which by then had taken over Air Inter) were found liable for pain and suffering of victims' families, and a subsequent trial was scheduled to determine monetary compensation (Esler, 2009).

In the wake of a June 1995 crash of an Ansett de Havilland Dash 8 near Palmston North in New Zealand, accident investigators turned the aircraft's cockpit voice recorder (CVR) over to criminal prosecutors. The crash killed four persons on the aircraft, but not the pilots, who faced possible charges of manslaughter. Pilots in New Zealand instituted proceedings to block the police use of the CVR, saying recorders should only be used for safety and educational purposes. Prosecutors prevailed and regained access to the CVR, but pilots soon began disabling CVRs on their flights. Officials have crafted a plan that would permit police use of

CVRs in future cases, provided New Zealand's High Court deemed it necessary (McKenna, 1999, pp. 47-48).

In May 1996, a ValuJet McDonnell Douglas DC-9 crashed into the Everglades not long after take-off from Miami. A carton with oxygen canisters had been placed in the forward cargo hold of the DC-9, and ignited shortly after takeoff. The fire and smoke incapacitated the crew, rendering the aircraft uncontrollable. All 110 people on board died. The official investigation determined that the oxygen generators were improperly packaged and labeled by ValuJet's contract maintenance provider, SabreTech, though correct packaging and labeling would have required them "to draw a verbal distinction between canisters that were 'expired,' meaning most of the ones they were removing, and canisters that were not 'expended,' meaning many of the same ones, loaded and ready to fire, on which they were expected to put nonexistent safety caps. Also involved were canisters that were expired and expended, and others that were not expired but were expended. And then, of course, there was the set of new replacement canisters, which were both unexpended and unexpired" (Langewiesche, 1998). Three SabreTech mechanics were indicted by a Florida court on criminal charges. The editor of *Aviation Week and Space Technology* "strongly believed the failure of SabreTech employees to put caps on oxygen generators constituted willful negligence that led to the killing of 110 passengers and crew. Prosecutors were right to bring charges. There has to be some fear that not doing one's job correctly could lead to prosecution" (North, 2000). In the ensuing trial, the mechanics were acquitted on the grounds that they "committed mistakes, but they did not commit crimes" (Esler, 2009). The jury did convict SabreTech and ordered it to pay a \$2.9 million fine, though an appeals court overturned this in 2005.

In December 1998, the crew of a Delta Airlines Boeing 767 had to abort its take-off from Amsterdam because of a Boeing 747 being towed across the runway in front of them. Low visibility procedures were in force at the time of the incident. Investigators found how ergonomic issues with a newly added panel, the surface movement radar displays as well as role ambiguities between coach and trainee controller all contributed to the confusion. Two years later, the coach, the trainee and the assistant controller involved were all charged under a section of the Netherlands Aviation Act, (§5.3), which provides that "it is prohibited to provide air traffic services in such a way that persons or property are endangered or could be endangered." Though a conviction was upheld in some sense, no punishment was imposed as the judge acknowledged that the prosecutor had used the incident as a "test case" (Ruitenberg, 2002).

In September 1999, a Dassault Falcon 900B, operated by Olympic Airways on behalf of the Greek government, was on approach to Bucharest, Romania, when one of the pilots tried to level the aircraft at 15,000 feet with the autopilot engaged. The autopilot disengaged, the artificial flight control feel system failed and a subsequent oscillation caused a violent upset in which passengers in the cabin not wearing seat belts were killed and one passenger and the flight attendant were injured (S. Dekker, 2006). A subsequent criminal trial was conducted in Greece, and the pilot who had been flying the aircraft was found guilty.

In July 2000, an Air France Concorde crashed in Paris after running over a titanium metal strip on the runway, causing tires on one of the main landing gear legs to explode and send fragments into a wing tank, igniting the fuel. The airliner



lifted off the runway but crashed into a nearby hotel seconds later, killing all 109 people on board as well as four on the ground (Esler, 2009). After the investigation, criminal charges were levied against a former regulatory official and two former executives of Concorde's manufacturer. In addition, Continental Airlines, whose DC-10 was suspected of dropping the titanium strip on the runway just before the Concorde took off, was placed under criminal investigation. In 2006, the French Supreme Court refused to dismiss the charges, and a trial was conducted in 2010.

In October 2000, a Singapore Airlines Boeing 747 crashed when taking off from Taipei, Taiwan in the dark and bad weather. The aircraft had run into construction equipment on runway 05R, which the crew had mistaken for runway 05L because of inadequate signage and taxiway lighting (burnt-out bulbs and inadequate spacing, among other problems). There was no surface movement radar at the airport. The crew was apprehended later the same evening and detained in Taiwan on suspicion of criminal negligence and manslaughter.

In October 2001, a Scandinavian MD-80 on its take-off run collided with a privately operated Cessna Citation business jet at Milan Linate airport in fog, killing 118 people. The airport surface environment radar system was not working, and taxiway markings had been poor for years. Five Italian officials, including the ex-manager of the Milan Linate Airport, the former director general of the ENAV Italian ATC agency, and a controller, were ultimately convicted of manslaughter and sentenced to between three and eight years in prison. In 2006, an appeals court reaffirmed the convictions (Learmount & Modola, 2004).

In July 2002, a DHL Boeing 757 collided with a Bashkirian Tupolev-154 over Uberlingen, Germany. The two aircraft collided at altitude over an intersection, with the loss of 71 lives. Four years later, negligent homicide charges were brought against eight Swiss air navigation services controllers and managers by Swiss prosecutors. The controller on duty at the time of the accident had been threatened with similar criminal charges, but was stabbed to death by the father of one of the crash victims.

In August, 2005, a Helios Airways Boeing 737 crashed in the mountains close to Athens airport, Greece, with 121 occupants. A failure of the cabin pressurization system led to the incapacitation and death of the crew and passengers. Several prosecutions were launched, including two for manslaughter charges, the first against five Helios officials in Cyprus and the second against six more in Greece (Esler, 2009). Trials were going on in both countries in 2010 (Mail, 2009).

In August 2005, a Tuninter ATR-72 was forced to ditch in the sea off the Sicilian coast after running out of fuel enroute to Djerba, Tunisia from Bari, Italy. The aircraft fuel gauges and indicators had been replaced, mistakenly, with those of the shorter, lighter (but otherwise identical) ATR-42 and had shown that there was enough fuel on board when the aircraft took off. 19 people died. The two pilots were convicted of multiple counts of manslaughter and air disaster, and sentenced to ten years in jail in 2009. In theory, they had the opportunity to reach the Palermo airport for an emergency landing. Five mechanics and managers were also found guilty, with the chief operating officer and maintenance chief sentenced to nine years each (RTE, 2009).



In September 2006, a Gol Linhas Aéreas Boeing 737-800 collided with an Embraer Legacy Business jet over the Amazon, killing 154 people on board the 737. The report, compiled by the Brazilian Air Force's Centro de Investigação e Prevenção de Acidentes Aeronáuticos, or CENIPA, held the Legacy flight crew (two Americans) and four Brazilian air traffic controllers liable for the deaths, resulting in a criminal trial against them (Esler, 2009). A separate investigation instead pointed to systemic and deep rooted problems in the country's military-run air traffic control system, which put the two aircraft on a collision course.

In March 2007, a Garuda Indonesia Boeing 737, overran the runway at Yogyakarta, Indonesia, and caught fire. Although 140 occupants escaped, 21 were killed (among them five Australians) and 12 were seriously injured. In its investigation report, the Indonesian National Transportation Safety Committee listed among probable causes the crew's failure to reject an unstabilized approach, the captain's failure to heed the first officer's repeated calls for a go-around as well as GPWS alerts, and the first officer's failure to take control of the airplane. A year after the accident, the captain was arrested and charged with criminal negligence, manslaughter, and violations of aviation regulations. The Indonesian court's indictment was based on the accident investigation findings, which were used as evidence against the captain (Esler, 2009). He was sentenced to two years in prison. One of the judges remarked that the sentence was about the prevention of future accidents rather than revenge.

In August 2008, a Spanair MD-82 crashed during take-off from Madrid-Barajas Airport, killing 154 people. The take off warning system did not warn the crew of a problem with the slats (high-lift devices on the leading edge of the wing). Mechanics who had worked on the aircraft just before the take off were facing manslaughter charges (Brothers & Maynard, 2008).



## ***Differential Effects of Likelihood Alarm Technology, Type of Automation, and Type of Task on Decision Making as Applied to Aviation and UAS Operations***

Ernesto A. Bustamante

and

Rylan M. Clark

*University of Idaho*  
*Moscow, ID 83844-3043*  
[ernestob\[at\]uidaho.edu](mailto:ernestob[at]uidaho.edu)

### *Abstract*

*The goal of this research was to examine the differential effects of likelihood alarm technology (LAT) on decision-making accuracy and bias as a function of the type of automation and task. The types of automation examined included miss-prone (MP) and false-alarm prone (FP) decision support tools (DSTs). The a-b signal detection theory model served as the framework for measuring decision-making accuracy (a) and bias (b). Results showed differential effects of LAT on a depending on the type of automation. (a) Participants were greater as a function of LAT when they interacted with the FP DST. Results showed differential effects of the type of automation depending on the type of task. (b) Participants were higher when they interacted with the MP DST while they performed the FP task. These findings had important theoretical implications for decision-making models and practical applications to the design of DSTs for aviation and UAS operations.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

## **Differential Effects of Likelihood Alarm Technology, Type of Automation, and Type of Task on Decision Making as Applied to Aviation and UAS Operations**

Adequate decision making is a critical component of aviation and unmanned aerial systems (UAS) operations. Given the complexity of aviation and UAS related tasks, human-automation interaction, especially with decision support tools (DSTs), has become an integral factor for ensuring the safety, effectiveness, and efficiency of pilots' and UAS operators' decision making. Human error accounts for 70-80% of aviation accidents (Wiegmann & Shappell, 2003), and more than half of UAS mishaps arise from inadequate human-systems integration (Tvarynas, Thompson, & Constable, 2005). Choosing what type of DST to incorporate within modern aviation technology has critical, and often underestimated, consequences for human-automation interaction. Therefore, determining DSTs characteristics is a vital issue that begins in the design stage, and it is important to be aware of the different influences of alarm technology, type of automation, and type of task configurations on pilots' and UAS operators' decision-making abilities. Consequently, the purpose of this research was to examine the differential effects of likelihood alarm technology (LAT) on decision-making accuracy and bias as a function of the types of automation and tasks.

DSTs help pilots monitor important automation displays and keep them 'in the loop' with the status of pertinent systems so that they can take necessary corrective actions. DSTs can also assist UAS operators in identifying targets and providing advisories for weapon deployment missions. There are numerous DSTs in the form of alarm systems designed to aid pilots' and UAS operators' important decision-making processes. When triggered, alarm systems function to create an attentional capture effect (Woods, 1995), which may cause pilots and UAS operators to focus on a particular subsystem or situation. DSTs in modern glass cockpits and UAS ground control stations serve such a vital role that designers must consider several factors: such as the type of alarm technology, automation, and task, which may play a role on the type of configuration to employ. Wilson (2005) asserted that DST technology reliably transfers between inhabited and uninhabited aircrafts. However, the nature of the DSTs' underlying alarm technologies, the types of automation, and the types of tasks associated with inhabited aircrafts may qualitatively and quantitatively differ from those associated with UAS, especially when it comes to human decision making. This study is part of a programmatic line of research seeking to establish empirical evidence for the costs and benefits, in terms of decision making, for different alarm technology configurations and types of automation as they apply to different types of aviation and UAS related tasks.

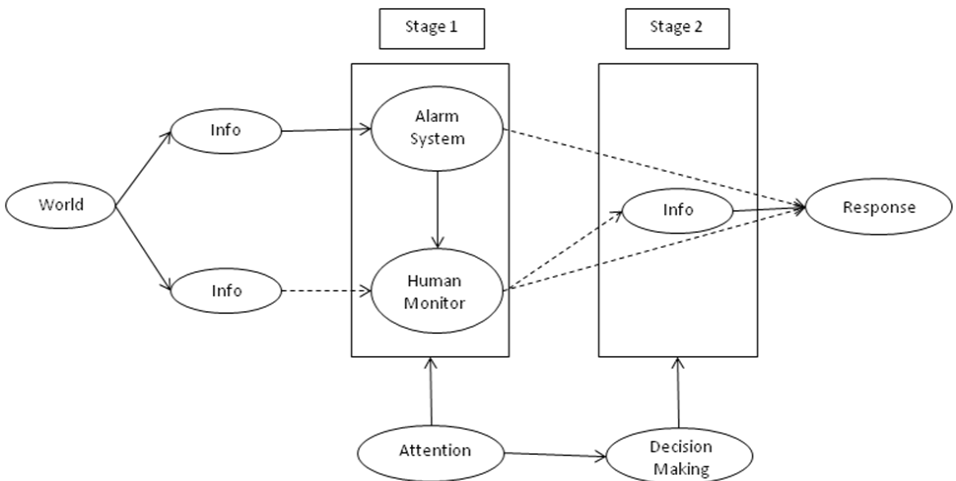
### **Decision Making with DSTs**

Decision making is a key component of pilots' and UAS operators' interaction with DSTs that manifests in how they react to advisories and make necessary corrective responses. Pilots' and UAS operators' decision making may be analyzed with an information-processing model composed of a sequence of stages: perception, attention, decision making, and response execution (Wickens, 1987). More specifically though, the two-stage signal detection theory (SDT) model developed by Bustamante (2008a) is particularly useful for examining pilots' and UAS operators' decision making while interacting with DSTs. The main reasons being are that the perceptual stage is not as critical because DST advisories, such as alarm

signals, are typically salient enough for pilots and UAS operators to perceive them (Edworthy & Stanton, 1995). Furthermore, response execution is generally a direct result of the decision making process (Wickens, 1987). However, attention and decision making are more relevant information processing stages, playing a focal role in detecting system problems and managing fault diagnosis (Moray, 1981).

### The Two-Stage Signal Detection Model of Decision Making

In his two-stage SDT model (see Figure 1), Bustamante (2008a) provided a framework for examining human decision making while interacting with DSTs. This model was derived from the theoretical foundations of Sorkin and Woods' (1985) signal detection analysis of systems with human monitors and Wickens' (1987) information processing model. Bustamante (2008a) also coupled his model with the theoretical foundations of the *a-b* SDT model of decision making (Bustamante, 2008b).



Note: Dashed lines indicate that the relationship may be optional depending on the type of setting.

Figure 1. Two-stage signal detection model of decision making.

The first decision-making stage, whereby pilots and UAS operators either acknowledge or ignore a perceived alarm signal, is mostly affected by attentional capacity. If pilots and UAS operators acknowledge an alarm, they are charged with searching for more information in an effort to diagnose the nature of the underlying problem and make a corrective decision. This second decision-making stage of the model is mostly shaped by information processing - deciding if the problem is real or not and implementing a chosen solution. Furthermore, either element of the human-automation system may bypass one or more of the information processing stages and make responses automatically based on the level of automation. The two-stage SDT model (Bustamante, 2008a), combined with measures of decision-making accuracy and bias derived from the *a-b* SDT model (Bustamante, 2008b), may be especially useful for evaluating the effectiveness of DSTs within the context of aviation and UAS operations.

## The *a-b* SDT Model of Decision Making

This model is based on the work of Bustamante (2008b), who offered alternative measures of decision-making accuracy (*a*) and bias (*b*) that do not rely on the underlying assumptions of traditional SDT. Instead, the *a* and *b* measures are based solely on an outcome matrix defined by the proportion of hits, false alarms, misses, and correct rejections. Within the *a-b* SDT model, decision-making accuracy is conceptually defined as the outcome due to making correct responses (i.e., hits and correct rejections) and mathematically estimated as their linear combination (see Formula 1). Decision-making bias, on the other hand, is conceptually defined as the outcome due to making affirmative responses (i.e., hits and false alarms) and mathematically estimated as their linear combination (see Formula 2).

$$a = .5 * p(HI) + .5 * p(CR) \quad (1)$$

$$b = .5 * p(HI) + .5 * p(FA) \quad (2)$$

The *a-b* SDT model of decision-making has several advantages over traditional SDT, which are relevant to aviation and UAS operations. First, given the lack of assumptions in comparison to traditional SDT, it is evident that the *a-b* SDT model is more parsimonious. One of the main reasons for this is that the *a-b* SDT model does not refer to an underlying decision-making continuum. Swets (1996) argued that the exact nature of the sensory excitation produced by either the noise or the stimulus could be quantified in terms of a single continuous variable. However, this argument does not apply well to domains where individuals and automated systems make decisions based on multiple sources of information and different decision-making algorithms, such as in aviation and UAS operations. Researchers have suggested that several factors may influence human decision making in a nonlinear fashion (Johnson, Bilimoria, Thomas, Lee, & Battiste, 2003). Examples include the use of cognitive heuristics (Kahneman, Slovic, & Tversky, 1982), experience (Bisseret, 1981), expertise (Klein, 1998), the amount of effort involved in choosing a particular alternative (Wogalter, Allison, & McKena, 1989), perceived urgency (Haas & Casali, 1995), risk perception (Ayres, Wood, Schmidt, & McCarthy, 1998), the emergency of the situation (Bliss & Gilson, 1998), workload (Bliss & Dunn, 2000), and trust (Lee & See, 2004). With regard to automated systems' decision-making, designers typically use highly complex algorithms that are also nonlinear, such as decision trees, Monte Carlo simulations, and neural networks (Canton, Refai, Johnson, & Battiste, 2005; Thomas, Wickens, & Rantanen, 2003; Yang, & Kuchar, 1997).

A second advantage of the *a-b* SDT model of decision making, which may be particularly relevant to aviation and UAS operations, is that DST designers and developers may interpret the *a* and *b* measures in an intuitively and readily applicable fashion. With regard to *a*, a score of 0 indicates the complete lack of ability to make accurate decisions. A score of .5 indicates decision-making performance at chance level, and a score of 1 indicates optimal decision-making accuracy. With regard to *b*, a score of 0 indicates a lack of affirmative responsiveness. A score

of .5 indicates an unbiased level of responsiveness, and a score of 1 indicates a complete response bias toward affirmative responses.

The most important advantage of the *a-b* SDT model of decision making, however, is that it provides metrics that more adequately measure the underlying detection and response processes, which lead to the estimation of distinct yet dependent decision-making outcomes (Bustamante, 2008b). Consequently, the *a-b* SDT model may serve as an adequate framework for examining the potential differential effects of likelihood alarm technology, type of automation, and type of task on decision making as applied to aviation and UAS operations.

## **Binary vs. Likelihood Alarm Technology**

As the name implies, binary alarm technology (BAT) emits two types of advisories intended to inform operators of the potential state of the world. Typically, DSTs equipped with BAT are either silent or activated, emitting one type of alarm signal when specified thresholds are exceeded. However, these systems may also provide operators with salient advisories regardless of the potential state of the world. These advisories may be comprised of a combination of variations of stimulus characteristics, such as color, signal word, and sound frequency. For example, an advisory intended to indicate a normal state of the world may consist of a green icon, with the signal word “OK” embedded in it, accompanied by a low-frequency sound. Likewise, an advisory intended to indicate an abnormal state of the world may consist of a red icon, with the signal word “ALARM” embedded in it, accompanied by a high-frequency sound.

One of the main issues associated with most alarm systems, especially in aviation and UAS operations, is that they tend to mostly emit false alarms (Bliss, 2003; Dixon & Wickens, 2006). This, in turn, tends to decrease human compliance (Stanton, Ragsdale, & Bustamante, 2009) and may lead to automation disuse (Parasuraman & Riley, 1997). The main problem, particularly related to BAT systems, is that they do not provide enough diagnostic information for operators to decide when it is critical to shift their focus of attention from primary to secondary tasks. More specifically, BAT systems tend to emit limited advisories for pilots and UAS operators to adequately divert their attention from the primary flight tasks to secondary subtasks that may require immediate corrective actions, such as monitoring the engine status of the aircraft they are flying or scanning aerial images to deploy weapons on enemy targets while flying UAS (Clark & Bustamante, 2008; Clark, Peyton, & Bustamante, 2009).

Likelihood Alarm Technology (LAT), on the other hand, is intended to provide operators with multiple advisories depending on the probability of abnormal states of the world. For example, in addition to the two levels of BAT, a DST equipped with LAT may emit an intermediate advisory intended to suggest a low probability of an abnormal state of the world, which may consist of a yellow icon with the signal word “WARNING” embedded in it accompanied by a medium-frequency sound. LAT is based on two fundamental human-automation interaction principles: probability matching and urgency mapping. As noted by Bliss, Gilson, and Deaton (1995), humans tend to match their frequency of responding to alarm signals based on the probability that they indicate an actual abnormal state of the world. Also, as emphasized by Edworthy and Loxley (1991), humans have a tendency to respond more often to alarm signals that they perceive to portray high urgency. The main



advantage of LAT over BAT, is that the former allows operators to better allocate their attentional resources and ultimately increase their decision-making accuracy and reduce their decision-making bias by responding more often to higher-likelihood advisories that suggest a higher urgency (Bustamante, 2005, 2007, 2008a; Bustamante & Bliss, 2005; Bustamante, Fallon, & Bliss, 2005; Clark & Bustamante, 2008; Clark, Ingebritsen, & Bustamante, in press; Clark et al., 2009).

## **False Alarm vs. Miss Prone Automation & Tasks**

Two critical aspects of DSTs are emitting alarm signals under normal states of the world (i.e., false alarms) and failing to emit alarm signals under abnormal states of the world (i.e., misses). Depending on which type of error it tends to exhibit, automation is typically conceived as 'false-alarm prone' (FP) or 'miss prone' (MP). FP systems tend to produce both high false-alarm and hit rates. MP systems, on the other hand, tend to generate high miss and correct-rejection rates. This tradeoff is influenced primarily by the accuracy of the DST and its pre-determined or adjustable alarm thresholds, which could be affected by the base rate of abnormal conditions and the costs associated with each type of error given the nature of the task (Bustamante, Bliss, & Anderson, 2007; Wickens & Dixon, 2007).

For instance, monitoring engine status in an inhabited aircraft could be conceived as a FP task. The main reason being is that the potential costs associated with failing to take corrective actions (e.g., engine failure, aircraft malfunction, and perhaps ultimately the loss of human lives) outweigh the potential costs associated with shifting attention from the primary flight tasks to diagnose normal engine functioning (e.g., unnecessary workload increase). On the other hand, conducting a weapon-deployment mission through the use of a UAS could be conceived as a MP task. The main reason being is that the potential costs associated with deploying a weapon on a non-military target (e.g., the destruction of civilian property and perhaps even lives) outweigh the potential costs associated with missing a military target (e.g., potential future military attacks from the missed target).

## **Goals of This Research**

The purpose of this research was to examine the differential effects of LAT on decision-making accuracy and bias as a function of the types of automation and tasks. The types of automation examined included FP and MP DSTs. The types of tasks included a simulated commercial aviation FP engine-monitoring task and a simulated UAS MP weapon-deployment task. The *a-b* SDT and the two-stage SDT models of decision making served to create the fundamental framework for measuring decision-making accuracy and bias.

Based on the previously discussed literature, we derive consistent hypotheses and expected results. We hypothesized a differential effect of LAT on decision-making accuracy depending on the type of automation. We expected LAT to enhance decision-making accuracy when implemented in the FP DST. We also hypothesized a differential effect of the type of automation on decision-making bias depending on the type of task. We expected higher decision-making bias when the type of automation mismatched the type of task.

## Method

### Research Design

We used a 2 x 2 x 2 between-groups fractional experimental design. We conducted two independent experiments. We systematically manipulated the type of alarm technology (BAT vs. LAT) and the type of automation (FP vs. MP) in the same manner for each experiment. The only difference between the two experiments was the type of task (FP vs. MP). As a method of experimental control, participants who completed Experiment A were not allowed to complete Experiment B and vice versa. Additionally, to avoid potential history and diffusion of treatment effects, we ran both studies concurrently yet physically isolated in the same laboratory. We used the *a-b* SDT model (Bustamante, 2008b) to calculate decision-making accuracy and bias during the first stage of the two-stage SDT model of decision making (Bustamante, 2008a). Prior to the beginning of data collection, we obtained proper approval from the Institutional Review Board. All researchers involved in data collection completed the National Institute of Health online training course for protecting human participants.

### Participants

An *a priori* power analysis revealed that approximately 25 participants would be necessary to obtain statistically significant effects at a .01 alpha level, assuming a power of .80 and a medium effect size for each factor (Cohen, 1988). Consequently, one-hundred university students participated in each experiment. The demographic characteristics of participants in each experiment were very similar. Fifty-three women and 47 men participated in Experiment A. They ranged from 18 to 42 years of age ( $M = 20.07$ ,  $SD = 3.06$ ). Forty-nine women and 51 men participated in Experiment B. They ranged from 18 to 39 years of age ( $M = 20.51$ ,  $SD = 3.68$ ). In both experiments, we randomly assigned each participant to one of four experimental conditions, characterized by the type of alarm technology and the type of automation. There were a total of 20 participants in each experimental condition in both experiments. All participants reported to have normal or corrected-to-normal vision and hearing. As incentives, we compensated participants with two research credits and provided a \$25 gift certificate to the participant with the highest level of performance. Throughout the beginning and completion of either experiment, researchers treated participants according to the American Psychological Association Ethical Guidelines.

### Materials and Apparatus

Two computer workstations equipped with 19" monitors, standard QWERTY keyboards, optical mice, and sound-attenuating headphones were used to allow participants to complete the primary and secondary tasks in each experiment. The presentation of the primary and secondary tasks was consistent with prior research (Bustamante, 2008a). The primary tasks were presented in participants' main field of view, and the secondary task was presented at a 90° angle to the left of participants' main field of view. The rationale for this configuration was to generate structural interference between the primary and secondary task displays, which is common in aviation and UAS operations.

Two tasks of the Multi-Attribute Task Battery (Comstock & Arnegard, 1992) were used to simulate the primary flight tasks (see Figure 2).

**Compensatory-tracking task.** The main purpose of this task was to simulate the key function that pilots and UAS operators need to perform to fly an airplane or a UAS, which is to maintain level flight.

**Resource-management task.** The main purpose of this task was to simulate another important function that pilots and UAS operators need to perform as they fly an airplane or a UAS, which is to make sure that they have an adequate level of fuel.

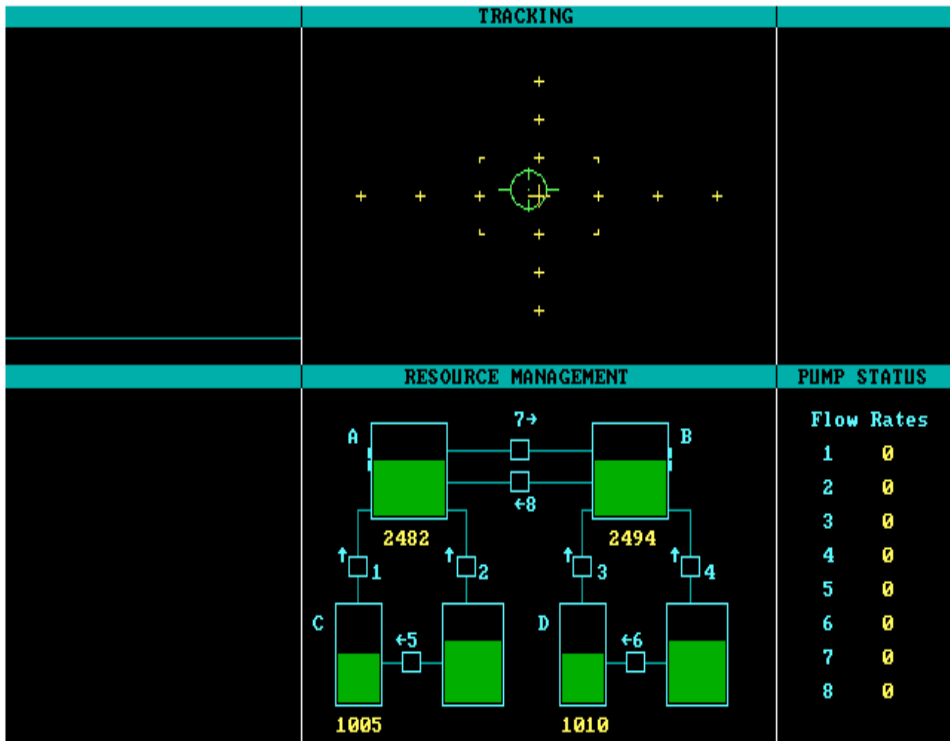


Figure 2. Multi-attribute task battery.

In addition to performing the primary flight tasks, participants had to perform a secondary task. In Experiment A, participants performed an engine-monitoring task (Stanton et al., 2009). The main purpose of this task was to simulate a crucial secondary function that pilots need to perform to maintain flight safety, which is to ensure that they have at least one fully functioning engine at all times. Participants performed this task with the aid of a simulated EICAS display (see Figure 3).



Figure 3. Simulated EICAS display.

The simulated EICAS display equipped with BAT provided one of two types of advisories (OK, ALARM) regarding the potential status of the engines. The simulated EICAS display equipped with LAT provided one of three types of advisories (OK, WARNING, ALARM). After receiving the EICAS advisory, participants could ignore it and continue performing their primary flight tasks, or they could acknowledge it and search for system-status information. If so, they had to evaluate the temperature and pressure levels of two engines and determine whether they needed to make a corrective response to repair the engines if both of them were malfunctioning (see Figure 4).

# System Status

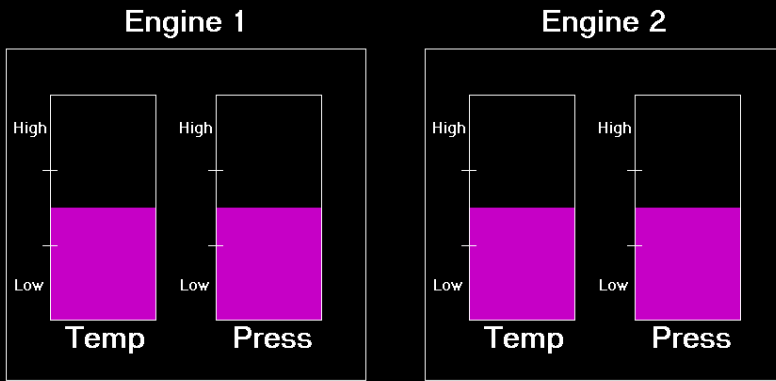
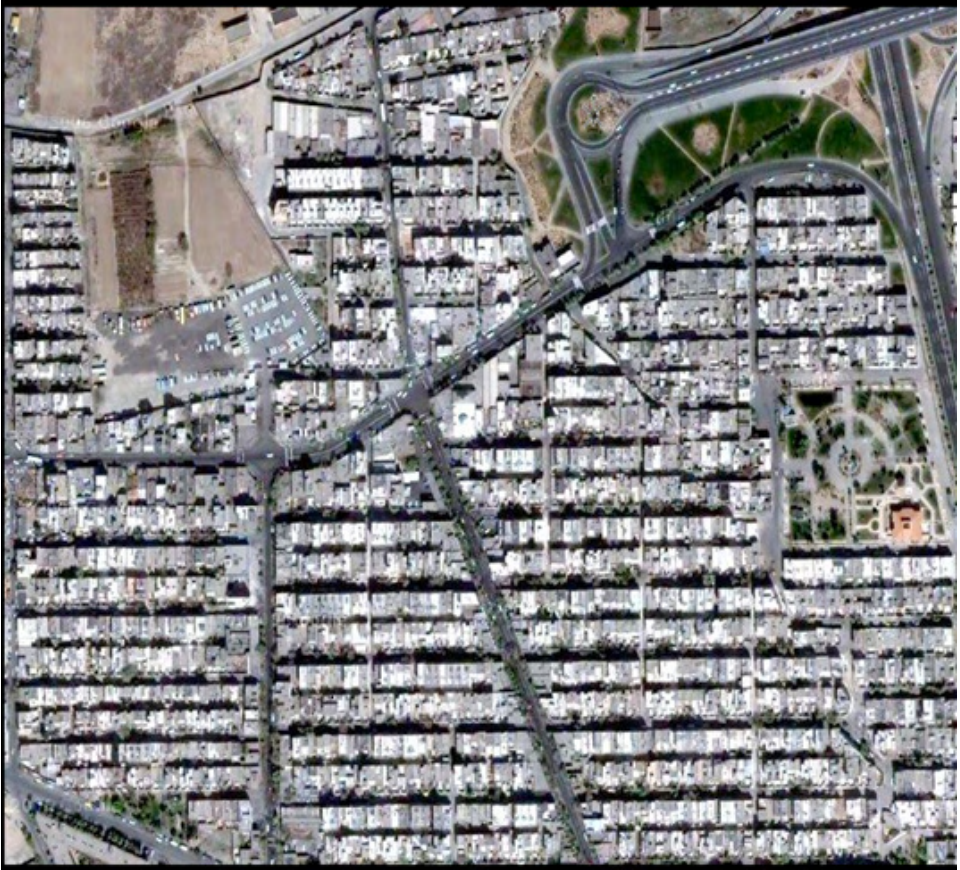


Figure 4. System-status information.

In Experiment B, participants performed a secondary weapon-deployment task (Clark et al., 2009). The main purpose of this task was to simulate one of the functions that UAS operators may need to perform in a combat situation, which is to deploy a weapon on a military target. To accomplish this task, participants interacting with the BAT DST received one of two types of advisories (OK, ALARM), and participants interacting with the LAT DST received one of three types of advisories (OK, WARNING, ALARM) regarding the potential presence of an enemy target. After receiving the advisory, participants could ignore it and continue performing their primary flight tasks, or they could acknowledge it and view an aerial image to find the enemy target. If participants considered that an enemy target was present, they had to deploy a weapon to destroy it (see Figure 5).





*Figure 5. Sample aerial image.*

### **Type of Automation and Alarm Technology**

The type of automation (FP vs. MP) was manipulated in a similar manner as Dixon, Wickens, and McCarley (2007). The FP system had a 100% hit rate and an 80% false alarm rate. The MP system, on the other hand, had a 20% hit rate and a 0% false alarm rate. As previously mentioned, the BAT system provided participants with one of two types of advisories. One of such advisories was composed of a green rectangle with the signal word “OK” embedded in it, accompanied by a 500-Hz simple sign wave sound. The other advisory was composed of a red rectangle with the signal word “ALARM” embedded in it, accompanied by a 3000-Hz simple sign wave sound. The LAT system provided participants with three types of advisories depending on the probability of engine malfunctions in Experiment A and the presence of the enemy target in Experiment B. Two of the three advisories had the same physical characteristics as the stimuli used for the BAT system. The LAT system, however, provided an additional intermediate advisory composed of a yellow rectangle with the signal word “WARNING” embedded in it, accompanied by a 1500-Hz simple sign wave. All advisories were presented at a constant amplitude level of 65 dB(A) through the set of sound-attenuating headphones.

## Procedure

Participants came to the laboratory individually and at different times. When they came into the laboratory, they first read and signed an informed consent form and completed a demographic questionnaire. Then, researchers explained to participants the nature of the experiment and provided them with the instructions on how to perform the tasks. Next, participants performed each of the tasks individually and concurrently to familiarize themselves with the nature of each task. As part of each experiment, participants completed three 30-min sessions, separated by 5-min breaks. During the first session, participants completed the required tasks without the aid of any DST. Throughout the second and third sessions, participants completed the required tasks with the aid of the respective DST. The main reason for having participants complete three sessions was to allow their performance to plateau at a reliable level. Given this, we examined participants' decision-making accuracy and bias during the third session only. Completion of either experiment lasted approximately two hours. After participants completed either experiment, researchers debriefed and thanked them for their participation.

## Results

### Decision-Making Accuracy

Results from a 2 x 2 x 2 between-groups ANOVA showed a statistically significant with a medium effect size ordinal interaction effect between the type of alarm technology and automation on decision-making accuracy,  $F(1, 192) = 34.85, p < .01$ , partial  $\eta^2 = .15$ , observed power = 1.00. LAT significantly increased participants' decision-making accuracy when they interacted with the FP system only. These results are graphically depicted in Figure 6. Table 1 contains the estimated marginal means and standard errors for the four conditions characterized by the type of alarm technology and the type of automation.

Table 1

*Estimated Marginal Means and Standard Errors of Participants' Decision-Making Accuracy*

Alarm Technology	Type of Automation			
	FP		MP	
	M	SE	M	SE
BAT	.55	.01	.51	.01
LAT	.71	.01	.53	.01



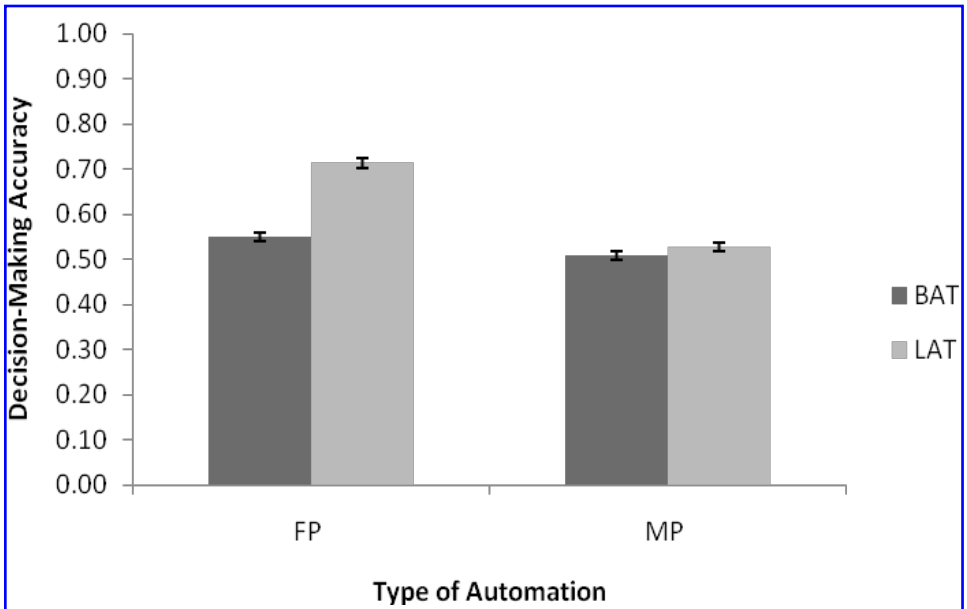


Figure 6. Decision-making accuracy as a function of type of automation and alarm technology.

Additionally, results showed a statistically significant main effect with a medium effect size of the type of technology on decision-making accuracy,  $F(1, 192) = 56.01, p < .01$ , partial  $\eta^2 = .23$ , observed power = 1.00. In general, decision-making accuracy was significantly greater for participants who interacted with LAT ( $M = .62, SE = .01$ ) than for those who interacted with BAT ( $M = .53, SE = .01$ ). Last, results showed a statistically significant main effect with a large effect size of the type of automation on decision-making accuracy,  $F(1, 192) = 83.81, p < .01$ , partial  $\eta^2 = .30$ , observed power = 1.00. In general, decision-making accuracy was significantly greater for participants who interacted with FP automation ( $M = .63, SE = .01$ ) than for those who interacted with MP automation ( $M = .52, SE = .01$ ).

### Decision-Making Bias

Results from a 2 x 2 x 2 between-groups ANOVA showed a statistically significant disordinal interaction effect between the type of automation and task on decision-making bias,  $F(1, 192) = 6.98, p < .01$ , partial  $\eta^2 = .04$ , observed power = .75. Differences in decision-making bias varied depending on the mismatch between the type of automation and task. These results are graphically depicted in Figure 7. Table 2 contains the estimated marginal means and standard errors for the four conditions characterized by the type of automation and task.

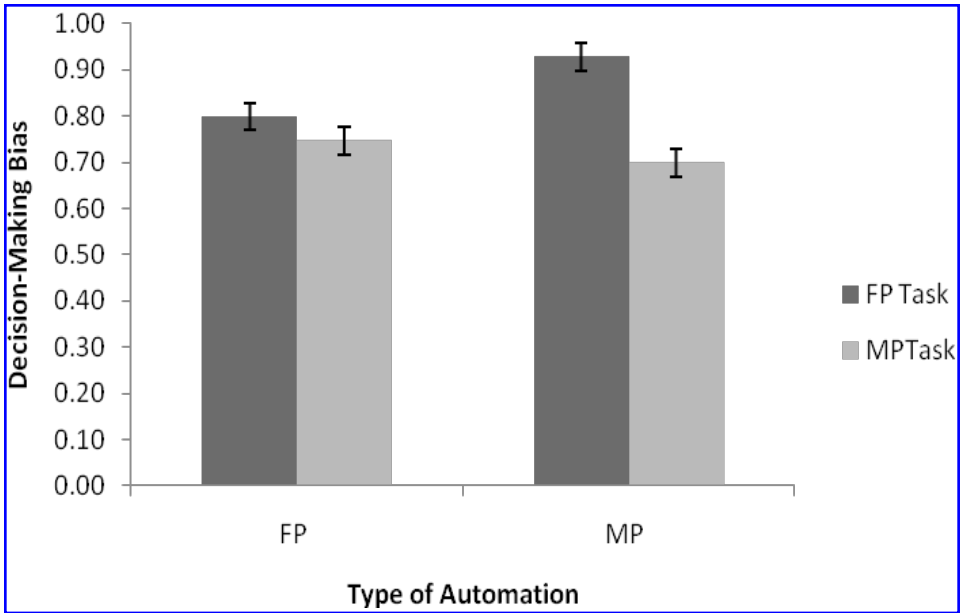


Figure 7. Decision-making bias as a function of type of automation and task.

Table 2

Estimated Marginal Means and Standard Errors of Participants' Decision-Making Bias

Type of Task	Type of Automation			
	FP		MP	
	M	SE	M	SE
FP	.80	.03	.93	.03
MP	.75	.03	.70	.03

Additionally, results showed a statistically significant main effect of the type of task on decision-making bias,  $F(1, 192) = 17.65, p < .01$ , partial  $\eta^2 = .08$ , observed power = .99. In general, decision-making bias was significantly greater for participants who performed the FP task ( $M = .86, SE = .02$ ) than for those who performed the MP task ( $M = .72, SE = .02$ ).

## Discussion

### Differential Effects of LAT and Type of Automation on Decision-Making Accuracy

The findings from this research effort provided empirical support for our main contentions and derived hypotheses regarding the differential effects of LAT and type of automation on decision-making accuracy. As expected, LAT enhanced

decision-making accuracy, especially when applied to FP automation. This trend was consistent with prior empirical research (Clark et al., 2009) and theoretical postulations for the use of LAT (Bustamante, 2008a). As previously mentioned, FP automation tends to emit a large proportion of false alarms. What previous literature, as well as the present findings, suggest is that LAT is a useful approach to mitigate the potential negative ramifications due to false alarms. The main goal of LAT is to provide individuals with more diagnostic information regarding the probability that different alarm signals actually indicate the presence of abnormal situations. Given this, it follows that LAT should be applied to FP systems because, unlike MP automation, such systems have a greater tendency to generate a high proportion of false alarms.

## **Differential Effects of the Type of Automation and Task on Decision-Making Bias**

The findings from this research effort also provided empirical support for our main contentions and derived hypotheses regarding the differential effects of the type of automation and task on decision-making bias. As predicted, decision-making bias differentially increased due to the mismatch between the type of automation and task. Results from this research showed that response bias toward acknowledging alarms was highest when individuals performed the FP task while interacting with MP automation. These findings may seem counterintuitive at first, perhaps due to the nature of the terminology used to characterize the types of automation and tasks. However, this effect was consistent with the fundamental notions of what MP tasks and FP automation entail. Considering that the costs associated with responding to false alarms in MP tasks outweigh the costs of responding to false alarms in FP tasks, it follows that individuals would be more biased toward acknowledging alarms in general while performing a FP task. Furthermore, it is also important to take into account the fundamental difference between FP vs. MP automation. Considering that MP systems rarely generate false alarms, it follows that individuals would be more biased toward acknowledging alarms in general while interacting with MP automation. Integrating these two theoretical premises, it was reasonable to expect individuals to be more biased toward acknowledging alarms emitted by a MP system while performing a FP task.

## **Theoretical Implications**

The findings from this research effort had important theoretical implications for the *a-b* SDT and the two-stage models of decision making. As previously argued, the main theoretical contention of the *a-b* SDT model of decision making is that it provides metrics that more adequately measure the underlying detection and response processes, which lead to the estimation of distinct yet dependent decision-making outcomes (Bustamante, 2008b). Also, as previously discussed, the two-stage SDT model of decision making (Bustamante, 2008a) emphasizes the attentional and decision-making stages of information processing as the critical factors in human interaction with DSTs. These two stages are theoretically mapped to the fundamental detection and response processes underlined in the *a-b* SDT model of decision making.

The differential effects of LAT, the type of automation, and the type of task on decision-making accuracy and bias provided empirical support for the integration of these two models. The factors that were intended to affect decision-making ac-

curacy, such as LAT, interacted with factors that were not intended to have this effect, such as the type of automation. Furthermore, the factors that were intended to have independent effects on decision-making bias, such as the type of automation and task, had an interaction effect instead. These interactions provided empirical support for the contention that decision-making accuracy and bias are distinct yet dependent outcomes of human interaction with DSTs.

The results from this study also had important theoretical implications for the principles of probability matching (Bliss et al., 1995) and urgency mapping (Edworthy, & Loxley, 1991), which served to create the fundamental framework for implementing LAT in the design of DSTs. The differential effects of LAT and the type of automation provided empirical support for taking advantage of probability matching and urgency mapping. As previously mentioned, LAT increased decision-making accuracy when it was coupled with FP automation. The potential reason for this might be that following the probability matching and urgency mapping principles, participants responded more often to high-likelihood alarms, which portrayed higher urgency, and less often to low-likelihood alarms, which portrayed lower urgency.

## Practical Applications

The findings from this research effort also had important practical applications for aviation and UAS operations. As previously mentioned, the main advantage of LAT depends upon its more diagnostic signals, which could help pilots and UAS operators make better decisions regarding their attention allocation. While performing predominantly monitoring tasks that require corrective actions, such as engine monitoring and weapon deployment, pilots and UAS operators could benefit from interacting with DSTs that appropriately capture their attention depending on the urgency of such tasks over the main primary flight tasks. Moreover, complex and high-risk aviation and UAS operations, which require accurate decision making regarding the allocation of attentional resources to different tasks, may become more manageable when assessing the effects of alarm configurations using these findings.

Currently, aviation and UAS designers primarily adhere to the *proximity compatibility principle* (Wickens & Carswell, 1995) grouping information relevant to a common task or mental operation together within an integrated DST display. Such displays are prevalent in aviation (Bliss, 2003) and UAS operations (Dixon & Wickens, 2006). In general, display integration makes more intuitive sense than relying on pilots' and UAS operators' ability to make decisions by integrating related information from different and scattered displays. A consequence of this high degree of display integration, however, is that not all relevant information is readily available to pilots and UAS operators as they must navigate through various display pages taking up a great deal of their attentional resources (Bustamante, 2008a). The findings from this research suggest that LAT may aid pilots and UAS operators stay 'in the loop' with regard to the states of world as they relate to the tasks at hand by allowing them to allocate their attentional resources more effectively.

Furthermore, consistent with extensive prior research (Bustamante, 2005, 2007, 2008a; Bustamante & Bliss, 2005; Bustamante, Fallon, & Bliss, 2005; Clark & Bustamante, 2008; Clark, Ingebritsen, & Bustamante, in press; Clark et al., 2009), the findings from this study provide compelling evidence for the potential

benefits of the application of LAT to the design and implementation of FP DSTs within integrated aviation and UAS displays. More specifically, FP DSTs equipped with LAT could provide pilots and UAS operators with the ability to respond faster and more accurately to potential threats while performing several tasks (Ratches, J. A., Walters, C. P., Buser, R. G., & Guenther, B.D., 1997).

### **Potential Limitations & Suggestions for Future Research**

The main overarching potential limitation of this study was its level of ecological validity, which is a common limitation of most laboratory research. Notwithstanding, it is important to keep in mind the delicate balance between low-fidelity laboratory research and high-fidelity field research. The former allows researchers to achieve high levels of internal validity, whereas the latter allows practitioners to achieve high levels of external validity. Although laboratory research may lack the high level of ecological validity that is more conducive for applying its findings to specific domains, it nevertheless may allow researchers to generalize their findings to applied settings given its higher level of internal validity.

More specifically though, the present research had specific potential limitations, which may intuitively lead to suggestions for future research. First, the way in which participants performed the tasks involved in the present study may be qualitatively different from the way in which certified pilots and UAS operators may perform their respective tasks in field settings. Second, participants in this research individually performed simulated aviation or UAS operations that are typically performed by more than one individual in field settings. Third, the nature of the simulated environments used in this study may not be readily applicable to actual aviation and UAS operations due to differences in their level of complexity. Last, an important component of human interaction with DSTs in critical situations is the imminent level of risk associated with making inaccurate and biased decisions. Due to ethical and practical constraints, it was unfeasible to simulate a level of risk qualitatively comparable to real-world aviation and UAS operations. Consequently, future research efforts should focus on validating the findings of this study using certified pilots and UAS operators, evaluating the performance of pilots and UAS operators at individual and team levels of analysis, and increasing the level of complexity and imminent risk by examining pilots' and UAS operators' performance in actual aviation and UAS operations.

### **Conclusion**

Decision-making accuracy and bias are critical components of human performance in aviation and UAS operations. The goal of this research was to examine the differential effects of LAT on decision-making accuracy and bias as a function of the type of automation and task. As expected, results showed a differential effect of LAT on decision-making accuracy depending on the type of automation. Also, as predicted, results showed a differential effect of the type of automation depending on the type of task. The findings from this research effort had important theoretical implications for the *a-b* SDT and the two-stage models of decision making and practical applications to the design of DSTs for aviation and UAS operations.

## Acknowledgment

We would like to acknowledge Dr. Stephen Rice for providing us with the aerial images used in Experiment B.

## References

- Ayres, T. J., Wood, C. T., Schmidt, R. A., & McCarthy R. L. (1998). Risk perception and behavioral choice. *International Journal of Cognitive Ergonomics*, 2(1-2), 35-52.
- Bisseret, A. (1981). Application of signal detection theory to decision making in supervisory control: The effect of the operator's experience. *Ergonomics*, 24(2), 81-94.
- Bliss, J. P. (2003). Investigation of alarm-related accidents and incidents in aviation. *The International Journal of Aviation Psychology*, 13(3), 249-268.
- Bliss, J. P., & Dunn, M. C. (2000). Behavioural implications of alarm mistrust as a function of task workload. *Ergonomics*, 43(9), 1283-1300.
- Bliss, J. P. & Gilson, R. D. (1998). Emergency signal failure: Implications and recommendations, *Ergonomics*, 41, 57-72.
- Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behaviour in response to alarms of varying reliability. *Ergonomics*, 38(11), 2300-2312.
- Bustamante, E. A. (2005). A signal detection analysis of the effects of workload, task-critical and likelihood information on human alarm response. In *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1513-1517). Santa Monica, CA: Human Factors and Ergonomics Society.
- Bustamante, E. A. (2007). Using likelihood alert technology in cockpit displays of traffic information to support free flight. In *Proceedings of the 14th International Symposium on Aviation Psychology* (pp. 100-102). Dayton, OH.
- Bustamante, E. A. (2008a). Implementing likelihood alarm technology in integrated aviation displays for enhancing decision making: A two-stage signal detection modeling approach. *International Journal of Applied Aviation Studies*, 8(2), 241-262.
- Bustamante, E. A. (2008b). The *a-b* signal detection theory model. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 68(9-B), 6371.
- Bustamante, E. A., & Bliss, J. P. (2005). Effects of workload and likelihood information on human response to alarm signals. In *Proceedings of the 13th International Symposium on Aviation Psychology* (pp. 81-85). Oklahoma City, OK: Wright State University.
- Bustamante, E. A., Bliss, J. P., & Anderson, B. L. (2007). Effects of varying the threshold of alarm systems and workload on human performance. *Ergonomics* 50(7), 1127-1147.

- Bustamante, E. A., Fallon, C. K., & Bliss, J. P. (2005). Improving decision-making with a likelihood alarm display [Abstract]. *American Psychological Association Divisions 19 & 21 Midyear Symposium*. Fairfax, VA, March 3-4.
- Canton, R., Refai, M., Johnson, W. W., & Battiste, V. (2005). Development and integration of human-centered conflict detection and resolution tools for airborne autonomous operations. In *Proceedings of the 13th International Symposium on Aviation Psychology* (pp. 92-97). Oklahoma City, OK: Wright State University.
- Clark, R. M., & Bustamante, E. A. (2008). Enhancing decision making by implementing likelihood alarm technology in integrated displays. *Modern Psychological Studies*, 14(1), 36-49.
- Clark, R. M., Ingebritsen, A., & Bustamante, E. A. (in press). Differential effects of likelihood alarm technology and false-alarm vs. miss-prone automation on decision-making accuracy and bias. *Proceedings of the Human Factors and Ergonomics Society 54<sup>th</sup> Annual Meeting*. San Francisco, CA: Human Factors and Ergonomics Society.
- Clark, R. M., Peyton, G. G., & Bustamante, E. A. (2009). Differential effects of likelihood alarm technology and false-alarm vs. miss prone automation on decision making. In *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* (pp.349-353). Santa Monica, CA: Human Factors and Ergonomics Society.
- Dixon, S. R. & Wickens, C. D. (2006). Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload. *Human Factors*, 46, 474 – 486.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Comstock, J. R., & Arnegard, R. J. (1992). *The multi-attribute task battery for human operator workload and strategic behavior research* (NASA Technical Memorandum No. 104174). Hampton, VA: National Aeronautics and Space Administration, Langley Research Center.
- Dixon, S. R. & Wickens, C. D. (2006). Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload. *Human Factors*, 46, 474 – 486.
- Dixon, S.R., Wickens, C.D., & McCarley, J.S. (2007). On the independence of compliance and reliance: Are automation false alarms worse than misses? *Human Factors*. 49(4), 564-572.
- Edworthy, J., & Loxley, S. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors*, 33(2), 205-231.
- Edworthy, J., & Stanton, N. (1995). A user-centered approach to the design and evaluation of auditory warning signals: 1. Methodology. *Ergonomics*, 38(11), 2262-2280.
- Haas, E. C. & Casali, J. G. (1995). Perceived urgency of and response time to multi-tone and frequency-modulated warning signals in broadband noise. *Ergonomics*, 38, 2313-2326.



- Johnson, W. W., Bilimoria, K. D., Thomas, L. C., Lee, H. Q., & Battiste, V. (2003). Comparison of pilot and automation generated conflict resolutions. In *Proceedings of the American Institute of Aeronautics and Astronautics Conference* (pp. 1-10).
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York, NY: Cambridge University Press.
- Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46, 50-80.
- Moray, N. (1981). The role of attention in the detection of errors and the diagnosis of failures in human-machine systems. In J. Rasmussen and W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* (pp. 185-198). New York, NY.
- Parasuraman, R., & Riley V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230-253.
- Ratches, J. A., Walters, C. P., Buser, R. G., & Guenther, B.D. (1997). Aided and automatic target recognition based upon sensory inputs from image forming systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, 1004 – 1019.
- Sorkin, R. D., & Woods, D. D. (1985). Systems with human monitors: A signal detection analysis. *Human Computer Interaction*, 1, 49-75.
- Stanton, N. S., Ragsdale, S. A., & Bustamante, E. A. (2009). The effects of system technology and probability type on trust, compliance, and reliance. In *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* (pp.1368-1372). Santa Monica, CA: Human Factors and Ergonomics Society.
- Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Thomas, L. C., Wickens, C. D., & Rantanen, E. M. (2003). Imperfect automation in aviation traffic alerts: A review of conflict detection algorithms and their implications for human factors research. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 344-347). Santa Monica, CA: Human Factors and Ergonomics Society.
- Tvarynas, A.P., Thompson, B.T., & Constable, S.H. (2005). "US Military Unmanned Aerial Vehicle Mishaps: Assessment of the Role of Human Factors using HFACS." 311th Performance Enhancement Directorate, US Air Force, Brooks AFB, TX, 2005.
- Wickens, C. D. (1987). Information processing, decision-making, and cognition. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 72-107). Oxford, England: John Wiley & Sons.
- Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors*, 37(3), 473.

- Wickens, C. D., & Dixon, S. R. (2007). The benefits of imperfect diagnostic automation: A synthesis of the literature. *Theoretical Issues in Ergonomics Science*, 8(3), 201-212.
- Wiegmann, D. & Shappell, S. (2003). A human error approach to aviation accident analysis: The human factors analysis and classification system. Aldershot, Great Britain: Ashgate Publishing Company.
- Wilson, J. R. (2005). The evolution of UAV avionics. *Military and Aerospace Electronics*, 16(9), 39-40, 42-45.
- Wogalter, M. S., Allison, S. T., & McKenna, N. A. (1989). Effects of cost and social influence on warning compliance. *Human Factors*, 31, 133-140.
- Yang, L. C., & Kuchar, J., K. (1997). Prototype conflict alerting system for free flight. *Journal of Guidance, Control, and Dynamics*, 20(4).



## ***Pilot Source Study: An Analysis of Pilot Backgrounds and Subsequent Success in US Regional Airline Training Programs***

Guy M. Smith

*Embry-Riddle Aeronautical University*

David A. NewMyer

*Southern Illinois University*

Elizabeth Bjerke

*University of North Dakota*

Mary Niemczyk

*Arizona State University*

Raymond A. Hamilton

*Auburn University*

*PO Box 730298*

*Ormond Beach, FL 32173-0298*

*386-236-8494*

[ypsilon\[at\]earthlink.net](mailto:ypsilon[at]earthlink.net)

### ***Abstract***

*The 2010 Pilot Source Study, commissioned to research the success of pilots in initial training for Part 121 operations, analyzed the training performance of 2,156 new-hire pilots in the years 2005-2009. Six regional airlines provided data that was mined from human resource and pilot training files. Five university researchers independently analyzed the data and integrated their results. The study expressed success in terms of fewer extra training events and fewer non-completions in regional airline training. Statistically, the best performing pilots were those who had flight instructor certificates, graduated from collegiate accredited flight programs, received advanced (post-Private) pilot training in college, graduated with collegiate aviation degrees (any aviation discipline), and had between 500 and 1,000 pre-employment flight hours. Pilot source characteristics that had no significance in regional airline pilot training success were: having a non-aviation college degree and having prior corporate pilot or airline pilot experience.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

## **Analysis of Pilot Backgrounds and Subsequent Success in US Regional Airline Training Programs**

Public and legislative attention is currently focused on the appropriate levels of training and the qualification requirements for United States airline pilots. Involved citizens and Congress have expressed concerns about pilot performance and professionalism, issues that were highlighted by the Colgan Air (operating as Continental Connection Flight 3407) accident in a DHC-8 on February 12, 2009, outside of Buffalo, New York. The accident focused attention on whether commercial copilots are adequately prepared prior to their training at a regional airline. Prompted by the Colgan Air accident, the US House of Representatives passed the *Airline Safety and Pilot Training Improvement Act of 2009* (H.R. 3371, 2009) to amend Title 49 of the United States Code with the intent to improve airline safety and pilot training. Similar legislation was introduced in the US Senate – *Enhancing Flight Crewmembers’ Training* (S. 1744, 2009) requiring the Federal Aviation Administration (FAA) Administrator to prescribe regulations to ensure that all crewmembers on air carriers have proper qualifications and experience. As of March, 2010, the language from H. R. 3371 and S. 1744 was been combined into two bills being considered by Congress under the general heading of “FAA Reauthorization,” namely S. 1451(2010) and H. R. 1586 (2010).

Consequently, the FAA issued an Advance Notice of Proposed Rulemaking (ANPRM) titled *New Pilot Certification Requirements for Air Carrier Operations* (FAA, 2010). The purpose of this notice was to gather information on whether current eligibility, training, and qualification requirements for commercial pilot certification were adequate for conducting domestic, flag, and supplemental operations (FAA). The ANPRM requested public comment on the necessity to improve pilot performance and professionalism standards with specific emphasis on training for commercial pilots involved in Part 121 operations. The FAA sought input and recommendations on five concept areas, each of which included a series of questions.

In the ANPRM, Question 2A asked, “Are aviation/pilot graduates from accredited aviation university degree programs likely to have a more solid academic knowledge base than other pilots hired for air carrier operations? Why or why not?” (FAA, 2010, p. 7). To answer this question thoroughly and accurately, a consortium of educators, regional airlines, the Aviation Accreditation Board International (AABI) and the University Aviation Association (UAA) commissioned a study to determine the performance outcomes of new pilot indoctrination for first officers in Part 121 operations.

### **Background**

Accreditation is a system for recognizing educational programs that meet a defined set of standards – granted by private organizations (AABI, n.d.). There are two types of Accreditors: (a) Institutional accreditors that review and accredit entire institutions; and (b) Program accreditors that review and accredit specific programs or subject area offerings within an institution (AABI). AABI is a program accreditor that focuses on collegiate non-engineering aviation education for both two-year and four-year programs. AABI is one of 46 specialized accreditation organizations recognized by the Council for Higher Education Accreditation (CHEA) (CHEA, 2010). In the FAA (2010) ANPRM, Question 2A requests information about

accredited aviation university degree programs; AABI is the only body that accredits its aviation university degree programs.

Another organization that represents collegiate aviation is the University Aviation Association (UAA), “the voice of collegiate aviation education to its members, the industry, government and the general public” (UAA, n.d., homepage). UAA is a nonprofit organization including aviation high schools, 2-year colleges, and 4-year universities that have aviation programs. UAA represents all segments of aviation education, including flight programs. UAA is not an accrediting organization; however, many of the colleges and universities that have AABI Accredited Flight Programs are active members of UAA.

On February 19, 2010, at a combined meeting of AABI and UAA, members were challenged by the two presidents to provide collegiate aviation support for the FAA Administrator’s goals on pilot qualification regulatory initiatives. This study was commissioned to research the success of new pilot indoctrination for first officers in Part 121 operations. The goal of the study was to provide empirical data concerning characteristics of the sources of pilot training that related to the pilots’ success in regional airline training. The ultimate goal is to make it possible for talented young people to select “airline pilot” as an aviation career and to support the aviation industry with a strong cadre of enthusiastic candidates in the pilot supply chain. With the support of AABI and UAA, researchers from five independent universities and six regional airlines developed this study to analyze the performance data of pilots hired at these carriers between 2005-2009.

### **Review of the Literature**

Over the years, significant research has been conducted on predicting pilot training success. Much of this research (Hunter & Burke, 1994; Carretta & Ree, 1996; Martinussen, 1996; Damos, 1996; Griffin & Koonce, 1996) focused on military pilot selection and training success. Due to the high cost of training failures and stagnant attrition numbers, militaries from numerous countries have conducted a wide range of studies to evaluate selection measures.

In a meta-analysis of 68 published studies, Hunter and Burke (1994) utilized a method of validity generalization to assess which predictor measures were most significant. The most significant predictive measures were found in the following: quantitative ability, spatial ability, mechanical ability, aviation information, general information, gross dexterity, perceptual speed, reaction time, biographical information, and job sample. In a separate meta-analysis, Martinussen (1996) compared samples from 50 studies conducted in 11 different countries. This research found that the best predictor of pilot performance was previous training experience and a combined index utilizing cognitive and/or psychomotor tests. Carretta and Ree (1996) analyzed the role of general cognitive ability in the selection process of military pilots that could be accomplished using numerous varying batteries.

Damos (1996) presented a critical analysis of pilot selection batteries. A major concern of the author was the use of the dichotomous pass/fail outcome variable used in the majority of the research. The author argued for utilizing a more defined operational performance measure to capture the role of a pilot. Another concern in relying on a dichotomous variable as the outcome measure was the reduction in predictive validity measures (Burke, Hobson, & Linsky, 1997). Burke et al.

found that a larger sample size was needed to guarantee a respectable statistical power.

Despite the vast amount of literature relating to military pilot selection, very little research was found on civilian pilot selection. In a study conducted within a collegiate aviation program; Mekhail, Niemczyk, Ulrich and Karp (2010) found significance when correlating scores obtained on the Table Reading Test to both flight hours to solo an aircraft and flight hours to obtain a private pilot certificate.

The pilot selection process at the regional airlines within the United States varies greatly from the *ab initio* training process utilized by both military and foreign air carriers. Pilots apply at the regional carriers after having obtained their pilot certifications and sufficient flight experience; thus, the selection variables differ from the traditional pilot selection studies. In a survey of key administrators at 11 regional airlines, it was found that the most important new-hire candidate traits included being a team player, being trainable, having good crew resource management skills, and having current flight experience (Fanjoy, Young, & Suckow, 2006). These traits were often assessed with a written knowledge test, a structured interview, and a flight simulator checkride. However, half of the respondents did not place a strong level of confidence on the ability of these instruments to predict candidate success.

In order to assure a better prediction of pilot success at a regional airline, Karp (2004) suggested utilizing a regional airline bridge training model. This training model would prepare collegiate flight education program graduates for a successful transition into the role of a regional pilot. This model included an integrated learning style, which would incorporate coursework beyond the basic flight training.

In a study conducted at one regional airline, Cortés (2008) correlated pilot background information to the subsequent success in initial training at the airline. The background variables mined in this study were the following: source of flight training, type of college degree completed, possession of a flight instructor (CFI) certificate, and total flight experience. Cortés defined success in initial training at the airline by the number of extra training events that a pilot required to complete the training program. It was found that the group with the best overall success at the regional airline consisted of individuals who graduated from an AABI Accredited Flight Program, possessed a CFI certificate, and had fewer than 500 hours of total flight time. The least successful in initial training were those trained at a commercial flight school or a Part 61 Fixed Base Operator (FBO).

### **Research Questions**

As a means to expand upon the research concerning pilot selection at the regional airlines, this study answered the following research questions:

1. What were the characteristics of pilots who were hired by the US regional airlines between 2005-2009?
2. How did these characteristics relate to their success in regional airline training programs?



## Methodology

### Participants

On February 19, 2010, at a combined meeting of the Aviation Accreditation Board International (AABI) and the University Aviation Association (UAA), a study was commissioned to research the success of new pilot indoctrination for first officers in Part 121 operations. Six regional airlines participated by providing access to their human resource and pilot indoctrination files; the regional airlines were American Eagle Airlines, Atlantic Southeast Airlines, Cape Air, Horizon Air, Mesa Airlines, and Trans States Airlines. Seven colleges or universities, matched with these airlines, assisted with data collection. The research project studied the performance of 2,156 new-hire pilots in the years 2005 – 2009.

### Procedures

There were three constraints on the study: (a) a requirement that all variables had to be common among the six regional airlines so their data could be combined; (b) an agreement that the analyzed dataset would not have identification data for a specific pilot or airline; and (c) a requirement for researchers to collect and analyze the data with a neutral perspective that did not attempt to favor any interested party.

SurveyMonkey (2010) was selected as an online data collection device because it uniformly organized the data, automatically collected the data in a spreadsheet, and provided a common vehicle for transmitting de-identified data from the regional airlines to the principal investigator. Representatives from the partner universities (professors, graduate research assistants, or interns) entered the airline data into SurveyMonkey. These data had to be mined from two separate departments in the regional airlines – the Training Department and the Human Resource Department.

Demographic data were gathered for the subject pilots, including: year hired, college degree, name of college, name of degree, military background, where the pilot received advanced (beyond private pilot) training, whether the pilot had previous experience as a flight instructor, total flight hours at the time of indoctrination, and previous experience as a corporate or airline pilot. To de-identify the data, two variables (whether the degree was in an aviation concentration, and whether the degree was from an AABI Accredited Flight Program) were derived so that the "name of college" and "name of degree" data could be removed. The outcomes studied were: (a) how many times did the pilot need to repeat the elements of indoctrination training, and (b) whether the pilot completed the full training program at the airline. The individual pilot and airline information are de-identified in the study.

Five independent university researchers from Arizona State University, Auburn University, Embry-Riddle Aeronautical University, Southern Illinois University, and the University of North Dakota independently analyzed the data using the SPSS data mining and statistical analysis software and integrated their results through a series of conference calls. Consensus among the researchers was reached by a process that considered inputs from each researcher, reconciled any conflicting arguments, and concluded that there was no opposition to substantial results

and conclusions. Additionally, a draft of the report was sent to their constituents requesting comment and feedback; there were no responses that would invalidate the results.

## Limitations

There were limitations on the type and amount of data that could be collected from a rich source of human resources and training data maintained by airlines. Data were collected from six airlines on new-hire pilots in the years 2005-2009; however, incomplete data sets from several airlines prevented an analysis by year hired.

Since there were no standard pilot evaluation processes or uniform training records, data were mined from an assortment of records – both paper and electronic. Some data were not available at all of the carriers. Additionally, airline human resources and training personnel rightfully guarded company records and required stringent control and protection of their data, even after researchers were granted access to some of it. Consequently, the study was limited to pilot characteristics and success data that were common across all six airlines.

Effect size (Cohen's  $d$  for  $t$ -tests and ANOVA; Cohen's  $w$  for Chi-Square) was included in the reporting of all significant results. Although the null hypothesis significance tests showed that the means were significantly different; the effect sizes were small to modest, meaning that the factor accounted for a small or modest percent of the relationship between pilot source data and regional airline training data. Small effect sizes were anticipated for this study because, in many cases, the outcome variables (associated with regional airline training) were removed by several years from the income variables (associated with the source of a pilot's foundational training). According to Trusty, Thompson, and Petrocelli (2004), "Small effect sizes for very important outcomes can be extremely important, as long as they are replicable" (p. 110).

## Results

The six regional airlines and their affiliated institutions entered 2,187 records into the online data collection device. Several records were purged because they contained obvious data entry errors (duplicate records, blank records, etc.), leaving 2,156 valid records for data analysis. The records from the six airlines were combined into a single dataset and all identifying information was removed. In the following analysis, the statistical assumptions and conditions were met unless otherwise noted.

### Outcome Variable: Extra Training Events

The dependent variable, *Extra Training Events*, as suggested by Cortés (2008), was defined as, "How many repeat training events at your airline did this pilot require BEFORE IOE (Initial Operating Experience)? NOTE: Training events - anything that required a PASS grade (Ground Schools, Exams, Procedure Trainers, Simulators, LOFT, etc.)". The variable, *Extra Training Events*, is described in Table 1.

Table 1

*Extra Training Events*

Extra Training Events	
Mean	0.950
Median	0
Mode	0
Standard Deviation	1.537
Range	12
Minimum	0
Maximum	12
Count	2156

Since the Pilot Source Study was a large sample (N = 2156), the variable was treated as a scale variable and parametric tests were considered robust. Graphical analysis of Extra Training Events suggests that one-tail *p* values are appropriate (Motulsky, 1999). According to Motulsky, for large samples (> 100) the *p* value will be nearly correct even if the population is fairly far from Gaussian.

**Outcome Variable: Completions**

The dependent variable *Completions* was defined as, “Did this pilot complete the training with your airline (including IOE)?” The dependent variable *Completions* was not parsed because the airlines would not disclose reasons for non-completion. The dichotomous variable *Completions* is described in Table 2.

Table 2

*Completed Training (Including IOE)*

Completed Training (Including IOE)		
Yes	2035	94%
No	121	6%
Total	2156	100%

**Predictor Variable: Flight Instructor**

The independent variable Flight Instructor was defined as, “INSTRUCTOR: Was this pilot an FAA certificated flight instructor (CFI, CFII, MEI, etc.?).” Of the 2,156 pilots, 1,583 (73%) were certificated flight instructors and 573 (27%) were not. Flight Instructor has confounding variables, most notably the number of hours spent in flight instructing. In a follow-on question, the surveyor instrument collected

*Hours-of-Dual-Given*; however, excessive missing data made *Hours-of-Dual-Given* unreliable.

A Chi-Square test of significance compared *Completions* for pilots who were flight instructors and for pilots who were not flight instructors. In Table 3, the results show that pilots who were not flight instructors had comparatively more non-completions.

Table 3

*Comparison of Number of Completions Between Flight Instructors and Other Pilots*

Completions		Flight Instructor (YES)	Flight Instructor (NO)
Complete (YES)	Observed/Expected $\chi^2$ Contribution	1509/1494 1%	526/541 4%
Complete (NO)	Observed/Expected $\chi^2$ Contribution	74/89 25%	47/32 69%

$\chi^2 (1,1) = 9.884, p = .0017, \text{Cohen's } w = .068.$

A two-sample one-tailed *t*-Test (assuming unequal variances) tested for differences in *Extra Training Events* between pilots who were flight instructors and pilots who were not flight instructors. Table 4 displays the results – Pilots who were flight instructors had fewer *Extra Training Events* than pilots who were not flight instructors.

Table 4

*Comparison of Extra Training Events Between Flight Instructors and Other Pilots*

	Flight Instructor (NO)	Flight Instructor (YES)
Mean	1.14	0.88
Variance	2.60	2.26
Observations	573	1583
df	955	
<i>t</i> Stat	-3.987***	
$P(T \leq t)$ one-tail	0.00004	
Cohen's <i>d</i>	.167	
<i>t</i> Critical one-tail	1.65	

\*\*\* $p < .001.$

**Predictor Variable: AABI Accredited Flight Program**

AABI Accredited Flight Program was derived from three entries in the online data collection device: (a) “COLLEGE: What college/university did the pilot graduate from? If unknown, enter U.” (b) “DEGREE TYPE: What undergraduate degree did the pilot have?” and (c) “DEGREE NAME: What was the name of the undergraduate college degree? If unknown, enter U.” These three entries were compared against the list of AABI Accredited Flight Programs dated September 18, 2009, provided to the researchers by AABI. It is important to note that AABI accredits programs, not institutions; so a pilot was counted in AABI Accredited Flight Program only if that pilot graduated from a college or university on the list *and* if the pilot’s degree type and degree name matched the program name of the AABI Accredited Flight Program on the list. Of the 2,156 pilots, 616 (29%) were graduates of AABI Accredited Flight Programs, while 1,540 (71%) were not.

A Chi-Square test of significance compared *Completions* for pilots who graduated from AABI Accredited Flight Programs and all other pilots in the dataset. The results in Table 5 show that graduates of AABI Accredited Flight Programs had comparatively fewer non-completions.

Table 5

*Comparison of Number of Completions Between AABI Graduates and Other Pilots*

Completions		AABI (YES)	AABI (NO)
Complete (YES)	Observed/ Expected $\chi^2$ Contribution	601/581 4%	1434/1454 2%
Complete (NO)	Observed/Expected $\chi^2$ Contribution	15/35 67%	106/86 27%

$\chi^2 (1, 1) = 16.434, p = .00005, \text{Cohen's } w = .087.$

A two-sample one-tailed *t*-Test (assuming unequal variances) tested for differences in *Extra Training Events* between pilots who graduated from AABI Accredited Flight Programs and all other pilots in the dataset. Table 6 shows the results – Pilots who graduated from AABI Accredited Flight Programs produced fewer *Extra Training Events*.

Table 6

*Comparison of Extra Training Events Between AABI Graduates and Other Pilots*

	AABI (NO)	AABI (YES)
Mean	1.08	0.64
Variance	2.69	1.42
Observations	1540	616

df	1545
t Stat	6.09***
P(T<=t) one-tail	0.0000
Cohen's d	.307
t Critical one-tail	1.65

\*\*\*p < .001.

### Predictor Variable: Source of Pilot Training

The independent variable, Source of Pilot Training, was defined as, "PILOT TRAINING: Where did this pilot get Advanced Pilot Training (beyond Private Pilot)?" The entries for advanced pilot training were: College = 994 (46%); Military = 55 (3%); Non-college Part 141/142 = 670 (31%), and Non-college Part 61 = 437 (20%). It should be noted that college flight programs are also taught under Part 61, Part 141, or Part 142; however, those data were not collected for this study.

A Chi-Square test of significance compared *Completions* among the four sources of pilot training – Table 7. Post hoc analysis ( $\chi^2$  Contribution) shows two significant results: pilots trained in Colleges had comparatively fewer non-completions and pilots trained in Non-college Part 141/142 programs had comparatively more non-completions.

Table 7

#### Comparison of Number of Completions Based on Sources of Pilot Training

Completions		College	Military	Non-college (Part 141 or 142)	Non- college (Part 61)
Complete	Observed/ Expected	966/938	49/52	612/632	408/413
(YES)	$\chi^2$ Contribution	3%	1%	2%	0%
Complete	Observed/ Expected	28/56	6/3	58/38	29/25
(NO)	$\chi^2$ Contribution	46%	9%	37%	3%

$\chi^2 (3, 1) = 30.163, p = .00000, \text{Cohen's } w = .118.$

A one-way Analysis of Variance (ANOVA) tested for differences in *Extra Training Events* among the four sources of pilot training. The results, shown in Table 8, suggest that pilots trained in Colleges had fewer *Extra Training Events* than pilots trained in Non-college Part 141/142 programs ( $p < .001$ ) and pilots trained in Non-college Part 61 programs ( $p < .05$ ).

Table 8

*Analysis of Variance for Extra Training Events Based on Source of Pilot Training*

PILOT TRAINING	Mean	SD	N		
College	.76	1.29	994		
Military	1.16	1.69	55		
Non-college Part 141/142	1.16	1.72	670		
Non-college Part 61	1.04	1.69	437		
Source of variation	SS	df	MS	F	Sig.
PILOT TRAINING	72.66	3	24.22	10.39***	.000
Error	5017.03	2152	2.33		
Total	7037.00	2156			
Cohen's <i>d</i>	.139				
Scheffe Tests		Significance			
College vs. Non-college Part 141/142***		.000			
College vs. Non-college Part 61*		.016			

\* $p < .05$ . \*\*\* $p < .001$ .

**Predictor Variable: Aviation Degree**

Aviation Degree was derived from a comprehensive variable in the online data collection device: "DEGREE NAME: What was the name of the undergraduate college degree? If unknown, enter U." A pilot was counted in Aviation Degree, if that pilot earned any degree that contained words like aviation, flight, airport, pilot, etc. It is important to note that this variable contained a wide variety of aviation disciplines; these were not all flight degrees. Of the 2,156 pilots, 1,144 (53%) had aviation degrees; the other 1,012 (47%) had either a non-aviation degree or no degree.

A Chi-Square test of significance compared *Completions* between pilots who graduated with a degree in aviation and all other pilots in the dataset. The results in Table 9 show that pilots who graduated with a degree in aviation had comparatively fewer non-completions; pilots with degrees other than aviation or with no degree had comparatively more non-completions.



Table 9

*Comparison of Number of Completions Between Pilots With an Aviation Degree and Other Pilots*

Completions		Aviation Degree (YES)	Aviation Degree (NO)
Complete (YES)	Observed/Expected $\chi^2$ Contribution	1095/1080 3%	940/955 3%
Complete (NO)	Observed/Expected $\chi^2$ Contribution	49/64 44%	72/57 50%

$\chi^2 (1, 1) = 8.127, p = .0044, \text{Cohen's } w = .061.$

A two-sample one-tailed *t*-Test (assuming unequal variances) tested whether there was any difference in *Extra Training Events* between pilots who graduated with a degree in aviation and all other pilots in the dataset. The results, depicted in Table 10, show that pilots who graduated with a degree in aviation had fewer *Extra Training Events* than other pilots in the dataset.

Table 10

*Comparison of Extra Training Events Between Aviation Graduates and Other Pilots*

	Aviation Degree (YES)	Aviation Degree (NO)
Mean	0.87	1.04
Variance	2.12	2.63
Observations	1144	1012
df	2047	
<i>t</i> Stat	1.71*	
$P(T \leq t)$ one-tail	0.04	
Cohen's <i>d</i>	.110	
<i>t</i> Critical one-tail	1.65	

\* $p < .05.$

**Predictor Variable: Total Flight Hours**

The independent variable, Total Flight Hours, was defined as, "HOURS: How many Total Hours did the pilot have at the beginning of training with your airline?" Six entries in Total Flight Hours had missing data; thus  $N = 2150$ . This scale variable is described in Table 11. Since the variance and range were so wide-ranging, the researchers agreed to treat Total Flight Hours as a categorical variable, also described in Table 11. The categories were chosen to be factors of 1,500 hours, the total pilot time required for an Air Transport Pilot certificate under Part 61.159.

Table 11

*Total Flight Hours Described as a Scale Variable and Categorical Variable*

Total Flight Hours (Scale Variable)		Total Flight Hours (Categorical Variable)	
Mean	1,312.51	Range	
Median	913	178 to 500 Hours	405
Standard Deviation	1,618.05	501 to 1,000 Hours	780
Variance	2,618,088.43	1,001 to 1,500 Hours	459
Range	21,498	> 1,500 Hours	506
Minimum	178		
Maximum	21,676		
Count	2,150		

A Chi-Square test of significance compared Completions based on the number of Total Flight Hours. The results in Table 12 show that pilots with 501 to 1,000 total flight hours had comparatively fewer non-completions.

Table 12

*Comparison of Number of Completions Based on Total Flight Hours*

Completions	HOURS	0-500	501-1000	1001-1500	> 1500
Complete	Observed/ Expected	387/382	753/736	422/433	466/477
(YES)	$\chi^2$ Contribution	0%	2%	2%	1%
Complete	Observed/ Expected	18/23	27/44	37/26	40/29
(NO)	$\chi^2$ Contribution	6%	38%	28%	23%

$\chi^2 (3,1) = 17.242, p = .001, \text{Cohen's } w = .089.$

A one-way Analysis of Variance (ANOVA) tested for differences in *Extra Training Events* among the four categories of Total Flight Hours. The results, shown in Table 13, suggest that pilots who had 501 to 1,000 total flight hours had fewer *Extra Training Events* than pilots with > 1,500 total flight hours.

Table 13

*Analysis of Variance for Extra Training Events Based on Total Flight Hours*

TOTAL FLIGHT HOURS	Mean	SD	N		
0 to 500	.92	1.42	405		
501 to 1000	.85	1.34	780		
1001 to 1500	.96	1.56	459		
> 1500	1.12	1.85	506		
Source of variation	SS	df	MS	F	Sig.
TOTAL FLIGHT HOURS	23.39	3	7.80	3.31*	.019
Error	5058.28	2145	2.36		
Total	7022.00	2149			
Cohen's <i>d</i>	.079				
Scheffe Tests	Significance				
501 to 1000 vs. >1500	.022*				

\* $p < .05$ .

**Predictor Variable: College Degree**

The independent variable, College Degree, was defined as, “COLLEGE DEGREE: Did this pilot have a college degree (any discipline) at the beginning of training with your airline? NOTE: Consider completed undergraduate degrees only.” The only data entry options for College Degree were: Associate Degree, Bachelor’s Degree, or No Degree. Of the 2,156 pilots, 245 (11%) had an Associate Degree; 1,563 (73%) had a Bachelor’s Degree; and 348 (16%) had no degree.

A Chi-Square test of significance compared Completions among the three options for College Degree. The results,  $\chi^2(2,2) = 2.41$ ;  $p = .300$ , showed that no relationship existed between the number of non-completions and the Types of College Degrees.

A one-way Analysis of Variance (ANOVA) tested for differences in Extra Training Events among the three entries for College Degree (Associate, Bachelor’s, None). The results,  $F(2, 2153) = 1.16$ ,  $p = .315$ , show no difference in the number of Extra Training Events based on having an Associate, Bachelor’s (any discipline) or no college degree.

**Predictor Variable: Military**

The independent variable, Military, was defined as, “MILITARY: What prior military experience did this pilot have?” The tallied results for Military were: None - 1941 (90%); Military Aviator, Pilot (Fixed Wing) – 61 (2.8%); Military Aviator, Pilot (Rotary Wing) – 7 (.3%); Military Aviator, Non-Pilot (e.g., NFO, WSO, Bomb-Nav) – 18 (.8%); and Military, Non-Aviator – 129 (6%). Of the 2,156 pilots, only 68 were former military pilots. Of note, the small number of military pilots (N = 68, 3% of the dataset) in this group corroborates the belief that military pilots usually seek employment with the major airlines rather than with the regional airlines.

A Chi-Square test of significance compared *Completions* among pilots with prior military experience and all other pilots in the dataset. The results,  $\chi^2(1,1) = 0.839$ ;  $p = .360$ , show no difference in completions between pilots with or without military experience.

A two-sample one-tailed *t*-Test (assuming unequal variances) tested whether there was a difference in *Extra Training Events* between pilots with previous military experience ( $M = 1.04$ ,  $SD = 1.56$ ), and all other pilots in the dataset ( $M = 0.94$ ,  $SD = 1.53$ ). The results,  $t(262) = 0.42$ ,  $p = 0.34$ , show that pilots with prior military experience had the same number of *Extra Training Events* as other pilots in the dataset.

### **Predictor Variable: Previous Experience as a Corporate or Airline Pilot**

The independent variable, Previous Experience, was defined as, "PREVIOUS EXPERIENCE: What previous corporate or airline pilot experience did this pilot have?" The selections for Previous Experience were: None, Previous Corporate Pilot, or Previous Airline Pilot. If the pilot had previous airline experience, a follow-up question asked, "If Previous Airline Pilot, what airline?" The qualitative data from this follow-on question was deleted from the dataset because the answers were indiscriminate and because the data held potential identification information. The tallied results for Previous Experience were: None:1658 (77%); Previous Corporate Pilot:148 (7%); and Previous Airline Pilot: 350 (16%).

A Chi-Square test of significance compared *Completions* among the three categories of previous experience. The results,  $\chi^2(2,1) = 4.76$ ;  $p = .092$ , show that pilots with previous airline or corporate experience had the same proportion of non-completions as pilots with no previous experience.

A one-way Analysis of Variance (ANOVA) tested whether there was a difference in *Extra Training Events* among the three categories of Previous Experience. The results,  $F(2, 2153) = 2.51$ ,  $p = .081$ , show that pilots with previous airline or corporate experience had the same number of *Extra Training Events* as pilots with no previous experience.

## **Summary and Discussion**

The 2010 Pilot Source Study began with the following research questions: (a) "What were the characteristics of pilots who were hired by the US regional airlines between 2005-2009?" and (b) "How did these characteristics relate to their success in regional airline training programs?"

### **Characteristics of New-hire Pilots**

The data that described the characteristics of pilots, who were hired by the US regional airlines, resides in the individual airline's human resources department in the form of pilot applications; interviews; and, in some cases, simulator evaluation reports; psychological test results; medical evaluations; etc. Because of the assortment of the data sources, the sensitivity of the data, and the need for uniformity of data; the pilot characteristics examined in this study are a small sample of the abundant data that may be available.

Using the data from the 2,156 pilots at the six contributing airlines, the characteristics of pilots who were hired between 2005 and 2009 by the US regional airlines were:

- 1,563 (72.5%) received a bachelor's degree, while 245 (11.36%) received an associate degree, and 348 (16.14%) had no degree at all.
- 1,144 (53.1%) had a degree in an aviation discipline.
- 616 (28.6%) were determined to have a degree from a collegiate flight program that was accredited under the Aviation Accreditation Board International (AABI) Program Criteria for Flight Education (AABI, 2008).
- 215 had a military background of which 68 (3.2%) were military pilots.
- 994 (46.1%) received their advanced pilot training (beyond Private Pilot) in a collegiate flight program (conducted under Part 61, 141 or 142); 670 (31.1%) received their advanced pilot training in non-college flight programs (conducted under Part 141 or 142); 437 (20.3%) received their advanced pilot training in non-college flight programs (conducted under Part 61); and 55 (2.6%) received their advanced pilot training in the military.
- 1,583 (73.4%) were flight instructors.
- All had records of accumulated flight hours that ranged from 178 to 21,676 hours, broken into four categories with the following distributions:
  - 1) 0 to 500 hours: 405 (18.8%)
  - 2) 501 to 1,000 hours: 780 (36.3%)
  - 3) 1,001 to 1,500 hours: 459 (21.3%)
  - 4) Above 1,500 hours: 506 (23.5%)
- 1,658 (76.9 %) had no prior corporate pilot or airline pilot experience, 350 (16.2%) had prior airline pilot experience, and 148 (6.9 %) had prior corporate pilot experience.

Another way to describe the characteristics of the 2,156 pilots in this study is that more than half of them had a baccalaureate degree, had an aviation degree, were flight instructors, had 1,000 or fewer hours of flight time, and had no prior airline pilot or corporate pilot experience.

### **Success in Regional Airline Training Programs**

Because of the assortment of the data in training departments, the sensitivity of training data, and the need for uniformity of data; only two success variables were mined from all of the airlines. These key outcomes were: (a) the number of extra training events (repeats) that the pilots experienced in initial airline training before their Initial Operating Experience (IOE) and (b) whether the pilots succeeded in completing their initial pilot training (including IOE). The study found the following:

- The number of extra training events experienced by the pilots were:
  - 1) Zero = 1,310 (60.8%)

- 2) One = 257 (11.9 %)
  - 3) Two = 298 (13.8 %)
  - 4) Three = 136 (6.3 %)
  - 5) Four = 75 (3.5%)
  - 6) Greater than four = 80 (3.7 %)
- A total of 2,035 (94 %) of the new-hire pilots completed initial training with a regional airline, while 121 (6 %) did not.

## **Relationships Between Pilot Characteristics and Training Success**

Appendix A is a statistical summary of the *2010 Pilot Source Study*. Through the application of ANOVA, Chi-Square, and *t*-Test statistics, the following conclusions were drawn about the relationship between the characteristics of pilots hired by US regional airlines between 2005 and 2009 and their success in regional airline training (as defined in the outcome variables, *Extra Training Events and Completions*):

Having a college degree (Associate or Bachelor's) did not produce a difference in the number of extra training events during initial training with a regional airline; nor did it produce a significant relationship with the number of non-completions in initial training. However, if the college degree was an aviation degree (any aviation discipline), then the relationship changed. Having an aviation degree produced fewer extra training events and comparatively fewer non-completions in initial training. More significantly, if pilots earned their college degree in an AABI Accredited Flight Program, they had fewer extra training events and fewer non-completions in initial training.

The source of advanced pilot training was defined in the online data collection device as "where the pilot earned his/her advanced training (beyond the Private Pilot Certificate)." Pilots, who received their advanced training in college, subsequently had fewer extra training events and comparatively fewer non-completions in regional airline training programs. Pilots with a military background did not have the same result; however, the small number of military pilots in the data set precludes any meaningful conclusions about military-trained pilots. Pilots in this dataset who received their advanced training in non-college Part 141/142 programs or in non-college Part 61 programs did not perform as well as their collegiate counterparts.

Previous flying experience, beyond advanced pilot training, produced interesting results. Pilots who attained their flight instructor certification had fewer extra training events and comparatively fewer non-completions in their initial training at the regional airline. On the other hand, having previous experience as a corporate pilot or as an airline pilot did not produce a difference in the number of extra training events nor did it produce a significant relationship with the number of non-completions in initial training.

Total flight hours was treated as a categorical variable rather than a continuous variable to negate the effects of large numbers for relatively few pilots at the top of the scale and because this study was mostly interested in the success of new-hire pilots with fewer than 1,500 hours. One category of pilots, those with 501 to 1,000 hours, had comparatively fewer extra training events than pilots in any other total flight hour category. This same category had comparatively fewer non-completions. The effect of Total Flight Hours, in order of performance was: Group 1 (501-1000 hours), Group 2 (178-500 hours), Group 3 (1001-1500 hours) and Group 4 (greater than 1500 hours). The most significant difference was between Group 1 and Group 4 for both *Extra Training Events* and *Completions*. This result is counter-intuitive; it is generally expected that more flight hours will yield better performance. Extraneous variables may be confounding the results for this cohort of new-hire regional airline pilots with more than 1,500 hours; however, no data collected for this study was able to explain the result.

### **Recommendations for Further Study**

For further research on this subject, it may be advantageous to pursue a larger, more comprehensive study of pilots hired at regional airlines that includes more regional airlines and more pilot subjects. Expanding the current study will provide a more complete examination of the characteristics of new-hire pilots and the relationships of these characteristics to their success in initial training.

One limitation of this study was the wide array of data and the varied data storage methods among the regional airlines. Before conducting a follow-on study, it would be advantageous for researchers to conduct preliminary work with additional cooperating airlines to develop an understanding of the strengths and limitations of the data available in human resource records and pilot training records that are routinely kept by the airlines.

The *2010 Pilot Source Study* was intentionally unbranded, unsponsored, and unfunded to make the study resistive to special interest criticism. As a result, the regional airlines and the cooperating universities absorbed the financial burden of collecting and analyzing the data. Researchers should pursue funding sources for further studies; otherwise, the cost of data mining at an even larger sampling of airlines could be prohibitive.

This study was limited to examining the effects of single variables on the two outcome variables. A future study that includes multivariate analysis of the relationships of pilot characteristics to success might provide deeper insight into the subject matter.

The data suggests that there might be value added to the development of pilot skills by a comprehensive education over a 2-year or 4-year college career. This appears to be a subject ripe for further study.

The subject of pilot characteristics and their relationship to regional airline training success seems to be a fitting subject for the application of Data Envelopment Analysis (DEA) and other business models, which would assess the ability to produce a student training output with a minimum resource level, required (Cooper, Seiford, & Tone, 2007).



Because there were areas where significant positive relationships were found between a particular pilot characteristic and success in initial regional airline pilot training, it is recommended that the components of any one of those characteristics (an AABI Accredited Flight Program or advanced flight training in college) be studied for additional depth of understanding of these relationships.

## References

- Airline Safety and Pilot Training Improvement Act of 2009*, H.R. 3371, 111<sup>th</sup> Cong. (2009).
- Aviation Accreditation Board International. (2008, July) *Accreditation criteria manual* (Form 201). Auburn, AL: Author. Retrieved April 12, 2010 from <http://www.aabi.aero/>
- Aviation Accreditation Board International. (n.d.). *Accredited programs*. Retrieved April 12, 2010, from <http://www.aabi.aero/programs.html>.
- Burke, E., Hobson, C., & Linsky, C. (1997). Large sample validations of three general predictors of pilot training success. *The International Journal of Aviation Psychology*, 7(3), 225-234.
- Carretta, T., & Ree, M. (1996). Central role of g in military pilot selection. *The International Journal of Aviation Psychology*, 6(2), 111-123.
- Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data development analysis* (2nd ed.). New York: Springer Science and Business Media.
- Cortés, A. (2008, March). *Newhire pilot yield: The pilot production pipeline and industry cyclicalilty*. Paper presented at the National Training Aircraft Symposium, Daytona Beach, FL.
- Council on Higher Education Accreditation. (2010, February). *CHEA Annual Report 2008-2009*. Retrieved April 12, 2010, from [http://www.chea.org/pdf/2008-2009 Annual Report12-29-09sbs2.pdf](http://www.chea.org/pdf/2008-2009%20Annual%20Report12-29-09sbs2.pdf).
- Damos, D. (1996). Pilot selection batteries: Shortcomings and perspectives. *The International Journal of Aviation Psychology*, 6(2), 199-209.
- Enhancing Flight Crewmembers' Training*, S. 1744, 111<sup>th</sup> Cong. (2009).
- FAA Air Transportation Modernization and Safety Improvement Act*, S. 1451, 111<sup>th</sup> Cong. (2010).
- FAA Air Transportation Modernization and Safety Improvement Act*, H.R. 1586, 111<sup>th</sup> Cong. (2010).
- Fanjoy, R., Young, J., & Suckow, M. (2006). The experience factor: A regional airline view of pilot candidate qualifications. *Collegiate Aviation Review*, 24(1), 67-72.
- Federal Aviation Administration. (2010, February 8). New Pilot Certification Requirements for Air Carrier Operations (ANPRM, Docket No. FAA-2010-0100; Notice No. 10-02). Washington, DC: Department of Transportation.
- Griffin, G., & Koonce, J. (1996). Review of psychomotor skills in pilot selection research of the U.S. military services. *The International Journal of Aviation Psychology*, 6(2), 125-147.
- Hunter, D., & Burke, E. (1994). Predicting aircraft pilot-training success: A meta analysis of published research. *The International Journal of Aviation Psychology*, 4(4), 297-313.
- Karp, M. (2004). Airline pilot training: A university to regional airlines bridge training model. *International Journal of Applied Aviation Studies*, 4(1), 185-204.

- Martinussen, M. (1996). Psychological measures as predictors of pilot performance: A meta-analysis. *The International Journal of Aviation Psychology*, 6(1), 1-20.
- Mekhail, A., Niemczyk, M., Ulrich, J., & Karp, M. (2010). Using the table reading test as an indicator for success in pilot training, *Collegiate Aviation Review*, 28(1), 101-114.
- Motulsky, H. J. (1999). *Analyzing data with GraphPad Prism*. San Diego CA: GraphPad Software Inc.
- SurveyMonkey [Web-based surveys]. (2010). Menlo Park, CA: Author.
- Trusty, J., Thompson, B., & Petrocelli, J. V. (2004). Practical guide for reporting effect size in quantitative research in the Journal of Counseling & Development. *Journal of Counseling and Development* 82(1), 107-111.
- University Aviation Association. (n.d.). Retrieved April 16, 2010, from <http://www.uaa.aero/>.

## Appendix A

### 2010 Pilot Source Study – Summary Results

INDEPENDENT (Predictor) VARIABLE	DEPENDENT (Outcome) VARIABLE	Statistical Tests	Test Statistic	Statistically Significant?	Conclusions
<b>FLIGHT INSTRUCTOR</b> (Yes, No) N = 2156	EXTRA TRAINING EVENTS (Range 0-12)	t-Test	$t(955) = 3.987$ , $p < .001$	YES ***	Pilots who were flight instructors had fewer extra training events than pilots who were not flight instructors.
<b>FLIGHT INSTRUCTOR</b> (Yes, No) N = 2156	<b>COMPLETION</b> (Yes, No)	Chi-Square	$\chi^2(1,1) = 9.884$ , $p < .01$	YES **	Pilots who were flight instructors had comparatively fewer non-completions.
<b>AABI ACCREDITED FLIGHT PROGRAM</b> (Yes, No, or No Degree) N = 2156	EXTRA TRAINING EVENTS (Range 0-12)	t-Test	$t(1545) = 6.09$ , $p < .001$	YES ***	AABI Accredited Flight Programs produced fewer extra training events
<b>AABI ACCREDITED FLIGHT PROGRAM</b> (Yes, No, or No Degree) N = 2156	<b>COMPLETION</b> (Yes, No)	Chi-Square	$\chi^2(1,1) = 16.43$ , $p < .001$	YES ***	AABI Accredited Flight Programs produced comparatively fewer non-completions
<b>SOURCE OF PILOT TRAINING</b> (Military, College Degree, Non-College - Part 141 or Part 142, Non-College - Part 61) N = 2156	EXTRA TRAINING EVENTS (Range 0-12)	ANOVA	$F(3,2152) = 10.39$ , $p < .001$	YES ***	Pilots trained in college had fewer extra training events than non-college pilots.
<b>SOURCE OF PILOT TRAINING</b> (Military, College Degree, Non-College - Part 141 or Part 142, Non-College - Part 61) N = 2156	<b>COMPLETION</b> (Yes, No)	Chi-Square	$\chi^2(3,1) = 30.16$ , $p < .001$	YES ***	Pilots trained in college had comparatively fewer non-completions.
<b>AVIATION DEGREE</b> (Yes, No, or No Degree) N = 2156	EXTRA TRAINING EVENTS (Range 0-12)	t-Test	$t(2047) = 1.71$ , $p < .05$	YES *	Aviation Degrees produced fewer Extra Training Events
<b>AVIATION DEGREE</b> (Yes, No, or No Degree) N = 2156	<b>COMPLETION</b> (Yes, No)	Chi-Square	$\chi^2(1,1) = 8.13$ , $p < .01$	YES **	Aviation degrees produced comparatively fewer non-completions.

INDEPENDENT (Predictor) VARIABLE	DEPENDENT (Outcome) VARIABLE	Statistical Tests	Test Statistic	Statistically Significant?	Conclusions
<b>TOTAL FLIGHT HOURS</b> (0-500 Hours, 501-1000 Hours, 1001-1500 Hours, >1500) N = 2150	EXTRA TRAINING EVENTS (Range 0-12)	ANOVA	$F(3,2145) = 3.31, p < .05$	YES *	Pilots with 501 to 1000 hours had the fewest extra training events.
<b>TOTAL FLIGHT HOURS</b> (0-500 Hours, 501-1000 Hours, 1001-1500 Hours, >1500) N = 2150	<b>COMPLETION</b> (Yes, No)	Chi-Square	$\chi^2(3,1) = 17.24, p < .01$	YES **	Pilots with 501 to 1000 hours had comparatively fewer non-completions.
<b>COLLEGE DEGREE</b> (Associate, Bachelor's, or None) N = 2156	EXTRA TRAINING EVENTS (Range 0-12)	ANOVA	$F(2,2153) = 1.16$	NO	Having a college degree did not produce a difference in number of extra training events.
<b>COLLEGE DEGREE</b> (Associate, Bachelor's, or None) N = 2156	COMPLETION (Yes, No)	Chi-Square	$\chi^2(2,2) = 2.41$	NO	There was no relationship between the number of non-completions and whether pilots had a college degree.
<b>MILITARY</b> (None, Military Pilot [FW], Military Pilot [RW], Military Aviator [Non-Pilot], Military [Non-Aviator]) N = 2156	EXTRA TRAINING EVENTS (Range 0-12)	t-Test	$t(262) = 0.42$	NO	Prior military experience had no effect on extra training events Note: The small # of military pilots (68) suggests that most military pilots go directly to the major airlines.
<b>MILITARY</b> (None, Military Pilot [FW], Military Pilot [RW], Military Aviator [Non-Pilot], Military [Non-Aviator]) N = 2156	<b>COMPLETION</b> (Yes, No)	Chi-Square	$\chi^2(1,1) = 0.84$	NO	There was no relationship between the number of non-completions and prior military experience. Note: The small # of military pilots (68) suggests that most military pilots go directly to the major airlines.

INDEPENDENT (Predictor) VARIABLE	DEPENDENT (Outcome) VARIABLE	Statistical Tests	Test Statistic	Statistically Significant?	Conclusions
<b>PREVIOUS EXPERIENCE</b> (None, Previous corporate pilot, Previous airline pilot) N = 2156	EXTRA TRAINING EVENTS (Range 0-12)	ANOVA	$F(2,2153) = 2.51$	NO	Pilots with previous airline or corporate experience had the same number of extra training events as pilots with no previous experience.
<b>PREVIOUS EXPERIENCE</b> (None, Previous corporate pilot, Previous airline pilot) N = 2156	<b>COMPLETION</b> (Yes, No)	Chi-Square	$\chi^2(2,1) = 4.76$	NO	Pilots with previous airline or corporate experience had the same proportion of non-completions as pilots with no previous experience.

\* = Significant

\*\* = Very Significant

\*\*\* = Exceptionally Significant

## ***Critical Incident Stress Management (CISM): An Effective Peer Support Program For Aviation Industries***

Jeffrey T. Mitchell, Ph.D., CTS

*Clinical Associate Professor  
Emergency Health Services  
University of Maryland*

*President Emeritus, International Critical Incident Stress Foundation*

Joerg Leonhardt

*Head of Human Factors  
German Air Navigation Services, DFS*

### *Abstract*

*The authors discuss the need for a comprehensive, integrative, systematic, and multi-component peer support program for aviation personnel. They also provide an overview of the basic principles and practices of one such program known as Critical Incident Stress Management (CISM). Successful peer support programs, modeled on the principles of CISM, have been instituted in employee classifications within the high-risk aviation field. Aviation is a high-risk organization because employees experience considerable physical and psychological threats in their work. A CISM program now serves air traffic controllers, pilots, flight attendants, and ground personnel. The article presents the basic elements of peer support teams serving high-risk organizations. Aviation is also a High Reliability Organization (HRO); an organization expected to function consistently and safely. Organized and well-trained CISM teams can help a high risk/high reliability organization become a high resiliency organization. The article details the core services provided by such teams and suggests the types and level of training necessary for the peer support personnel who serve on CISM teams within the aviation industries.*



## **Critical Incident Stress Management (CISM):**

### **An Effective Peer Support Program for Aviation Industries**

*High-risk organizations* are those in which the employees face potential exposures to extreme stress or significant threats to their physical and emotional well-being. Highly stressful or threatening circumstances are viewed as a probable hazard associated with employment in these professions. Military and emergency services personnel such as law enforcement officers, fire fighters, search and rescue personnel, disaster workers, nurses, doctors, and other medical staff are among the most well known psychologically high-risk professions. They are not, however, the only professions in which there are considerable physical and emotional risks. The physical and psychological risks that are associated with mining, logging, transportation, and commercial fishing operations are well documented (Staff, Air Safety Week, 2001; Parker, 2006; US Department of Labor, 2008 a, b; 2009; English, 2010).

Aviation related industries, as important sub-groups of the overall transportation industry, are well within the definition of high-risk organizations. The responsibilities and pressures within the aviation industries are elevated and the physical and psychological risks associated with substandard performances or accidents are great. Being on a list of high-risk organizations should not be a badge of honor for any industry. Risks must be managed and reduced wherever possible. Aviation industries must constantly strive to be high reliability organizations (HRO). These organizations must function reliably, consistently, effectively, and, above all else, safely. To their credit, aviation industries have been very successful in becoming High Reliability Organizations. Fortunately, the current overall safety record of the aviation industries depicts a low volume of significant critical incidents or accident rates compared to the number of flights handled on a daily basis.

### **Critical Incidents in Aviation**

*Critical incidents* are events that are so unusual or powerful that they can overwhelm the coping capacity of those who are exposed to them. A critical incident generates a strong sometimes tumultuous emotional reaction, which is called a crisis or occasionally a *crisis reaction*. Aviation industries are justified in their concern about crisis reactions. They are typically accompanied by cognitive, emotional, physical, and behavioral manifestations of stress. When people reach a state of intense emotional disturbance, their thinking may become disorganized and unclear. Exaggerated feelings can dominate one's reactions. When stress levels are high, mistakes may occur more frequently, and accident rates and personal injuries tend to climb. Psychological disequilibrium may appear, that is, a person's thinking ability is suppressed, and one's feelings intensify to a point of being nearly out of control. In those circumstances, crisis intervention may be required to rebalance the person and assist him or her in resolving the situation. *Crisis intervention is the active, supportive, and temporary assistance given by family members, friends, colleagues, and trained peer support personnel to others who are experiencing a period of acute distress* (Neil, Oney, DiFonso, Thacker, and Reichart, 1974; Slaikeu, 1984).

Violent passengers, serious injuries to ground crew members, severe turbulence, injured or ill passengers (particularly children), deaths of passengers, hostage takings, loss of separation between aircraft, and serious equipment failures

are just some of the critical incidents that may impact the emotional lives of people within the aviation industries. Despite the fact that airlines and related aviation industries have been quite successful in becoming high reliability organizations, such stressful events are still a relatively common occurrence. The ultimate critical incident for the aviation industry would, of course, be a major air disaster, but, thankfully, such horrific events are infrequent. Although they are high reliability organizations, aviation industries remain *high-risk organizations* for their employees. Crisis support programs like the Critical Incident Stress Management program have been designed to enhance the resiliency of employees in high-risk organizations.

*Critical Incidents in Aviation:*

- Work related deaths of colleagues
- Serious work related injuries to colleague
- Suicide of a colleague
- Accidental killing or wounding an aviation employee or a passenger
- Events involving a high degree of personnel threat
- Aircraft “Close calls” in the sky or on the ground
- Severe turbulence
- Seriously ill or injured child on board
- Major medical emergency while airborne
- Threatening passenger
- Violence from any source while in flight
- Aircraft damaged or impaired while in flight
- Fire on board an aircraft
- Hazardous materials leak on an aircraft
- Bomb threat
- Disasters involving aircraft
- Aircraft crashes
- Other distressing experiences

## CISM: Definitions, Models, And Tools

Critical Incident Stress Management (CISM) is a *comprehensive, integrated, systematic, and multi-component crisis intervention program*. It is interesting to note that the initials, CISM, can be used in two ways as it is in the previous sentence. First, CISM represents the title of the program, Critical Incident Stress Management. Second, the same initials, CISM, can be viewed as a description of the program - comprehensive, integrated, systematic, and multi-component. CISM is versatile, practical, and effective. It is a common sense stress management system. It can do much to alleviate distress and to maintain healthy levels of function for aviation staff members. Since its development in the 1970's, it has spread rapidly into many different types of agencies, organizations, and services in over a thousand communities around the world. The United Nations has recently developed its own internal CISM program to assist UN workers throughout its multi-national community (United Nations Department of Safety and Security, 2007 a, b, c).

CISM is best described as a “package” of crisis intervention tactics that are strategically woven together. These are the main objectives of a CISM program:

- *Mitigate* the impact of a traumatic event;
- *Facilitate* normal recovery processes in normal people, who are having normal reactions to traumatic events;
- *Restore* individuals, groups and organizations to adaptive function; and
- *Identify* people within an organization or a community who would benefit from additional support services or a referral for further evaluation and, possibly, psychological treatment.

CISM is a broad collection of support services and it is *not* considered a form of psychotherapy, nor is it a substitute for psychotherapy. A basic model of a CISM program includes several key components:

- Planning, education, resiliency building, policy development, and peer support team preparation.
- Assessment of the magnitude or severity of a critical incident as well as its impact on the personnel.
- Strategic planning so that the proper *targets* of an intervention are identified and the correct *types* of interventions are selected

for application at the most advantageous *time*. In developing a strategic action plan, the *themes* that might influence decision making must be considered. Finally, the most appropriate *team* must be selected to provide the best services.

- A flexible, multi-tactic approach utilizing specific models for individual or group interventions (See table 2).
- Follow-up services including phone calls, visits to work sites or homes, and advice to supervisors or administration when necessary.
- Links to professional referrals when personnel would benefit from such contacts.

As noted above, a CISM program encompasses a package of many tactics or tools that may be necessary under different circumstances. Table 2 lists the types of interventions, the targets, and timing of the interventions, as well as the potential goals of the interventions.

Table 2

*Multi-Component Elements of CISM*

Intervention	Timing	Target population	Potential goals
<b>I. Pre-event Planning/ Preparation</b>	Pre-event	Anticipated target/ victim population	Anticipatory guidance, foster resistance, resilience.
<b>II. Surveillance &amp; Assessment</b>	Pre-event & during event		
<b>III. Strategic Planning</b>	Pre-event & during event		
<b>IV. Individual Crisis Intervention</b> including "psychological first aid" (PFA) & SAFER-R	As needed	Individuals as needed	

---

**V.****Large Group Crisis Intervention**

A. "RIT"- Rest, Information and Transition (also known as "demobilization")	Shift disengagement end of initial deployment	Disaster response personnel	Decompression, ease transition, screening, triage, education and meet basic needs.
B. Respite center	Ongoing, large-scale events	Emergency personnel, large groups	Respite, refreshment, screening, triage and support.
C. Crisis Mgmt. Briefing (CMB) / large group "psychological first aid"	As needed	Heterogeneous large groups	Inform, control rumors, increase cohesion.

**VI.****Small Group Crisis Intervention**

A. Small Group Crisis Mgmt. Briefing (sCMB).	On-going & post-event; may be repeated as needed	Small groups seeking information and/or resources	Information, control rumors, reduce acute distress, increase cohesion, facilitate resilience, screening and triage.
B. Defusing (Also known as Immediate Small Group Support ISGS ) <i>Small group "psychological first aid"</i>	within 12 hours post events	Small homogeneous, primary groups	Stabilization, Ventilation, reduce acute distress, screening, information, increase cohesion, and facilitate resilience
C. CISD (also known as "Powerful Event Group Support" [PEGS])	Post-event: ~1-10 days for acute incidents; ~3-4 wks. post-disaster recovery phase	Small homogeneous, primary groups with equal trauma exposure (e.g., workgroups, emergency services, and military)	Increase cohesion, ventilation, information, normalization, reduce acute distress, facilitate resilience, screening, and triage. Follow-up essential.

---

<b>VII. Family Crisis Intervention</b>	Pre-event or post event as needed	Families	Wide range of Interventions (e.g., pre-event pre- paration, individual crisis intervention, sCMB, PEGS or other group processes.)
<b>VIII. Organizational/ Community Intervention</b>	As needed	Organizations/ communities affected by trauma or disaster response.	Improve organizational, community preparedness. Leadership consultation.
<b>IX. Pastoral Crisis Intervention</b>	As needed	Individuals, small groups, large groups, congre- gations, & com- munities who desire faith-based presence/crisis intervention	Faith-based support
<b>X. Follow-up and/or Referral Facilitating access</b>	As needed	Intervention recipients and exposed individuals	Assure continuity of care

(Everly, G.S., Jr. & Mitchell, J.T., (2008). *Integrative Crisis Intervention and Disaster Mental Health*. Ellicott City, MD: Chevron Publishing Corporation. Used with permission.)

### The Primary Principles of Crisis Intervention

It is important that members of a CISM or aviation-based Critical Incident Response Program (CIRP) team adhere closely to a basic set of seven operational principles when providing crisis intervention services of any type.

*Proximity* is the first principle of crisis intervention. Crisis work is often provided in surroundings familiar to the people who need support. In the aviation world, many crisis intervention tactics are applied at airports or, in some cases, on board a parked aircraft or in an airport conference room.

The second principle is *immediacy*. People in a crisis appreciate help as soon as possible after being exposed to a traumatic experience. The longer they wait, the less likely crisis intervention will be effective (Lindy, 1985). The third principle is that of *expectancy*. Early in the intervention, the provider of crisis intervention should instill some hope that it is possible to manage and resolve the situation (Salmon, 1919; Kardiner and Spiegel, 1947; Solomon and Benbenishty, 1986). The fourth principle is *brevity*. The luxury of abundant time is usually absent in a crisis. Actions to assist people in crisis are generally brief (Parad & Parad, 1968).

The fifth principle is *simplicity*. People do not handle complexity very well in the midst of a crisis. Therefore, crisis support personnel should focus on selecting

solutions that are easy for a distressed person to apply. Simple, well thought out interventions will be the most effective in the majority of cases.

It is usually helpful if crisis workers have some creativity when they are working in extraordinary and challenging circumstances. The ability to be *innovative* in the face of unusual, threatening, and disturbing situations is a key to good crisis intervention and the sixth major principle of crisis intervention. Appropriate helpful interventions may have to be developed on the spot.

Finally, the seventh principle is that whatever crisis interventions are chosen in a crisis should be *practical*. Impractical solutions are no solutions at all. People struggling through a painful experience may view those who make impractical suggestions as insensitive and uncaring.

Peer support team members are encouraged to *keep the seven primary principles of crisis intervention in mind*. Crisis intervention is helpful when it is applied in accordance with those principles. Crisis support personnel should avoid establishing outcome expectations that are far beyond its capabilities. In other words, crisis intervention cannot cure disease or significant mental disturbance. It is not a cure for Post Traumatic Stress Disorder, a condition that is among the most serious reactions to a critical incident.

### **Aviation Staff Support Programs**

The management of the potential physical and psychological risks to aviation employees involves the establishment of well-organized and efficient crisis support programs. The aim of such programs should be to build a *high resiliency organization* that can resist overwhelming stress, bounce back from a traumatic event, and recover its personnel to healthy, functional levels. Fortunately, when it comes to support services for an organization's employees, aviation is not facing an uncharted course. The age of pioneering efforts in aviation crisis support is over. Crisis intervention in the aviation industries has a history that is now over two decades old and these programs have taken on great importance as they have consistently demonstrated their value. Experience clearly indicates the positive effects when crisis intervention services are applied in aviation's daily critical incidents as well as in major accidents (Vogt & Leonhardt, 2006).

This remainder of this article briefly reviews the history of crisis support services in aviation and then provides practical insights into the development of peer support programs, the training requirements for crisis support personnel, and the strategies and techniques of effective CISM programs. Additionally, by referring to studies such as the cost benefit analysis of a European air navigation service provider (Vogt, Leonhardt, Köper & Pennig, 2004), this article will outline the benefits of providing CISM services to aviation employees.

### **History of Critical Incident Stress Management in Aviation**

The history of crisis support services in aviation must be viewed in the context of crisis intervention in general. There is a rich history of crisis intervention that dates back to the wars and disasters of the past century and a half (APA, 1964; Artiss, 1963; Caplan, 1961, 1964; Crocq et al., 2007; Frederick, 1981; Freeman, 1979; Lindemann, 1944; Roberts, 2005; Salmon, 1919; Stierlin, 1909). It was not



until the Franco-Prussian war in 1870-71 that the approach to helping people in a state of crisis became somewhat organized and structured (Crocq, 1999). From that time forward, crisis intervention increasingly became a more well-thought-out *active, temporary, and supportive* process to assist people who are experiencing acute emotional distress (crisis) caused by a critical incident. "Crisis Intervention," "psychological first aid," "emotional first aid," and "early psychological intervention" are a few of the terms that evolved during the last century that are currently used synonymously to describe the supportive assistance provided by trained people to their colleagues and friends in times of acute distress (Neil, et al., 1974). Beginning in the mid-1970's a Comprehensive, Integrated, Systematic and Multi-component (CISM) crisis support program emerged (Mitchell and Everly, 2001; Mitchell, 2004; Mitchell & Mitchell, 2006). Today, Critical Incident Stress Management is the most well known and most widely utilized crisis intervention program in the world (Raphael & Wilson, 2000).

Modern theoretical foundations and core principles of contemporary crisis intervention theory were developed by Lindemann (1944) and elaborated upon by Gerald Caplan during the 1960s (Caplan, 1961, 1964, 1969). Caplan suggested that "Emotional first aid" or psychological first aid could be learned and applied by people who do not have formal training in psychology, social work, or psychiatry. He believed that *excellent help in a crisis could come from paraprofessionals or from friends helping friends*, family members reaching out to other family members and communities supporting their own members. Caplan laid the groundwork for the supportive crisis intervention services that are currently provided by members of the clergy, by emergency services, and by aviation personnel (Caplan, 1961, 1964, 1969).

Many aviation services including airlines, air traffic control systems, ancillary aviation services, and airports around the world have organized crisis support teams during the last two decades. Each airline, airport, air traffic control system, or ancillary service initiated their programs under vastly different circumstances and each has its own history of support services for its personnel.

### **Aviation Psychological Services**

Psychological services within the airline industry appear to have followed three general patterns and focused on three different issues within those patterns. The first pattern was the *screening of personnel*. Psychological testing of potential service members was important to assure the selection of the best people. Improper screening could contribute to accidents and, potentially, the loss of lives (McCarthy, 2002). By the 1960s, the National Aeronautical and Space Administration (NASA) developed sophisticated psychological screening methods for the assessment of potential astronauts (Miller, 2002). The field of aviation psychology mirrored NASA's approach. Psychological screening became an integral part of the medical screening of most potential commercial pilots.

The second pattern in aviation psychology, and a much resented one, was the pattern of *on-the-job evaluations*. This pattern included the use of aviation psychological services to protect the airline and the public and, although not stated, to enforce discipline and control. If a pilot's behavior was challenging to administration or if supervisors perceived it to be unusual or even remotely potentially dangerous, a pilot received a psychological evaluation. The results of the evaluation were

sometimes used by aviation administrators to bar a pilot from flying or to fire the person altogether. The process is seen as less threatening today in comparison to its earlier development.

The third pattern of aviation psychology, *psychological assistance in difficult times*, emerged in the mid 1980s and was positively viewed by aviation personnel. Crew resource-management programs, for example, improved cockpit communications and decreased safety risks. Other psychological support services included alcohol and drug rehabilitation programs that allow pilots to fly again once they successfully complete treatment. Some airlines also developed family support programs for their employees (O'Flaherty, 1995).

Prior to 1980, mental health professionals outside of the aviation industry predominantly expressed most of their interest in aviation by a rather myopic focus on the needs of disaster survivors or grieving loved ones of the deceased (Duffy, 1979; Freeman, 1979, Frederick, 1981). This particular focus allowed only limited resources to be applied to the establishment of support programs for employees. Despite this lack of attention to the needs of employees, by the end of the 1980's, most air carriers did have Employee Assistance Programs to help employees with issues requiring general counseling and substance abuse intervention, but peer-based crisis intervention services were still virtually unknown. One notable exception to the description of the mediocre atmosphere surrounding employee crisis support services was the attention that was beginning to be paid to aircraft crash, rescue, and firefighting crews and to law enforcement personnel working at air crash scenes (Forstenzer, 1980, Mitchell, 1982).

### **Airline Training in Critical Incident Stress Management**

Alaska Airlines and Wien Air Alaska were the first airlines to send key pilots and human resource personnel for training in Crisis Intervention and stress management at a conference in Anchorage, Alaska in 1982. American Airlines arranged for similar training for twenty-five pilots and other selected employees in 1987. The American Airlines program was the first formal, aviation specific, crisis and stress management program on record (LiBassi, 1987). Other airlines and military aviation units quickly followed suit and developed Critical Incident Stress Management teams to support their personnel. The very first military aviation focused CISM team was established by the United States Air Force at Rhine-Main US Air Force Base near Frankfurt, Germany in 1988 (Mitchell, 2006).

The Air Line Pilots Association (ALPA) became involved in developing crisis support programs for its members in the early 1990s. ALPA strongly supported the development of Critical Incident Response Programs (CIRP) within airlines. By May of 1994, ALPA's Executive Board unanimously passed a resolution formally recognizing and funding the program (Tompkins et al., 1996). Today, most major airlines around the globe, including many cargo carriers have Critical Incident Stress Management teams which they have renamed "Critical Incident Response Programs." Positive reports from aviation industry personnel have been consistent over the nearly twenty-year history of these crisis programs. (Dillenbeck, 1996; Bailey & Hightower, 1996).

ALPA's endorsement of CISM programs was a major step forward in the development of standardized comprehensive, integrated, systematic, and multi-

component crisis intervention and stress management programs for air carriers around the world. ALPA leadership contributed enormously to the expansion of crisis intervention programs to all employees within the airline industry. By keeping The National Transportation Safety Board (NTSB) and the Federal Aviation Administration informed of the uses and positive effects of CISM services, ALPA also influenced a more positive atmosphere toward stress management services for stressed pilots and other employees in the view of the accident investigators and the flight regulations personnel (ALPA, 1997).

Flight attendant associations and unions as well as those of mechanics, baggage handlers, and other aviation workers have also endorsed the CISM program. In a somewhat unusual program, Frankfurt International Airport and DFS (Deutsche Flugsicherung<sup>1</sup>), the German air traffic control system, work together in a mutual aid arrangement. They share CISM training opportunities, meet together regularly, and mutually support each other. In the unfortunate event of a major emergency both the Frankfurt Airport and DFS will share resources to provide CISM services to employees and to the traveling public.

### **Air Traffic Controller Systems Initiate CISM Programs**

Tragedy was the backdrop for the development of Critical Incident Stress Management programs for air traffic controllers. Transport Canada initiated a well-received air traffic controller CISM program in 1988. The program provided stress education and individual crisis intervention services to distressed air traffic controllers. In 1989, United Airlines experienced the Sioux City, Iowa crash. Canadian air traffic controllers mobilized and crossed the border into the United States to assist their American colleagues by means of CISM support services (Dooling, 2003).

The Sioux City disaster was among several factors that stimulated the development of the US Federal Aviation Administration's air traffic control CISM team. The CISM team is unique and serves as a model program in that the National Air Traffic Controllers Association, the air traffic controllers union, and the Federal Aviation Administration jointly sponsors it. In fact, both management and union jointly developed a 20-minute videotaped program on Critical Incident Stress Management for air traffic controllers (NATCA / FAA, 2003).

Of particular note is the endorsement of CISM services by the International Federation of Air Traffic Controllers' Associations (IFATCA). Standardized CISM services became a priority for this organization's constituents during the last decade. IFATCA has been a driving force in the establishment of CISM programs in many air navigation services throughout Europe and elsewhere. Most air navigation services credit IFATCA with convincing their management that CISM is a benefit for air traffic controllers.

Eurocontrol, the air traffic control organization for European nations, has for a decade, sponsored several important studies and conferences on Critical Incident Stress Management. It encourages its member nations to develop support programs for air traffic control personnel. A prominent example of one system that did so is the Deutsche Flugsicherung (DFS), the German air navigation system. Its CISM program was developed in 1998. The team has responded to many smaller

---

<sup>1</sup> DFS is the German Air Navigation Service Provider

events, but also experienced a trial by fire in two major incidents. One was an airplane crash into terrain and the other was the collision of two aircraft over Lake Constance near Ueberlingen in Switzerland in 2001 (Leonhardt, 2006).

### **The Peer Support Concept**

Many of the crisis intervention tools within the CISM field are applied by trained peer support personnel. The use of peer support personnel is one feature that makes CISM programs unique. Peers enjoy immediate credibility among their fellow workers. They share professional understanding of the work processes that exist within specialized groups such as, pilots, air traffic controllers, flight attendants, aircraft crash and rescue firefighters, and law enforcement personnel. Peer support personnel understand the post incident stress reactions of their colleagues. They can support personnel in the same profession and they can normalize their reactions to the event. Peers are readily available and have relevant insights into the organization and the profession (Leonhardt, 2006; Mitchell, 1983).

To enable peers to provide the best support services, there must be clear procedures and guidelines. The various structured crisis intervention tools within CISM, therefore, serve useful purposes for peer support personnel who do not have university degrees and certifications in mental health. Peers are highly motivated to assist distressed colleagues and they are very serious about their crisis work. Specially trained peer support personnel have proven to be extremely effective in delivering excellent support service to their colleagues (Hassling, 2000; Jenkins, 1996; North et al., 2002; Nurmi, 1999; Robinson & Mitchell, 1993; Vogt et al., 2004; Vogt, Pennig & Leonhardt, 2007; Western Management Consultants, 1996).

### **Training and Standards for Aviation Peer Support Personnel**

A typical CISM training program for peer support personnel is multi-layered and provides adequate opportunities for peers to practice the skills presented in the various levels. The primary training for peer support personnel is a course detailing the skills for supporting individuals in crisis. This course not only provides the information necessary to assist individuals in a crisis, it also encompasses sufficient practice sessions to build practical skills. Peer support personnel become more skillful with additional training. A ten-day training program can be presented in two-day modules over a one to two year time frame. The modules in a CISM peer support training program include, but are not limited to,

- Skills for supporting individuals in crisis
- Group crisis support skills
- Skills for suicide prevention, intervention, and recovery
- Advanced group processes and procedures
- Strategic management of crises

Aviation peer support personnel should be encouraged to participate in a broad spectrum of crisis and stress management programs. Additional crisis support responsibilities are added as individuals become knowledgeable and skillful regarding CISM.

### **Practical Guidelines for Effective Aviation Peer Support Programs**

1. Every successful crisis-oriented aviation staff support program is *comprehensive*. That is, it has elements in place before, during, and after traumatic events. Additionally, it is *programmatically*. That is, it is endorsed by the administration and built into the fabric of the organization. Administration and union “buy-in” or acceptance of a support program is essential to the program’s survival. Although they remain independent units, peer support programs must communicate, coordinate, and link their efforts with the organization’s leaders and with human services, employee assistance, and psychological resources within an organization.
2. A key to effective peer support programs is the presence of dedicated enthusiastic leaders who actively work toward enhancing the peer support team at every opportunity.
3. Well-developed aviation peer support programs are *integrated*. That is, all of the elements of a program are interrelated and blended with one another. The combined effects of an integrated program are far more powerful than any single element.
4. Aviation personnel are best sustained by a *systematic* program or “support package,” which has phases, segments, or logical steps. Therefore, peer support programs should take a few simple steps such as, resting personnel and talking with them on an individual basis, before increasing the complexity, number, and duration of the available crisis interventions after a distressing event.
5. Effective aviation peer support programs must be *multi-tactic* in approach. Many different types of support services must be available since every person will have a somewhat different response to a highly stressful event. Each person will have different needs after a traumatizing experience.
6. Although it is part of a Comprehensive, Integrated, Systematic, and Multi-tactic (CISM) approach, *linkages to a wide range of resources* is an important enough element of an aviation crisis support program. Additional assistance, beyond the capabilities of the peer support team, may be necessary.

7. Crisis support teams are most effective when they are run and staffed by *peer support personnel* and backed up by mental health professionals who are trained in Critical Incident Stress Management (CISM) (Mitchell, 2004).

### **Effectiveness of CISM**

The literature to date suggests that crisis intervention, in the form of a Critical Incident Stress Management program, has positive effects on the reduction of stress symptoms. When applied by properly trained personnel, who adhere to the standards of practice, CISM programs may play a preventative role against the development of long-range psychological problems (Bohl, 1991, 1995; Boscarino, Adams, & Figley, 2005; Campfield and Hills, 2001; Chemtob, Tomas, Law, & Cremniter, 1997; Deahl et al., 2000; Dyregrov, 1998; Eid, Johnsen, & Weisaeth, 2001; Everly, Flannery & Mitchell, 2000; Everly, Flannery, Eyer, & Mitchell, 2001; Flannery, 2005; Jenkins, 1996; Nurmi, 1999; Richards, 2001; Robinson and Mitchell, 1993; Vogt et al., 2004; Vogt & Leonhardt, 2006; Vogt et al., 2007; Wee, Mills, & Koelher, 1999; Western Management Consultants, 1996).

In one particularly interesting series of studies directly related to the aviation industry, Vogt and Leonhardt (2006) and Vogt et al. (2004, 2007) performed a cost-benefit analysis of the German Air Traffic Control Services' (Deutsche Flugsicherung, DFS) CISM program. The authors describe the applications of the CISM program within air-traffic control systems after "loss of separation" incidents (when two radar blips merge into one on the screen). An unforeseen loss of separation between aircraft is a major threat to the professional self-image of an air traffic controller officer. Such incidents are generally distressing reminders to the controllers of the ever-present risk of life-threatening aviation accidents. Prior to the development of a CISM support program, German Air Traffic Control Officers lost an average of three days from work per incident because of distress. About 30 of these lost of separation events occur per month in the crowded skies above Europe (Vogt & Leonhardt, 2006; and Vogt et al., 2004, 2007).

An analysis by Vogt and Leonhardt (2006) and Vogt et al. (2004, 2007) indicates that, since the introduction of the CISM program in 1997, no air traffic controller reported a single lost day for stress related to a loss of separation critical incident. The estimated fiscal benefits associated with the prevention of absenteeism or reported stress-related illness while at work actually exceeded the program's costs several times. The study concludes that a combination of factors assures the effectiveness of peer provided crisis intervention. They include a clear model, well-trained providers, and adherence to the standards and protocols of good practice (Vogt & Leonhardt, 2006; and Vogt et al., 2004, 2007).

### **Conclusion**

This article began with a discussion of high-risk organizations in which the employees are subjected to high levels of stress and circumstances that threaten their physical and mental health. It then noted that aviation's high-risk organizations are also High Reliability Organizations (HRO). Consistency in performance, efficiency, and a safety consciousness are the hallmarks of HROs. The authors review evidence that when well-designed Critical Incident Stress Management programs are in place and properly trained aviation peer support personnel adhere to the stan-

dards of CISM practice, that a new meaning for “HRO” may be applicable. That is, aviation personnel will function effectively within *high resiliency organizations* that can resist stress, manage it when it arises, and restore personnel to maximal levels of performance and unit cohesion.

From the airport facility to the airlines and from the ground staff to the pilots, air traffic controllers, and flight attendants, the studies and reports from within the aviation industry suggest that excellent help in a crisis could come from para-professionals or from friends helping friends. Whether it is a situation involving personal grief, a family crisis, or an air disaster, aviation CISM team members respond with skill, care, and concern. Crisis support personnel have demonstrated the enormous value of crisis intervention teams. It is likely that the number of support programs will expand in the coming years. Peer support personnel make a difference in the lives of others.



## REFERENCES

- Air Line Pilots Association (ALPA). (1997). *Critical Incident Response Program Guide*. Herndon, VA: ALPA.
- American Psychiatric Association (APA). (1964). *First Aid for Psychological Reactions in Disasters*. Washington, DC: American Psychiatric Association.
- Artiss, K. (1963). Human behavior under stress: From combat to social psychiatry. *Military Medicine*, 128, 1011-1015.
- Bailey, W. & Hightower, S. (1996). *Critical Incident Response Program*. Dallas, Texas: Southwest Airlines Pilots' Association, Transport Workers Union.
- Bohl, N. (1991). The effectiveness of brief psychological interventions in police officers after critical incidents. In J.T. Reese and J. Horn, and C. Dunning (Eds.) *Critical Incidents in Policing, Revised* (pp.31-38). Washington, DC: Department of Justice.
- Bohl, N. (1995). Measuring the effectiveness of CISD. *Fire Engineering*, 125-126.
- Boscarino, J.A., Adams, R.E. & Figley, C.R. (2005). A Prospective Cohort Study of the Effectiveness of Employer-Sponsored Crisis Interventions after a Major Disaster. *International Journal of Emergency Mental Health*, 7(1), 31-44.
- Campfield, K. & Hills, A. (2001). Effect of timing of critical Incident Stress Debriefing (CISD) on posttraumatic symptoms. *Journal of Traumatic Stress*, 14, 327-340.
- Caplan, G. (1961). *An approach to Community Mental Health*. New York: Grune and Stratton.
- Caplan, G. (1964). *Principles of Preventive Psychiatry*. New York: Basic Books.
- Caplan, G. (1969). Opportunities for school psychologists in the primary prevention of mental health disorders in children, In A. Bindman and A. Spiegel (Eds.) *Perspectives in Community Mental Health* (pp.420-436). Chicago: Aldine.
- Chemtob, C., Tomas, S., Law, W., and Cremniter, D. (1997). Post disaster psychosocial intervention. *American Journal of Psychiatry*, 134, 415-417.
- Crocq, L. (1999). *Les Traumatismes psychiques de guerre*. Paris, France ; Odile Jacob.
- Crocq, L., Dalligand, L., Villerbu, L., Tarquinio, C., Duchet, C., Coq, J.M., Chidiac, N., Vitry, M. (2007). *Traumatismes psychiques : Prise en charge psychologique des victims*. Paris, France: Elsevier-Masson.
- Deahl, M., Srinivasan, M., Jones, N., Thomas, J., Neblett, C., & Jolly, A. (2000). Preventing psychological trauma in soldiers. The role of operational stress training and psychological debriefing. *British Journal of Medical Psychology*, 73, 77-85.
- Dillenbeck, P. (1996). *Critical Incident Response Program for FedEx Flight Crewmembers*. Memphis, TN: Federal Express Corporation.

- Doolling, M. (2003). Critical Incident Stress Management (A videotaped presentation at Eurocontrol's conference on "Training of controllers in Unusual / Emergency Situations, 1996). Luxembourg: EUROCONTROL, Institute of Air Navigation Services.
- Duffy, J. (1979) The role of CMHCs in airport disasters. Technical Assistance Center Report, 2(1), 1,7-9.
- Dyregrov, A. (1998). Psychological debriefing: An effective method? *TRAUMATOLOGY*, 4, (2), Article 1.
- Eid, J., Johnsen, B. H., & Weisaeth, L. (2001). The effects of group psychological debriefing on acute stress reactions following a traffic accident: a quasi-experimental approach. *International Journal of Emergency Mental Health*, 3, 145-154.
- Everly, G.S., Jr., Flannery, R.B., and Mitchell, J.T. (2000). Critical Incident Stress Management (CISM): A review of the literature. *Aggression and Violent Behavior*, 5(1), 23-40.
- Everly, G.S., Jr., Flannery, R. B., Jr., Eyer, V. & Mitchell, J.T. (2001). Sufficiency analysis of an integrated multi-component approach to crisis intervention: Critical Incident Stress Management. *Advances in Mind-Body Medicine*, 17, 174-183.
- Flannery, R.B. (2005). *Assaulted Staff Action Program (ASAP): Fifteen years of Empirical Findings*. Paper presented at the 8th World Congress on Stress Trauma and Coping: Crisis Intervention: Best Practices in Prevention, Preparedness and Response. Baltimore, Maryland, USA February 16-20.
- Forstenzer, A. (1980). Stress, the psychological scarring of air crash rescue personnel. *Firehouse*, July, 50-52; 62.
- Frederick, C. J. (Ed.) (1981). *Aircraft Accidents: Emergency Mental Health Problems*. Washington, DC: National Institute of Mental Health, U.S. Department of Health and Human Services.
- Freeman, K. (1979). CMHC responses to the Chicago and San Diego airplane disasters. *Technical Assistance Center Report*, 2(1), 10-12.
- Hassling, P. (2000). Disaster management and the Gothenburg fire of 1998: when first responders are blamed. *International Journal of Emergency Mental Health*, 2(2) 267-273.
- Inglish, P. (2010). *Top 10 Most Dangerous Jobs 2007-2010*. Hub Pages Elite. Retrieved from [http://Hubpages.com/hub/Most\\_Dangerous\\_Jobs](http://Hubpages.com/hub/Most_Dangerous_Jobs).
- Jenkins, S.R. (1996). Social support and debriefing efficacy among emergency medical workers after a mass shooting incident. *Journal of Social Behavior and Personality* 11, 447-492.
- Kardiner A. and Spiegel, H. (1947). *War, Stress, and Neurotic Illness*. New York: Hoeber.
- Leonhardt, J. (2006). Critical Incident Stress Management (CISM) in Air Traffic Control (ATC). In J. Leonhardt and J. Vogt (Eds.) *Critical Incident Stress Management CISM in Aviation* (pp 81-91). Aldershot, UK: Ashgate Publishing Company.

- LiBassi, M.D. (1987). Employee Trauma Recovery Program. *Personnel Resources, Local Procedures Manual*. Dallas, Texas: American Airlines.
- Lindemann, E. (1944). Symptomatology and management of acute grief. *American Journal of Psychiatry*, 101, 141-148.
- Lindy, J. (1985). The trauma membrane and other clinical concepts derived from psychotherapeutic work with survivors of natural disasters. *Psychiatric Annals*, 15, 153-160.
- McCarthy, P. (2002). *Dr. Patrick McCarthy's brief Outline of the History of I/O Psychology*. Retrieved from [http://www.mtsu.edu/~pmcarth/io\\_hist.htm](http://www.mtsu.edu/~pmcarth/io_hist.htm).
- Miller, K. (2002). *How Astronauts Get Along*. Houston, Texas: Retrieved from [http://science.nasa.gov/science-news/science-at-nasa/2002/21oct\\_getalong/](http://science.nasa.gov/science-news/science-at-nasa/2002/21oct_getalong/).
- Mitchell, J.T. (1982). The psychological impact of the Air Florida 90 Disaster on fire-rescue, paramedic and police officer personnel. In R Adams Cowley, Sol Edelstein and Martin Silverstein (Eds.) *Mass Casualties: a lessons learned approach, accidents, civil disorders, natural disaster, terrorism*. Washington, DC: Department of Transportation (DOT HS 806302), October, 1982.
- Mitchell, J.T. (1983) When disaster strikes ... The critical incident stress debriefing process. *Journal of Emergency Medical Services*, 13(11), 49 – 52.
- Mitchell, J.T. (2004). Characteristics of Successful Early Intervention Programs. *International Journal of Emergency Mental Health*, 6(4), 175-184.
- Mitchell, J.T. and Everly, G.S., Jr., (2001). *Critical Incident Stress Debriefing: An operations manual for CISD, Defusing and other group crisis intervention services*, Third Edition. Ellicott City, MD: Chevron.
- Mitchell, S.G. and Mitchell, J.T. (2006). Caplan, Community and Critical Incident Stress Management. *International Journal of Emergency Mental Health*, 8(1), 1-10.
- Mitchell, J.T. (2006). A brief history of Crisis Intervention in Aviation. In Vogt, J. & Leonhardt, J. (Eds.). *Critical Incident Stress Management in Aviation*. Hampshire, UK: Ashgate Publishing Limited.
- NATCA / FAA (2003). *You are Not Alone: Critical Incident Stress Management Program*. Washington, DC: NATCA / FAA.
- Neil, T., Oney, J., DiFonso, L., Thacker, B. and Reichart, W. (1974). *Emotional First Aid*. Louisville: Kemper-Behavioral Science Associates.
- North, C., Tivis, C., McMillen, J., Pfefferbaum, B., Cox, J., Spitznagel, E., Bunch, K., Schorr, J., and Smith, E. (2002). Coping, functioning, and adjustment of Rescue workers after the Oklahoma City bombing. *Journal of Traumatic Stress*, 15, 171-175.
- Nurmi, L. (1999). The sinking of the Estonia: The effects of Critical Incident Stress Debriefing on Rescuers. *International Journal of Emergency Mental Health*, 1, 23-32.
- O'Flaherty, J. (1995). *Trans World Airlines: Trauma Response Team, Section Leader Manual*. St Louis, MO: Trans World Airlines, Inc.
- Parad, L. and Parad, H. (1968). A study of crisis oriented planned short –term treatment: Part II. *Social Casework*, 49, 418-426.

- Parker, J.G., (2006). "Watching the Clock", *Safety and Health*, April 2006, pp.44-47.
- Raphael, B. and Wilson, J. (2000). Introduction and Overview: Key issues in the conceptualization of debriefing. In Beverly Raphael and John Wilson (Eds.) *Psychological Debriefing*. Cambridge, UK: Cambridge University Press.
- Richards, D. (2001). A field study of critical incident stress debriefing versus critical incident stress management. *Journal of Mental Health*, 10, 351-362.
- Roberts, A. (2005). Bridging the Past and Present to the Future of Crisis Intervention and Crisis Management. In Allen Roberts (Ed.) *Crisis Intervention Handbook: Assessment, Treatment and Research*. Third Edition. New York: Oxford University Press.
- Robinson, R. and Mitchell, J. (1993). Evaluation of psychological debriefings. *Journal of Traumatic Stress*, 6, 367-382.
- Salmon, T.S. (1919). *War neuroses and their lesson*. New York Medical Journal, 108, 993-994.
- Slaikeu, K.A. (1984). *Crisis intervention: A handbook for practice and research*. Boston, MA: Allyn and Bacon, Inc.
- Solomon, Z. and Benbenishty, R. (1986). The role of proximity, immediacy, and expectancy in frontline treatment of combat stress reaction among Israelis in the Lebanon War. *American Journal of Psychiatry*, 143, 613-617.
- Staff. (2001). Aviation Workplace Safety Lags Other Industries. *Air Safety Week*, May 21.
- Stierlin, E. (1909). *Psycho-neuropathology as a result of a Mining Disaster* March 10, 1906. Zurich: University of Zurich.
- Tompkins, M., Morin, W., Campbell, A., Hudson, D., Sumwalt, R., Johnson, J., Anding, J.D. & Bracken, J. (1996). *Air Line Pilots Association: Critical Incident Response Program Manual*. Herndon, VA: ALPA.
- United States Department of Labor, Bureau of Labor Statistics. (2008a). *Fatal Occupational Injuries Resulting from Transportation incidents and Homicides, All United States*. Washington, DC: US Department of Labor.
- United States Department of Labor, Bureau of Labor Statistics. (2008b). *Fatal Occupational injuries by event or exposure, 2007-2008*. Washington, DC: US Department of Labor.
- United States Department of Labor, Bureau of Labor Statistics. (2009). National Census of Fatal Occupational Injuries in 2008. *Bureau of Labor Statistics News*, August 20, 2009.
- United Nations Department of Safety and Security CISMU Staff (2007a). UNDSS CISMU Certification Training For Counselors. *DSS Newsletter*. New York: United Nations Secretariat, Department of Safety and Security.
- United Nations Department of Safety and Security CISMU Staff (2007b). New crisis and stress management training programme launched. *I Seek* (May 2, 2007) New York: United Nations Secretariat.

- United Nations Department of Safety and Security CISMU Staff (2007c). *Certification training in Crisis and Stress Management*. New York: UN Department of Safety and Security, Consultative Working Group on Stress in Collaboration with the International Critical Incident Stress Foundation, the American Academy of Experts in Traumatic Stress and the Comité National de L'urgence Medico-Psychologique. New York: United Nations, Department of Safety and Security.
- Vogt, J. & Leonhardt, J. (Eds.). (2006). *Critical Incident Stress Management in Aviation*. Hampshire, UK: Ashgate Publishing Limited.
- Vogt, J., Leonhardt, J., Koper, B. & Pennig, S. (2004). Economic Evaluation of CISM – A Pilot Study. *International Journal of Emergency Mental Health*, 6(4), 185-196.
- Vogt, J., Pennig, S. & Leonhardt, J. (2007). Critical Incident Stress Management in air traffic control and its benefits. *Air Traffic Control Quarterly*, 15(2), 127-156.
- Wee, D.F., Mills, D.M. & Koelher, G. (1999). The effects of Critical Incident Stress Debriefing on emergency medical services personnel following the Los Angeles civil disturbance. *International Journal of Emergency Mental Health*, 1, 33-38.
- Western Management Consultants. (1996). *The Medical Services Branch CISM Evaluation Report*. Edmonton Alberta: WMC.

## ***Changing General Aviation Flight Training by Implementing FAA Industry Training Standards***

Michele Summers Halleran

and

Michael E. Wiggins

*Embry-Riddle Aeronautical University  
Daytona Beach, FL*

### ***Abstract***

*In an effort to reduce the accident rate in General Aviation, the Federal Aviation Administration (FAA) has developed a new training program called FAA/Industry Training Standards (FITS). The FITS program began in 2001 and represented the aviation industry, academia, and the FAA. FITS philosophy is based on the idea that most general aviation accidents can be attributed to deficiencies in three key areas; aeronautical decision-making, risk management, and situational awareness. The FITS program is leading an effort to change current flight training processes to embrace technology and target the pilot's decision-making skills with each training curriculum. It is centered on three key elements: scenario-based training, single pilot resource management, and learner-centered grading. FITS emphasizes the ultimate goal of increasing safety on all levels of performance and teaches the pilot much earlier in training to develop decision making skills. This paper examines the development, components, application, and outcomes of the FITS program.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

# Changing General Aviation Flight Training by Implementing FAA Industry Training Standards

## Background

The end of the 20<sup>th</sup> Century saw the genesis of a new era in general aviation. A new generation of small aircraft with computerized cockpits, complete with new primary flight displays, multifunction displays, autoflight systems, and digital datalink technologies to take full advantage of global positioning system (GPS) navigation capabilities was just around the corner. Very light jets were looming on the horizon promising private pilots the opportunity to fly at altitudes and speeds only known to commercial and corporate aviation at the time. Capabilities and technologies found only in commercial airliners and sophisticated corporate aircraft would soon find their way into general aviation. It was clear to some in the leadership of the Federal Aviation Administration's (FAA) General Aviation and Commercial Division that changes to pilot training would be necessary to meet the challenges these new aircraft and avionics would bring (Wright, 2002).

The change from "round-dial" cockpits to new "glass" cockpits has brought a new set of challenges and skills needed to properly manage a flight in technically advanced aircraft (TAA). This is the same challenge the airline industry faced in the 1980s and 1990s when "glass" cockpits and flight management systems were introduced into commercial aviation. In addition to traditional skills of flying, navigating, and communicating, pilots in the newer aircraft now have to manage automation, information displays, and other new technologies (FAA, 2009a). As airspace complexity and air traffic density increase, pilots must be situationally aware at all times, understand risk assessment, and have a complete understanding of automation management. Traditional maneuvers-based flight training and testing does not focus on the proper development of these skills, a fact noted by the airline industry in the late 20th Century, which caused the airlines to adopt their current training methods (FAA, 2009a).

In an effort to reduce the accident rate in general aviation and meet the challenges brought forth by these TAA, the FAA has developed a new style of training program called FAA/Industry Training Standards (FITS). The FITS program began in late 2001 by developing the FITS team; a group representing the aviation industry, academia, and the FAA. Just to illustrate that how serious the FAA was about this initiative, the FAA awarded a grant worth \$900,000 USD in June 2002 (total grant as of 9/1/09 \$2.3 million USD) to The Center of Excellence for General Aviation Research (CGAR). The CGAR is a consortium of universities to develop pilot training methods to change general aviation flight training in the United States. In total, over 16 industry partners, including manufacturers, insurers, training providers, and trade associations along with six academic institutions, including the CGAR, collaborated with the FAA to create FITS.

One key aspect of FITS was that it was not to be regulatory but voluntary in nature. No regulations were changed or mandated in the creation of FITS. A training program will not be "approved" by the FAA as a program submitted under 14 CFR part 141 or 14 CFR part 142 but will be "accepted" as meeting FITS requirements. Guidance documents have been given to both industry and to FAA aviation safety inspectors regarding seeking and gaining FITS acceptance. A program



seeking FITS acceptance will have to meet existing regulatory requirements but poses no additionally mandated burden on training providers.

## **FAA/Industry Training Standards Training Philosophy**

A study of accidents involving TAAs conducted in 2003 (FAA) found that in some cases the increased complexity of the new technologies can create distractions for pilots, that pilots need specific and focused training to better utilize the capabilities of TAAs, and that current training requirements and methods do not provide that needed training. This study made several recommendations that help frame the tenets of FITS. These recommendations included: 1) provide realistic training in realistic scenarios that focus on the “physical” airplane and the “mental” airplane, 2) assessing and managing risks, 3) understanding limitations of TAA systems and system management.

Previous training philosophies assumed that newly certified pilots generally remain in the local area until their aviation skills are refined. This is no longer true with the advent of technically advanced aircraft. Offering superior avionics and performance capabilities, many of these new aircraft travel faster and further than their predecessors. As a result, a growing number of entry-level pilots are suddenly capable of long distance/high speed travel and its inherent challenges. Flights of this nature routinely span diverse weather systems and topography requiring advanced flight planning and operational skills. Advanced cockpits and avionics, while generally considered enhancements, require increased technical knowledge of newer systems and avionics and new skills managing automation and computerized navigation systems. Without these skills, the potential for an increased number of pilot-induced accidents is daunting. A different method of training is required to accelerate the acquisition of these skills during the training process.

Research has proven that learning is enhanced when training is realistic. In addition, the underlying skills needed to make good judgments and decisions are teachable (Schuetz, 2003, Funk, 1998). Both the military and commercial airlines have embraced these principles through the integration of Line Oriented Flight Training (LOFT) and Crew Resource Management (CRM) training into their qualification programs. Both LOFT and CRM lessons mimic real-life scenarios as a means to expose pilots to realistic operations and critical decision-making opportunities. The most significant shift in these programs has been the movement from traditional maneuver-based training to incorporate scenario-based training.

FITS incorporates CRM, which we call Single-pilot Resource Management (SRM), and LOFT-type scenarios, called Scenario-Based Training (SBT) from the very first training flight. The FITS programs incorporate SRM and SBT in training programs for all levels of pilots, from pre-solo through CFI. The difference between airline CRM training and FITS SRM training is that SRM is taught throughout all levels and is not just taught to advanced pilots.

General aviation pilots are often the only pilot on board when flying. However, the number of resources they have to manage in TAAs can rival those of the most sophisticated airliner. Taking a lesson from the airline industry’s approach to CRM training, FITS is focusing on single-pilot resource management (SRM). SRM is defined as the art and science of managing all the resources (both on-board

the aircraft and from outside sources) available to a single-pilot (prior and during flight) to ensure that the successful outcome of the flight is never in doubt (FAA, 2009a).

Maneuver-based training emphasizes the mastery of individual tasks or elements. Regulations, as well as Practical Test Standards (PTS), drive completion standards. Flight hours and the ability to fly within specified tolerances determine competence. The emphasis is on development of motor skills to satisfactorily accomplish individual maneuvers. Only limited emphasis is placed on decision-making. As a result, when the maneuver-based trained pilot flies in the real world environment, he or she has had limited training in real world flying and decision-making. In the traditional private, commercial, and instrument training programs, students are required to fly a very limited number of cross-country flights.

SBT and SRM are similar to LOFT and CRM training. However, each is tailored to the pilot's training needs. These techniques use the same individual tasks that are found in maneuver-based training, but script them into scenarios that mimic real life cross-country travel. Guidance regarding the development of scenarios is available on the FITS website along with many sample syllabi and lesson plans.

One key concept that was blended into the development of the FITS model was the idea of developing the pilot's self-assessment skills thereby promoting life-long learning skills. The method chosen to accomplish this is called learner-centered grading (LCG). Taking note of trends in traditional classroom education and other areas of skill-based training, LCG is a method by which both the student and instructor independently grade the lesson, then compare, and discuss each assessment (FAA, 2009a). The idea is to train students so that they can accurately assess their own performance on each flight and continually learn after formal training has long ceased.

The concepts of SBT, SRM, and LCG are the foundation for the FITS approach to training pilots. By emphasizing the goal of flying safely, the pilot in training correlates the importance of individual training maneuvers to safe mission accomplishment. The concepts for FITS also include important findings from research in the FAA's Alaska Capstone Program and the Safer Skies program, and NASA's Advanced General Aviation Transport Experiment (AGATE) and Small Aircraft Transport Systems (SATS) programs (Wright, 2002).

## **Teaching Methods**

### **Scenario-Based Training**

For SBT to be effective there must be a purpose for the flight and consequences if it is not completed as planned (FAA, 2009a). Again, each training flight should have a realistic purpose. There should be realistic consequences if it is not completed as planned. The student and the instructor discuss the following information well in advance of every training flight: purpose of flight, scenario destination(s), desired pilot in training learning outcomes, desired level of pilot in training performance, desired level of automation assistance, and possible in-flight scenario changes (Summers, Ayers, Connolly, & Robertson, 2007).

## Developing Scenario-Based Training

Prior to the flight, a Certified Flight Instructor (CFI) will brief the scenario to be planned by the pilot in training (PT). The CFI will review the plan and offer guidance on how to make the lesson more effective relative to the goals and objectives for the lesson. Discussion, in part, will reflect ways in which the CFI can most effectively draw out a pilot in training's knowledge and decision processes. This enables the CFI to analyze and evaluate the PTs level of understanding. With the guidance of the flight instructor, the pilot in training should make the flight scenario as realistic as possible. This means the pilot in training will know where they are going and what will transpire during the flight. While the actual flight may deviate from the original plan, it allows the pilot in training to be placed in a realistic scenario. After discussion with the instructor, the pilot in training will plan the flight to include: reason to go flying, route, destination(s), weather, Notices to Airmen (Notams), desired pilot in training learning outcomes, and possible alternate scenarios and/or emergency procedures.

A typical training flight for a private pilot student might consist of doing a soft-field takeoff and departure, flying out to the practice area where a few stall maneuvers are practiced. After the stalls are completed, the student is then directed to descend to a lower altitude to practice a few ground reference maneuvers followed by a simulated engine failure. Finally, the flight returns to the airport so the student can practice a few short and soft-field takeoffs and landings. The flight is planned, guided, and sequenced by the instructor. This is not what the pilot who is training can expect when they are certified and will be flying on their own. The same can be said for instrument training. Not many real Instrument Flight Rules (IFR) flights depart and head directly to a holding pattern somewhere and intentionally execute instrument approaches with the intent of missing the approach and not landing. The regulations require that only one actual IFR cross-country flight be conducted for certification as an instrument pilot (14 CFR Part 61.65.). For a private pilot, only 3 hours of cross country training with an instructor are required plus one night dual cross country (14 CFR 61.109) (FAA, 2010).

The problem with the aforementioned maneuvers-based training approach is that realistic decisions are not utilized nor are even necessary. When the student pilot is "turned loose" after training, a new and different set of mental skills are needed, yet are often not practiced or even fully understood by the pilot. To be good at making sound decisions and practicing good resource management, pilots in training need to use those skills each and every flight in as realistic setting as possible. Hence, SBT is utilized in FITS curricula to develop and practice these skills.

Students should be involved in the development and planning of the scenarios as much as possible, as if they were planning a real flight. The instructor moves from a role of directing the flight to one of assisting and guiding the student through the scenario. The student should make all of the decisions pertaining to the flight, including continuing or diverting to the fullest extent possible. Yes, a student on their first few flights may not be able to make all of the decisions, but they should be making as many as they possibly can. This is one reason the method of debriefing is used in learner-centered grading. The instructor and student can discuss the student's decisions and processes and the tools used to make those decisions.

## Example of Scenario Based Training

A scenario provides a purpose of the flight and suggests a consequence if the flight does not go. It is designed to put the flight into a realistic context and provide a basis for realistic decision-making. A well-crafted scenario will contain many opportunities for the student to exercise good SRM skills as well as flying and practicing maneuvers. The following is an example of a scenario from the generic FITS private-instrument syllabus, lesson 8. This is a pre-solo lesson.

Scenario: A friend of yours is a contractor and needs to pick up some architectural prints in a city 60 miles from your location. They need these prints today or they will lose a large client and traffic going into the city will prohibit travel by car before the close of business. You were planning to fly that way to gain some additional IFR experience. You offer to fly your friend to the airport nearest the architectural firm and drop your friend off. You will then fly to another airport 40 miles away as intended earlier and practice flying approaches. You will then return to the city and pick up your friend and both of you will fly home together. The weather is marginal VFR along the entire route of flight (FAA, 2009a, p. 94).

The ultimate learning situation is real life and SBT tries to create as realistic setting for the lesson as possible. Simply, SBT puts learning in context. In other words, the more realistic the situation and scenario, the more learning is enhanced (Summers et al., 2007).

## Single Pilot Resource Management

Single Pilot Resource Management (SRM) is defined as “the art and science of managing all the resources (both on-board the aircraft and from outside sources) available to a single-pilot (prior and during flight) to ensure that the successful outcome of the flight is never in doubt.” (Summers et al., 2007, p. 13). SRM includes the concepts:

- Aeronautical Decision Making (ADM)
- Risk Management (RM)
- Task Management (TM)
- Automation Management (AM)
- Controlled Flight Into Terrain (CFIT) Awareness
- Situational Awareness (SA)

A carefully crafted scenario will require the pilot being trained to exercise all six elements of SRM before and/or during the lesson. Pilots will be expected to start assessing all of the risks that might be encountered during the flight during the planning process and continue to assess risks as the flight progresses. Pilots are expected to utilize the technology available, including traffic and terrain awareness displays, datalinked information, and autoflight systems to enhance situational awareness, manage workload, and make proper decisions.

## The SRM Decision Process

Pilots in training may have a wide range of experience and certifications. Therefore, there is no one “best” scenario. In FITS scenarios, there is not one right answer, rather each pilot is expected to analyze each situation in light of their experience level, personal minimums, and current physical and mental readiness level and make their own decisions. The result is there can be multiple successful outcomes to a lesson. This is one aspect that instructors need to consider and not insist on only one possible course of action by the student.

The SRM scenarios developed by the FITS team incorporate several maneuvers and flight situations into realistic flight scenarios. The scenarios are much like the Line Oriented Flight Training employed by the major corporate and airline training organizations for years. Table 1 gives an example of the performance, standards, and conditions for using SRM.

Table 1

### *Single Pilot Resource Management*

<b>Performance</b>	<b>Standards</b>	<b>Conditions</b>
The training task is:	The pilot in training will:	The training is conducted during:
1. Task Management (TM)	Prioritize and select the most appropriate tasks (or series of tasks) to ensure successful completion of the training scenario.	Note: All tasks under SRM will be embedded into the curriculum and the training will occur selectively during all phases of training. SRM will be graded as it occurs during the training scenario syllabus.
2. Automation Management (AM)	Program and utilize the most appropriate and useful modes of cockpit automation to ensure successful completion of the training scenario.	Note: All tasks under SRM will be embedded into the curriculum and the training will occur selectively during all phases of training. SRM will be graded as it occurs during the training scenario syllabus.

3. Risk Management (RM) and Aeronautical Decision-Making (ADM)	Consistently make informed decisions in a timely manner based on the task at hand and a thorough knowledge and use of all available resources.	Note: All tasks under SRM will be embedded into the curriculum and the training will occur selectively during all phases of training. SRM will be graded as it occurs during the training scenario syllabus.
4. Situational Awareness (SA)	Be aware of all factors such as traffic, weather, fuel state, aircraft mechanical condition, and pilot fatigue level that may have an impact on the successful completion of the training scenario.	Note: All tasks under SRM will be embedded into the curriculum and the training will occur selectively during all phases of training. SRM will be graded as it occurs during the training scenario syllabus.
5. Controlled Flight Into Terrain (CFIT) Awareness	Understand, describe, and apply techniques to avoid CFIT en-counters: a. During inadvertent encounters with Instrument Meteorological Conditions (IMC) during VFR flight. b. During system and navigation failures and physiological incidents during IMC flight.	Note: All tasks under SRM will be embedded into the curriculum and the training will occur selectively during all phases of training. SRM will be graded as it occurs during the training scenario syllabus.

---

*Note: Reprinted with permission of the author from "Managing Risks through Scenario Based Training, Single Pilot Resource Management, and Learner Centered Grading," by M. Summers, F. Ayers, T. Connolly, and C. Robertson, 2007.*

## **The 5 P Check**

In order to be effective, pilots must understand and use SRM in their daily flights. The FITS team developed a recurring SRM check referred to as the "5 Ps": the Plan, the Plane, the Pilot, the Passengers, and the Programming (Summers et al., 2007). The "Plan" involves assessing how the plan is progressing and how the plan is being affected by factors such as weather and traffic delays as the flight progresses. The "Plane" is assessing the condition of the plane and its systems, fuel state and capabilities affecting the progress of the flight. Factors such as a degraded or inoperative system or the lack of a system, such as de-icing capability, can have an influence on the risks facing the flight and influence the pilot's decisions. The "Pilot" is a self-assessment that considers such factors as fatigue,

ratings, currency, and experience affecting the flight. The “Passengers” is where the pilot assesses the passengers and determines how they may affect a flight. A passenger recovering from a cold may have ear problems during the flight. A passenger who is a pilot may be able to assist during periods of high workload or during an emergency situation. Finally, the “Programming” assesses the state of the navigation and automation systems during the flight to insure that it is current, in the proper mode, and is properly programmed for the current and subsequent stage of the flight.

Each of these areas can present areas of risk or opportunity for the pilot during the planning and execution of a flight. Assessing and managing risk through good ADM requires the pilot to diligently practice situational awareness. Assessing each of these five areas on a continual basis throughout a flight helps maintain that situational awareness and ultimately enhance good ADM. The “5 Ps” should be evaluated at key points during the flight, when the situation requires a change to the flight, or when an emergency arises. These decision points include, pre-flight, pre-takeoff, hourly or at the midpoint of the flight, pre-descent, and just prior to the final approach fix or for VFR operations, just prior to entering the traffic pattern (Summers et al., 2007). The “5 Ps” have been used successfully in group seminars as well as during flight and simulator training to show how to continually assess risk and promote good decision-making.

### **Learner Centered Grading**

The third component of the FITS training method is debriefing the flight using learner-centered grading. One of the goals of using this debriefing method is to promote self-assessment skills to enhance life-long learning and as a process of continuous self-improvement. Actively engaging the student in the assessment process fosters greater retention and understanding (FAA, 2009a).

A debriefing using LCG is a bit different from the traditional instructor-led debriefing. In LCG, the student is given a copy of the grading sheet and asked to grade his or her performance. The instructor will also grade the student’s performance. When each is done, the student and instructor compare their grades and discuss the differences. It is during this discussion that active learning takes place as well as the student gaining better self-assessment skills (FAA, 2009a). Instructors are encouraged to use open-ended questions during the debriefing to help stimulate the discussion. This method of debriefing is more along the lines of a facilitated debrief found in LOFT-type training.

### **Desired Outcomes**

A key component of LCG is the grading scale that is utilized. Traditional grading schemes, such as “excellent,” “good,” “satisfactory,” “marginal,” and “unsatisfactory,” are not used. These grading schemes are often based on empirical measurements and can be somewhat subjective. What may be “excellent” to one instructor might be only “satisfactory” to another. Also, these methods often do not take into account student progress and what might be excellent early in training might be marginal later in the course. The LCG method attempts to bring more realism into the evaluation process and focuses more on achieving desired outcomes in readily identifiable and measurable terms.



Because one of the goals of FITS training is to enhance risk management and decision-making, traditional grading methods would not work very well. How does one empirically differentiate between a good and very good grade for a student learning how to do a maneuver? How does one differentiate between a very good and excellent, or a 4 and a 5 or a B and an A, when it comes to making a decision? These grading methods, while seemingly empirical, are really very subjective and dependent on an instructor's opinion.

With LCG, grades are divided into two categories, Maneuver, or task grades, and SRM grades. The two grading methods are described below. In this grading description, the student is referred to as the PT or pilot-in-training. The instructor is referred to as the CFI for Certificated Flight Instructor. Notice that a key indicator of the differences between the grade levels is the degree of instructor intervention that is required to accomplish the task.

#### Maneuver Grades (Tasks)

Describe – at the completion of the scenario, the PT will be able to describe the physical characteristics and cognitive elements of the scenario activities. Instructor assistance is required to successfully execute the maneuver.

Explain – at the completion of the scenario, the PT will be able to describe the scenario activity and understand the underlying concepts, principles, and procedures that comprise the activity. Significant instructor effort will be required to successfully execute the maneuver.

Practice – at the completion of the scenario, the pilot in training will be able to plan and execute the scenario. Coaching, instruction, and/or assistance from the CFI will correct deviations and errors identified by the CFI.

Perform – at the completion of the scenario, the PT will be able to perform the activity without assistance from the CFI. Errors and deviations will be identified and corrected by the PT in an expeditious manner. At no time will the successful completion of the activity be in doubt. ("Perform" will be used to signify that the PT is satisfactorily demonstrating proficiency in traditional piloting and systems operation skills)

Not Observed – Any event not accomplished or required

#### Single Pilot Resource Management (SRM) Grades

Explain – the pilot in training can verbally identify, describe, and understand the risks inherent in the flight scenario. The pilot in training will need to be prompted to identify risks and make decisions.

Practice – the pilot in training is able to identify, understand, and apply SRM principles to the actual flight situation. Coaching, instruction, and/or assistance from the CFI will quickly correct minor deviations and errors identified by the CFI. The pilot in training will be an active decision maker.

Manage/Decide – the pilot in training can correctly gather the most important data available both within and outside the cockpit, identify possible courses of action, evaluate the risk inherent in each course of action, and make the appropriate decision. Instructor intervention is not required for the safe completion of the flight.

Not Observed – Any event not accomplished or required.  
(Summers et al., 2007, p23-24)

These grades are consistent throughout the course and do not change as lessons progress. These grades are based on the overall desired outcome for each course. Early lessons may have the desired outcome for particular items shown as Explain or Practice whereas the later lesson will have the desired outcomes for those same items as Perform or Manage/Decide. Figure 1 shows the Desired Outcome Grade Sheet for Lesson 2 of the FITS Generic FIS-B, TIS-B, and ADS-B syllabus (FAA, 2009b, p39). The Desired Performance column listed the completion standards for the lesson and the Task Grades and SRM Grades columns are where the demonstrated performance is recorded.

## Discussion

### Validation

The FITS training concepts were validated in at least 15 studies conducted by CGAR member universities and by training providers in various parts of the United States. Research studies, presentations, guidance, curricula, and generic training syllabi can be found on the FITS website at [http://www.faa.gov/education\\_research/training/fits](http://www.faa.gov/education_research/training/fits).

Initial research results from those studies indicate that the pilots completing these FITS courses have the same or better level of aeronautical competency over traditionally trained pilots and a higher level of skills in the key areas of single pilot resource management without sacrificing the piloting skills developed in traditional maneuvers-based training and testing (Federal Aviation Administration, 2009a). Also, the time required to achieve these skills is no more, and sometimes less, than in traditional maneuvers-based training programs (Federal Aviation Administration, 2009a).

Lesson 02			Task Grades				SRM Grades		
Desired Outcome Grade Sheet			Not Obs	Explain	Practice	Perform	Explain	Practice	Manage
Scenario Activity	Task	Desired Performance							
Preflight Preparation	<a href="#">Weather Information</a>	Perform							
	Flight Planning	Perform							
	SRM	Manage/Decide							
Preflight Procedures	Aircraft Preflight	Perform							
	ADS-B Equipment Check and Certification	Perform							
	Cockpit Check and Organization	Perform							
	SRM	Manage/Decide							
	Use of ADS-B for Taxi Awareness/Runway Incursion Prevention	Manage/Decide							
Takeoff and Departure Operations	Pre-Takeoff Procedures	Perform							
	Normal Takeoff and VFR/IFR								
	Departure Procedures and Navigation	Perform							
	Operation of ADS-B Avionics	Perform							
En route Operations	SRM	Manage/Decide							
	VFR/IFR Navigation	Perform							
	Operation of ADS-B Avionics/ADS-B Malfunctions	Perform							
	Use of Automation	Manage/Decide							
	Collision Avoidance using ADS-B/TIS-B	Perform							
	Display FIS-B Products	Perform							
	SRM	Manage/Decide							
Arrival Operations	Pre-Arrival Procedures	Perform							
	VFR/IFR Navigation and Arrival Procedures	Perform							
	Operation of ADS-B Avionics	Perform							
	Use of ADS-B for Spacing, Sequencing and Merging into Traffic Patterns and Approaches	Manage/Decide							
	Display of FIS-B Products	Perform							
	SRM	Manage/Decide							
Approach and Landing Operations	VFR Traffic Pattern/ Instrument Approach	Perform							
	Operation of ADS-B Avionics	Perform							
	Display of FIS-B Products	Manage/Decide							
	Normal Landing	Perform							
	SRM	Manage/Decide							

Note: Reprinted with permission from "FITS generic scenario based private/instrument certification syllabus for piston aircraft," by FAA, 2009.

Figure 1. Desired Outcome Grade Sheet.

## Implementation

The FAA is moving to establish FITS as the standard in general aviation flight training. Consequently, the entire training system (training and checking) needs to be available to the general aviation community. This includes training tools (syllabi) and Practical Test Standards changes that support FITS training. In addition, more guidance has been developed to help Certified Flight Instructors, Designated Pilot Examiners, and FAA Aviation Safety Inspectors train and evaluate pilots trained under a FITS program. The FITS team has produced over 20 documents, including sample syllabi, to provide guidance for instructors and training providers to understand and implement FITS. The FAA is beginning the process of updating the Practical Test Standards beginning this year. Sample FITS lesson plans are being developed for posting on the FAA's website for instructors and designated examiners to access.

The are four levels of FITS acceptance depending upon the type of training being conducted (FAA, 2007). Detail guidance for FITS acceptance is provided on the FITS website and in the FAA orders provided to aviation safety inspectors.

These four levels are:

- Accepted FITS Flight Syllabus- required training in an airplane, simulator, or advanced aviation-training device.
- Accepted FITS Syllabus (Non-Flight)- used to enhance a certain set of skills such as the use of a new glass cockpit display or to better develop SRM skills without actually flying, but uses a live instructor to conduct the training.
- Accepted FITS Self-Learning Program- a computer based program using a CD, DVD, or online program to conduct training for a specific application or purpose.
- Accepted FITS Supporting Material- materials that do not meet all of the requirements for any of the other three levels but can be used as part of a FITS curriculum or lesson (FAA, 2007).

Over the past seven years, the FAA/Industry Training Standards (FITS) program has grown exponentially. This growth has been driven by the general aviation industry desire to participate rather than FAA mandates. Not only is the aviation industry integrating FITS tenets into its training programs, but the rapid growth of glass cockpits in smaller general aviation airplanes has created a demand to expand FITS training concepts into more industry training programs. The FITS team worked directly with many aircraft manufacturers (or their chosen training providers) to develop appropriate training programs. Today, all of the major general aviation aircraft manufacturers, including Cessna Aircraft, Diamond Aircraft, and Cirrus Designs, offer a single engine and/or twin-engine aircraft with an accepted FITS training program (Glista, 2005). The FITS research team also worked with numerous universities to establish an accepted FITS training program. More than a dozen universities have implemented the FITS methodology into their flight-training curriculum. To date, there are approximately 150 FITS accepted training programs being used in the United States by over 30 training providers and manufacturers.

## Conclusion

The FITS program is not a regulatory or government mandated program. It was created with a unique partnership between industry, academia, and the FAA. Utilizing the lessons learned by the airline industry during the development of CRM and LOFT, academic studies about learning, the Alaska Capstone program, and NASA's AGATE and SATS programs, the FITS team developed a program designed to help general aviation pilots meet the challenges of the new technologies being introduced into the cockpit and National Airspace System.

The FITS program is centered on three key elements; SBT, SRM, and learner-centered grading (LCG). The FITS program emphasizes SRM with the ultimate goal of increasing safety on all levels of performance. FITS teaches the pilot much

earlier in training to develop decision making skills that are practiced with every flight lesson.

The fact that so many organizations have voluntarily decided to incorporate FITS methods into their training and seek FITS acceptance is testament to its success. The FITS program stands out as an example of how industry, academia, and the regulating body can work together to create workable and useful solutions that make sense and can be another part of the effort to increase safety in general aviation.

## References

- Federal Aviation Administration (2003). *General aviation technically advanced aircraft: FAA/Industry safety study*. Retrieved on August 8, 2009, from [http://www.faa.gov/training\\_testing/training/fits/research/media/TAA%20Final%20Report.pdf](http://www.faa.gov/training_testing/training/fits/research/media/TAA%20Final%20Report.pdf).
- Federal Aviation Administration (2007). *FITS acceptance criteria and levels of accepted FITS training*. Retrieved on August 8, 2009, from [http://www.faa.gov/training\\_testing/training/fits/guidance/media/FITS\\_levels\\_criteria.pdf](http://www.faa.gov/training_testing/training/fits/guidance/media/FITS_levels_criteria.pdf).
- Federal Aviation Administration (2009a). *FITS generic scenario based private/instrument certification syllabus for piston aircraft*. Retrieved on August 8, 2009, from [http://www.faa.gov/training\\_testing/training/fits/training/generic/media/pvt-inst.pdf](http://www.faa.gov/training_testing/training/fits/training/generic/media/pvt-inst.pdf).
- Federal Aviation Administration (2009b). *FITS generic ADS-B, TIS-B, and FIS-B syllabus*. Retrieved on August 8, 2009, from [http://www.faa.gov/training\\_testing/training/fits/training/generic/media/FIS\\_TIS\\_ADS.pdf](http://www.faa.gov/training_testing/training/fits/training/generic/media/FIS_TIS_ADS.pdf).
- Federal Aviation Administration. (2010). *Federal aviation regulations and aeronautical information manual*. Aviation Supplies and Academics: Newcastle, WA.
- Funk, K. (1998). *Training approaches and considerations for automated aircraft: A summary of training development experiences*. Corvallis: Oregon State University.
- Glista, T. (2005, April-May). *FAA/Industry training standards program update*. *FAA Aviation News*. Retrieved on August 10, 2009, from [http://www.faa.gov/training\\_testing/training/fits/guidance/](http://www.faa.gov/training_testing/training/fits/guidance/).
- Schuetz, S. (2003). *Scenario-based training for safe schools*. Retrieved on August 18, 2009, from <http://www.nlectc.org>.
- Summers, M. M., Ayers, F., Connolly, T., & Robertson, C. (2007). *Managing risks through scenario based training, single pilot resource management, and learner centered grading*. Retrieved on August 18, 2009, from [http://www.faa.gov/training\\_testing/training/fits/guidance/media/RM\\_thorough\\_SBT.pdf](http://www.faa.gov/training_testing/training/fits/guidance/media/RM_thorough_SBT.pdf).
- Wright, R. A. (2002) *Changes in general aviation flight operations and their impact on system safety and flight training: A white paper*. Retrieved on August 18, 2009, from [http://www.faa.gov/training\\_testing/training/fits/guidance/](http://www.faa.gov/training_testing/training/fits/guidance/).

## ***The Reciprocal Development of Expertise in Air Traffic Control***

Christine Owen

and

Wendy Page

*Faculty of Education  
University of Tasmania  
Hobart  
Australia*

[Christine.Owen@utas.edu.au](mailto:Christine.Owen@utas.edu.au)

### ***Abstract***

*Learning in the high reliability domain has vital importance where public safety is dependant on the expertise and proficiency of practitioners. Air traffic control is one example of work that may be characterized as a technologically complex, high stress, real-time environment with little margin for error. This study investigates learning in the air traffic control workplace in the context of workplace changes, which required all experienced controllers to act as instructors of trainee controllers. Initially prescribed as a strategy to enhance organizational flexibility and to increase workplace efficiency, the initiative was shown to have unintended effects. The results reported show how an incidental and unintended reciprocal learning opportunity (through engaging in on-the-job-training) enriches the reflective learning process for instructors. Organizational strategies, which serve to refresh the expertise of practitioners while addressing the needs of workplace training, offer wide-ranging benefits in the high reliability domain where reflective practices are necessarily limited by the intensity and immediacy of the work.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

## **The Reciprocal Development of Expertise in Air Traffic Control**

In the past few decades, a range of studies have served to highlight the importance of learning in the workplace, which has been shown to occur in a myriad of ways (Engestrom, 2001; Eraut, Alderton, Gerald, & Senker, 1998) including formal on-the-job training programs, as well as through informal and incidental learning (Collin, 2006). Given the dynamic changes in work, work organization, and the increasing attention directed at workplace learning in all its forms, an understanding of the role of the development of expertise and its place in formal and informal workplace learning becomes critical. This is especially so in the high reliability domain, where, though disturbances may be regarded as “opportunities for use and expansion of individual competencies and for organizational innovation and change” (Grote, Weichbrodt, Gunter, Zala-Mezo, & Kunzle, 2008, p.19), the latitude given to local actors to handle uncertainties in a flexible manner is necessarily constrained by the need for standardization and centralization to prevent system failure.

Since high reliability organizations, such as air traffic control, “are unforgiving environments where errors can have devastating consequences,” (Fogarty & Buikstra, 2008, p.199). They propose that, in order to maximize the safety climate, organizational processes and practices should be more closely examined. Further, Hudson (2007) warns of the danger of increasing academic engagement in the measurement instruments of safety, without undertaking a more comprehensive engagement with the industry itself.

While Hubbard (2008) maintains that air traffic control is a field that is particularly resistant to observation by outsiders, Hudson (2007) notes that, “failures arising from the culture of an organization have become seen as the reason why major accidents happen” (p.698).

Safety will continue to be critical as high reliability industries, such as air traffic control, become more complex and technologically driven (Fogarty & Buikstra, 2008). It is imperative that strategies to enhance expertise are supported in high reliability domains such as air traffic control, where the organization of work is such that its intensity and immediacy limit reflection, thus constraining learning. Nor is it possible to learn by experimentation. It follows that a consideration of the enhancement of expertise in the context of workplace learning and culture in air traffic control is compelling.

### **Maintaining and enhancing expertise**

It is now generally accepted that, in the future, almost everyone will be required to both continue their learning and even intermittently relearn aspects of their professional skills. Therefore, understanding how expertise is acquired, maintained, and extended within the workplace becomes of particular importance.

In the Cambridge Handbook of Expertise and Expert Performance, an expert is defined as “one who is very skillful and well-informed in some special field” (Ericsson, Charness, Feltovich, & Hoffman, 2006, p.3). According to this perspective, expertise refers to the attributes (including knowledge and skills) that distinguish experts from novices. However, the changing contexts of work and workplace learning, the assumptions underpinning traditional theories of skills, and the de-



velopment of expertise are being challenged in three ways. First, much theory of expertise is based on the assumption that the knowledge base is stable, even static (e.g., Chi, Glaser, & Farr, 1988; Ericsson et al., 2006). Second, there is the assumption that expertise is acquired individually when cognition becomes understanding, which is then internalized (e.g., Chi, Glaser, & Farr 1988). Third, there is the assumption that expertise is extended unidirectionally—as a transmission from more knowledgeable others to the less knowledgeable. That is, novices learn from those more proficient and proficient people become expert by seeking out more knowledgeable others and engaging in mentoring, deliberate practice, and reflection (e.g., Zimmerman, 2006).

This paper will demonstrate that each of these assumptions needs to be challenged, in order to provide a better account of how expertise may be acquired, maintained and developed in contemporary and dynamic workplaces like that of air traffic control.

For the purposes of this paper, we adopt Pillay and McCrindle's (2005) definition of expertise. They argue that professional expertise develops within a given domain of knowledge only as a result of contextualized training and practice. They define expertise as the "ability to combine domain knowledge with appropriate professional tools and strategies to solve problems within the socio-cultural context of the profession" (Pillay & McCrindle, 2005, p.67). The challenge for contemporary organizations undergoing changes to the nature of work and its organization, then, is to understand how expertise is acquired, maintained, and developed in the context of everyday practice.

Essentially the paper proceeds as follows: First it will provide a brief review of theories of expertise and why they need to be challenged. Then it will provide an outline of the type of work undertaken in the domain of air traffic control. Next the methods used in collecting the data and conducting the qualitative analysis will be outlined. A number of key themes pertaining to expertise will then be discussed which will include implications for the future of work and learning.

## **Expert cognition and the role of experiential learning**

Most theories of expertise contend that expertise is predicated on a highly evolved knowledge base, as well as highly automated skills that are developed over years of practice. This highlights the importance of the developmental nature of skill acquisition and the role of experience.

In studying the development of expertise in nursing, for example Benner, Tanner, and Chesla (1996) identified five stages of skill acquisition:

1. The *novice* or beginner, for whom performance is typically rule governed.
2. The *advanced beginner*, who can identify recurring meaningful situational components.
3. The *competent*, who can begin to see their actions as part of a long-range goal or plan to address the problem.

4. The *proficient*, who perceives situations as wholes.
5. The *expert* nurse, who typically is said to operate from a deep understanding of the total situation and whose responses become automatic.

The strength of this research is that it shows the various stages of growth within the developmental stages of expertise, between being a novice and becoming an expert. It also highlights the difficulty of capturing descriptions of expert performance because the expert is no longer aware of the rules, maxims, strategies he or she is calling upon to guide expert performance (Benner et al. 1996; Benner, 2001, 2004, 2005) because they perform automatically. Thus, the expert may not be able to articulate the elements of their practice. In these cases, individuals who have been performing the task for a long time forget which maxims and rules they are invoking in undertaking the work, making the skills learned opaque, even to themselves.

However, these static approaches to expertise have recently been challenged. One of the strongest criticisms of cognitive or information-processing approaches to expertise (e.g., Engestrom, 2004) is that proponents of these models assume a knowledge base that is stable, even static, perceiving expert performance to be domain specific and surprisingly limited (see for example, Ericsson et al., 2006). This gap has led to an increasing call to develop strategies to investigate “the way human practices emerge at work: as societally located and socially intelligible actions of reasoning and communication” (Engestrom & Middleton, 1996, p.3).

The second assumption, that expertise is acquired individually has also been challenged because it ignores the role of social context in the development of expertise. These issues have been discussed recently in the aviation industry. Hoover (2008), for example, calls for a consideration of social learning theory in flight instruction as “social learning is affected by the culture in which the individual is enmeshed and cognitive development results from shared experiences and interactions with individuals or groups that include both instructors and more competent peers”(p.364). Hoover cites the concept of reciprocal teaching, based on Vygotsky’s (1978) theory of social cultural learning, as offering a richness and depth currently absent from flight training curricula. It is from this perspective, which we have sought to examine the role of expertise and the enhancement of workplace learning, both formal and informal, in the setting of air traffic control.

### **Situating expertise in practice and the role of culture**

It is well established within socio-cultural theories of learning that all individuals, groups, and organizations operate in dynamic environments influenced by contexts. Socio-cultural theories of learning emphasize both the situated nature of the context as well as the importance of the collective and culture to learning. According to socio-cultural theories, knowledge is not only mentally structured but is practice-based, embedded in the everyday experiences of acting, negotiating and problem solving within the participatory process of working (Lave, 1993). Such learning is seen further as intertwined with the technical performance of work. Social networks are, thus, conceived as a shared social practice (Collin, 2006; Gherardi, 2001; Schulz, 2005), the concept of “legitimate peripheral participation” (Lave & Wenger, 2005) being used to characterize the ways in which novices

become competent members of the community of professional practice as a collaborative enterprise.

The distinction is drawn between cognitive information processing approaches to expertise and socio-cultural approaches in that the focus of attention switches from the individual to a community of practice. In this way, Konkola, Toumi-Grohn, Lambert, and Ludvigsen (2007) argue that expertise is not just developed inside the practitioner's head but also "expands the structures of knowledge to include not just mental and symbolic representations but also physical artifacts and recurring patterns of social practice" (p. 213-214).

While these approaches provide a more satisfactory account of learning and transfer in the development of expertise, they are still problematic. As Konkola, et al. (2007) notes, there remains an assumption that the community of practice is stable. It is also important to observe that the community of practice is depicted as benign; meaning that expert others are willing and able to support less knowledgeable others. However, in workplaces this is frequently not the case (see for example Collinson, 1992; Owen, 2009). Communities of practice sometimes encompass cultures that can work against valuing learning-oriented work practices and, thus, the development of particular forms of expertise. Moreover, both information processing and socio-cultural views still assume a unidirectional flow in the development of expertise from novice to expert. If it is accepted that knowledge is embedded in artifacts and cultures that are reciprocally determined (Pillay & McCrindle, 2005) then why insist on a one-way flow when it comes to developing expertise?

Engestrom (2004) suggests that, because of the increasing complexity and abstraction found within contemporary workplaces, a new interpretation of expertise in work organizations is required. He is particularly concerned with organizations undergoing transformational change where little may be known about the problem at hand and in need of resolution.

For Engestrom (2004, p.163) there is a need to develop a new generation of "collaborative and transformative expertise" that is based on the capacity of working communities to cross boundaries, negotiate, and improvise in order to reshape their activities. Its rules include "transparency and reciprocity" (p.163). Improvisation and reshaping activity is needed to address disturbances or breakdowns that result in tasks that are impossible to resolve. "Experts must face, diagnose, and resolve novel situations for which they have little or no directly applicable practice" (Engestrom, 2004, p.146). The conditions outlined by Engestrom are important and are likely, as he contends, to become increasingly frequent as dynamic change increasingly affects workplace practices. Thus, it is contended that the skills of transformative expertise as a new idiom are likely to be increasingly required.

Nevertheless, it is also critical to note that not all tasks will be *impossible to achieve* and, thus, require organizational transformation. Institutions and organizations will still face problems that are mundane, requiring attention to the developmental acquisition of learning through experiences in everyday practice.

A discussion of the type of undertaken in the air traffic control workplace will outline both the type of learning and the kinds of expertise required, as well as

summarizing some of the challenges of change which occurred in this workplace at the time of the study.

### **Air Traffic Control Work**

The goal of air traffic control work and the tasks of air traffic controllers are to maintain separation between aircraft in a way that is safe and allows for expeditious flow of air traffic. Air traffic controllers both direct the flow of traffic and provide in-flight information to assist aircrew in the operation of their aircraft.

Even though air traffic control practice involves applying just three standards or rules of separation (i.e., keeping aircraft separated vertically, laterally, and longitudinally), the work is complex because of a range of other factors. For example, aircraft may be required to divert from an original flight plan due to poor weather or, in a desire to get above or below poor weather conditions, aircraft may request flight level changes. Crosswinds or tailwinds may alter an aircraft's performance, resulting in the aircraft not performing as anticipated. Other environmental conditions (e.g., bushfires, fog) can also alter the flight plan. The performance of each aircraft can also vary and this may be due to the aircraft profile, its payload, company policy, and even how a pilot "drives" the aircraft.

The organization of air traffic control work is divided into the phases of the flight and is shared between:

- **Tower**, which provides airport control and surface movement control, and
- **Control Centre**, which provides
  - Approach control (responsible for aircraft approaching and departing the airport- approx 30 miles from the airport);
  - Area or "Enroute" control (aircraft travelling to and from their destination in what is typically describes as the cruise component of the flight), and
- **Arrivals control** (preparing for landing which commences approx 120 miles from the airport).

In Australia, technological developments had led to the centralisation of work activity to two Centres (one for Northern and Southern Australia) at Brisbane and Melbourne. This decision led to a reduction of air traffic control presence at other major airports. National industrial relations policies were also being introduced at the time that aimed at reducing job role specialisation and increasing labour flexibility. These were sometimes referred to as "multi-skilling" initiatives and at other times "multi-tasking" arguably because the changes were not enhancing the skill levels of workers, just getting them to do different jobs (Marginson, 2005). Given that in some communities the term "multi-tasking" can mean personnel undertaking several jobs at once, the term "multi-skilling" will be used here. In Air Services Australia, for example, this was evident in the implementation of industrial relations changes that resulted in the restructuring of work practices by combining all operational staff into the one industrial agreement and in requiring that every rated controller was required to become a "Full Performance Controller" (FPC) gener-

ally within three to five years of achieving an initial rating. To do this the controller must achieve a rating on three airspace sectors and maintain “currency” (i.e., be up to date and therefore able to operate) on these sectors. In addition, it was also a requirement that every Full Performance Controller become a workplace instructor of others within six months of achieving their FPC rating. Other industrial changes included flattening the organizational hierarchy from seven levels of Air Traffic Controller pay-scales to two (Journeyman Controller and Full Performance Controller), putting all controllers in essence on the same scale. Prior to this controllers would commence on an en-route sector to gain experience of controlling aircraft in flight and then progress as their experience grew to other sectors closer to the airport that arguably required faster decision-making skills and had greater demands (which had in turn been represented in a higher pay scale for Approach).

In summary, air traffic control work is mediated by technologies and involves a high level of responsibility and reliability. It involves multiple agents, higher order thinking (where the path of action is not fully specified in advance and yields multiple solutions), in a context of imperfect information, uncertainty, and constantly changing conditions, time pressure which at times can create a strong sense of urgency. In this type of workplace, there is little capacity when engaged in the work to stop it in order to enable reflection and debriefing – something that can be done in other work domains – because in this context the work cannot be stopped. Thus, enabling resilience through flexibility is necessarily constrained by safety considerations. Since reflection in action is impossible, new and innovative ways to enhance expertise offer coherence and efficiency within the context of organizational transformation, without compromising safety systems.

### **Research Design**

This study of learning in the air traffic control workplace was conducted over a five year period and utilized an ethno-methodological (e.g., Garfinkel, 2007) qualitative research design, where a stratified sample of one hundred air traffic controllers were interviewed, in some cases on multiple occasions, across three air traffic control Centers in Australia. The basic premise of the research was that the lived experience of people at work is significantly influenced by their contexts (in organizations, most commonly conceptualized as structures and cultures), that these changing contexts have implications for learning, and are in turn reproduced or transformed by people. The research questions examined are included in Appendix A. Two that are pertinent in this paper are “How does learning occur in the air traffic control workplace?” and “In what ways do organizational structures and cultures enable and constrain learning in the workplace?” For space limitations, in this paper only incidental instructor learning is discussed.

Considerable debate exists about the kinds of strategies that are appropriate to rigor in qualitative research. At the heart of these debates are issues of which epistemological position the debater comes from: a more subjectivist (Hassard, 1990; Van Maanen, 1988) epistemological position or an objectivist one (Burrell & Morgan, 1979). Given the goal of this research was to uncover the lived experience of participants the approach taken here is one of attending to the participants’ perspectives. As Gillett (1995) concluded “once one sees the tasks of understanding human behaviour as involving interpretation and empathy rather than prediction or control, the self-reports of the subject become very important” (p. 111).

Nevertheless, it is still important to adopt a conscious interpretative stance, and to have verifiable ways of establishing the credibility of the data.

The standards typically applied within professional communities of qualitative researchers relate to the degree to which the “trustworthiness,” “credibility,” “transferability,” “dependability,” and “confirmability,” are empirically verifiable. These are argued to be the qualitative equivalents of “internal validity,” “external validity,” “reliability,” and “objectivity” (Denzin and Lincoln, 2000). In this research these standards were applied in the following ways.

To ensure “trustworthiness” of data, researchers need to make use of multiple and different sources, methods and theories (Patton, 2002) and to interview an extensive number of informants to provide supporting evidence. In this study interviews were conducted with a stratified sample of the main roles of air traffic control work, within three different geographical locations, attending also to ensure controllers with differing demographic characteristics (e.g., experience) were included. “Credibility” of the findings was strengthened by prolonged engagement in the field and building trust with participants (Yin, 2003). In this study, this occurred over a five year period.

The strength of the theoretical argument is also a key feature of demonstrating “transferability” as well as credibility (Strauss & Corbin, 1994). Transferability of data occurs at two levels: the degree to which the findings are transferable within the population studied and the degree to which the results are transferable to other populations. Thick, rich description adds to verification according to Denzin and Lincoln (2000) because it allows the reader to make decisions regarding transferability to other settings because of the degree of detail in the information provided. “Dependability” of results relates to the issue of ensuring data collected is stable and consistent over time. Dependability is enhanced by strategies such as involvement in the field for extensive periods of time and extensive interviewing. Dependability is also enhanced by collecting data as part of an iterative process (Yin 2003). Dependability and confirmability are enhanced if threats to inaccuracy in data collection are reduced. These strategies can include, for example, use of audio- or video-tape and verbatim transcripts. “Confirmability” relates to the issue of whether there is a correspondence between what the study’s participants meant and what the researcher inferred. One way of ensuring confirmability is through member checks of the data (Denzin & Lincoln, 2000). Eight air traffic controllers involved in different parts of the organization research the manuscript of the full study. An indication of the veracity of the findings for the Australian Air Traffic Control population emerged some time after the completion of the study when the Air Services Australia library contacted the author to request six copies of the full study; such was the interest and demand in reading the findings.

## **Research Methods Employed**

Of the 100 interviews conducted, 25 were conducted with controllers acting in the role of instructor; 27 occurred with controllers who were reporting on their experiences of being both controllers as well as instructors; 36 were trainees (16 abinitio i.e., with no previous experience) and 12 in roles of human resource management. This paper draws on the 52 interviewees who were able to report on the experience of instructing. These interviews were conducted in Approach (n=13), Arrivals (n=12, Enroute (n=17) and Tower (n=10).



In collecting the data, three geographical sites were chosen – Brisbane, Melbourne, and a major airport in another capital city (Perth). At each of these facilities, interviews were conducted with personnel involved in all sectors of air traffic control work as well as with others involved in off-the job training and in positions of human resources and management.

The approach taken was an inductive theory-building one based on a semi-structured interview process. This is, a series of prepared questions guided the initial interview (see Appendix B), though not all questions were asked of all participants. This is typical in grounded theory and theory building iterative research processes (see Charmas, 2006; Denzin & Lincoln, 2000; Patton, 2002; Corbin & Strauss, 2008). At times interview topics diverged to follow a particular point made by the participant (which, if they proved pertinent in the subsequent coding and analysis process were then followed up and tested with other interviewees. Follow-up interviews, where they occurred, addressed questions that may not have been asked initially or were used to check out the respondent’s views on a particular theme that had emerged in another interview and needed to be further investigated. These strategies of triangulation of themes across multiple data sources enhance the veracity of the research method process

Interviews conducted were between thirty minutes and three hours in duration. Interviews were undertaken in the geographical locations typically in batches of 3-5 day periods. The audio taped interviews were then transcribed and coded before going back out into the field. When the interviews were transcribed, the data were entered into a software program for qualitative data analysis. When all the data had been collected, a process described by Tesch (1990) as *de-contextualising* was used which involved segmenting the data into meaningful units (i.e. a segment of text that is comprehensible by itself and contains one idea, episode, or piece of information). These segments become the beginning of an organising system or *pool of meanings* to which the data belong. This assembling is termed *re-contextualisation* and results in categories which are further refined to concepts or themes. The themes discussed in this paper, and their proportional representation across the interview sample are included in Table 1.

Table 1

*Themes of instructor learning and developing expertise*

Theme	Sub-theme	Number of participants who discussed	
		N	%
Developing controller expertise	Developing autonomy	35	68
	Developing automaticity	36	69



	Unlearning automaticity	22	42
	Shifting gears- re-learning how to learn	46	88
Instructor Learning	Developing instructional metacognition	20	38
	Instructor learning processes	41	79
	Reciprocal learning	33	63
	Developing expertise	33	63
Socio-cultural enablers	Organizational structures	25	48
	Organizational cultures	31	60
Socio-cultural constraints	Organizational structures	37	71
	Organisational cultures	43	83

It should be noted that each interviewee had the opportunity to discuss a range of themes.

## Results and Discussion

Given the space limitations in this paper, each of the themes relevant to instructor learning will be discussed and an example of the kind of comment made within that theme will be provided as an exemplar. The ways in which controllers first learn to develop their expertise is outlined first, followed by the strategies instructors engage in to redevelop their learning. Finally, the social cultural aspects that were identified to enable or constrain reciprocal learning and instructor development of expertise will then be outlined.

### Developing expertise as a controller

According to those interviewed, job content knowledge in air traffic control is built up through experience over a considerable number of shifts as controllers build up a reservoir of handling air traffic control situations or what (Suchman, 1996) referred to as learning "routine trouble".

**Developing autonomy.** Indeed, for many controllers, once they are licensed and able to operate on their own and to develop their own autonomy is when development of their own expertise begins.

Interviewer: "Someone has said to me that you don't develop technique until you have got your rating because they are not your techniques - you are just using somebody else's."

Respondent: "That's true. When you are finally rated and you waddle off with a rating in your pocket [that] is when you really start the learning session. All that has happened up until that point is somebody has sat

with you and beaten you across your head enough times to keep you out of trouble. Someone to say, “yes, they are safe to leave on their own, they shouldn’t bang two together.” And from then on, when you sit on your own, as you do when you go solo flying, and then you really start to figure out what is going on.” (Int. 83, Instructor).

**Developing automaticity.** Incidental learning, as the controller above explains, allows the controller to build up his or her own techniques and “style” of controlling — of working out what best works for them in solving the problems presented. In time, these experiences become automatic, allowing the controller to become efficient. In work environments that occur in real-time where the work at the console cannot be stopped, active reflection can indeed even slow the work process down. Direct experience and automaticity are needed to achieve efficiency. However, such automaticity can also become a problem, as the next controller explains.

“There was one person that we had, their skills had actually gone backwards when we moved them onto the next sector because they’d been there [on the old sector] for too long. Their thought processes and all [the skills] they had, had just gone into, how can I put it, autopilot, and they were just doing it by autopilot. But when they were moved onto the new sector, they had to learn again and it was like starting from scratch again because their ability to learn new skills was much lower than it would be if they were still in that learning process.” (Int. 53, Team Leader).

**Unlearning automaticity.** This change in sector requires the controller to stop working automatically and to re-attend consciously to the task at hand. When the controller becomes a trainee again, he or she needs to “shift gears” from acting automatically to acting with a deliberate conscious intention toward learning. Systematically moving from one sector to another and becoming a trainee again is also important for the maintenance of certain generic controller skills.

As the above extracts illustrate reflection is enhanced as a controller is deliberately placed outside his comfort zone into what Vygotsky (1978) called the “zone of proximal development,” thus, producing dissonance for the controller. In this respect, the multi-skilling introduced as part of the changes in organizational policies enhanced reflection and learning. By forcing the controller to become conscious of their practice, they have to reacquire the ability to learn and this supports the development of expertise.

**Shifting gears: (re)learning how to learn.** As discussed, structural changes in the organization of work included another initiative aimed at enhancing workforce flexibility through multi-skilling: that of requiring all controllers to become instructors. Once again, an unintended effect of doing so is to enhance the reflective strategies of the controller-instructor. Being able to reflect and think rather than act and (automatically) do is also one of the first things a new instructor needs to learn, in order to be able to monitor the emerging traffic scenario, as the following instructor explains.

“The hardest thing at first (when being an instructor) is to be able to concentrate on what is going on. You literally have to teach yourself to actually keep a full picture of what is happening. ‘*This has to be done now, he didn’t do it. I have got to remember that he didn’t do it.*’ That is the first

thing that you have got to teach yourself. That is very hard - being able to keep [a handle on] what is going on” (Int. 78, Instructor).

**Reciprocal learning.** The process of instructing sets up opportunities for controllers-as-instructors to observe and reflect on the job of controlling — something that is difficult to do when engaged in the temporally and cognitively demanding work that occurs as air traffic control undertaken in real-time that cannot be stopped. When a controller is instructing another, he or she sits or stands behind the person working at the console and watches the trainee controller undertaking the work. The practice of watching another do the work of controlling enables the controller to view the work activity from a different perspective. Controllers commented that having to become an instructor forced them to re-examine their own knowledge base. Stepping back from the job at the console had the advantage of giving the controller a different (wider) perspective.

“I think it helps because you are going back to basics for yourself. You probably even, in fact, improve your own performance greatly on training because you start calculating things again and find a lot of things that you would perhaps put in your little *judgment block* in the back of the mind, that you use all the time weren’t quite right. Maybe when you put them in they weren’t quite right. So with the trainee you are watching all the time and thinking, ‘*What would I do here? What would I do here?*’ so that you can come in if he is going to ask, ‘*What do I do now.*’ You say ‘*Do this*’ instead of having to think ‘*Oh God what do you do now?*’ - having it all worked out in your mind” (Int. 93, Instructor).

Thus by potentially setting the expert adrift from their accustomed sense of work practice, a “mundane accomplishment of directionality can be made explicit” (Engestrom, 2004, p.156). The expert is thrust beyond their automatic practice toward an explicitly conscious awareness of their practice.

Respondent: “The ‘conscious competence’ and ‘unconscious competence.’ I think that we really needed to know that. I thought that was really the first thing - what a good idea it was and then I realized that it is so much that when you are doing the job for years and you get a trainee [you realize that], not only do you know what you are doing but you do it automatically.” (Int. 79, Instructor).

Reflection-in-action is possible partly because of the overt distance from intuitive practice which instruction requires, thereby facilitating an aspect of *conscious enquiry*, which real time urgency and immediacy obscured in the practice of air traffic control work (Westrum, 1997).

In the context of work that has the characteristics of real-time intensity and complexity, the requirement to take on the role of instructing sets up the controller to engage in the process of observing a job and, in doing so, enhances the controller’s own opportunities for learning. Although this was not the intention of the organizational change, some instructors and team leaders commented that the instructional role is sometimes used as a strategy to enhance the performance of the instructor.

Interviewer: “What impact does having a trainee have on you and the way you do the job in terms of your currency?”

Respondent: "In some ways it improves you because you become more aware of what is going on around you and you look at other ways of doing things and being forced to sit back and watch it in detail - you analyse it a lot more. There is an old saying on Approach '*he's been doing the job for a year now; it is time for a trainee so he can really look at the whole thing.*' Quite often, training officers are selected from the point of view to improve the training officer not for the trainee's benefit. There is quite a bit of that going on and there's a lot of truth in it" (Int. 65, Instructor).

In considering the response above, a possible first reaction to this would be to suggest that the main purpose of becoming an instructor should be to enhance the trainee's performance not the instructors. However, from the perspective of the whole system of work activity it is argued that this would be an unduly limited view.

### **The Development of Reciprocal Expertise**

It is acknowledged that "explaining one's thinking to another leads to deeper cognitive processing" (Bereiter & Scardamalia, 1989), with reflection seen as an essential ingredient for learning (Brookfield, 1995; Schon, 1991). Historically, the term "reciprocal teaching" has been associated with peer learning in schools (Palinscar & Brown, 1984). However, this paper extends the concept to include a consideration of workplace learning. In applying it to a high reliability domain where learning from experimentation is necessarily limited due to safety considerations, such a suggestion has broad-ranging possibilities. In this context, the notion of reciprocal learning is important, because it enables the expert to step outside of their accepted codes of reference to refresh expertise and reframe their awareness for the purpose of teaching.

Embedding the role of instruction within the job descriptions of all controllers has had, in this organization, considerable benefits. Data collected from the interviews conducted in all three centres and in all air traffic control sectors revealed that undertaking the role of instructor is a structured task which improves the controller's own performance because it forces the controller instructor to think about, justify and articulate the bases for the decisions he or she would be making if doing the work. Performance enhancement aspects of the instructional role of controllers should be acknowledged within structures of work organization, thus institutionalizing a potential *reciprocal learning* opportunity.

Indeed, there is a paucity of relevant research on the question of instructor learning, with few studies investigating the potential benefits to instructors of the teaching role to enhance and refresh expertise. The novice/expert relationship has largely been seen as a unilateral one (Carrington, 2004), in which the primary benefit resides in the novice, while little consideration has been given to the possible reconfiguration/ transformation of existing knowledge that is afforded by taking on a teaching role.

It is imperative that strategies to enhance expertise are supported in high reliability domains, where the organization of work is such that its temporal dimension (intensity and immediacy) limit reflection, thus constraining learning. It is suggested that the expert may become less flexible "as practice becomes increasingly tacit and spontaneous, the practitioner may miss important opportunities to think

about what he is doing" (Schon, 1991, p.61), thus, potentially relying on flawed and error-prone decision-making processes. Similarly, Mitchell (2009) relates how a pilot's flying skills may diminish with the rise of automation in cockpits. "A short-coming of present automation systems is that they are not yet intelligent enough to cope with extraordinary situations. These situations are still left to the skills of experienced or talented pilots" (p.23). Complacency and over-reliance on automation may, thus, serve to reduce the flexible response required by expert practitioners in unexpected situations.

## **Socio-Cultural Contextual Influences on the Development of Expertise**

Nevertheless, there is still a needed relationship between developing expertise and experience, which can become problematic in workplaces undergoing structural and cultural change. In air traffic control, for example, the reciprocal development of expertise is also constrained by other changes in organizational and recruitment practices, which have led to a concentration of inexperienced controllers on particular en-route sectors.

**Structural impediments.** However, reciprocal development of expertise is not going to be a solution if the locus remains in the interactions between inexperienced instructors and their trainees, as the following instructor explains:

"I think, people training after six months... I mean it's okay because we are pretty fresh out, we still know our theory but you can teach people all the theory in the world and you can tell them how to do the job, but if you haven't got that experience base to hand on to them, they are really missing out I think. Then *they* haven't got that experience base. So then when their six months is up, they train and they have got no experience / [and] the level of what people can teach gets narrower and narrower. Whereas if you have got someone who has got years of experience they can pass on that knowledge and then they can pass on ...I think the nice-to-know stuff is getting less and less because all we know is the facts and so that is all we can teach. We can't pass on our years of knowledge." (Int. 49, Instructor).

For some instructors, particularly those in the Enroute sectors, the "dilution" of experience is a concern because it limits the alternatives that can be passed on to the trainee. Simply put, controllers-as-instructors "fresh out" do not have the job content knowledge to draw on. Therefore, controllers need opportunities to develop working alternatives in a range of situations before they feel comfortable taking on a trainee because their own experience level and lack of options is likely to limit the experiences they can provide to their trainees.

**Structural enablers.** This raises the question of how, under these circumstances organizations can make available the requisite variety (Weick, 2001) needed for trainees to access when engaging in on-the-job learning and after completing their certification. One means of facilitating the access to depth of job content knowledge for trainees is to make the process of facilitating trainee accredited learning the responsibility of the whole team, so that the depth of experience available within the team can be used as a resource.

There are ranges of ways in which this could be achieved that are not limited to simply sharing the instructional duties around within the team. In appropriate team cultures (one where inquiry-related behaviors are a norm of practice), team meetings could set aside time for trainees to share their experiences to date and seek input from other members. Such communicative practices could benefit the whole team because they may also provide new insights for team members, thus enhancing collective memory, or drawing attention to divergent practices that may need to be discussed and addressed.

Ensuring within an organizational structure that teams comprised members with a range of background experiences also would enhance the job content knowledge that would be available in the team.

**Cultural impediments and enablers.** However, in addition to being difficult if teams comprise only inexperienced controllers, it will also be problematic in cultures where openness and inquiry are inhibited. In masculinist cultures, for example, where learning is not part of work identity (Owen, 2009), simply changing policies to introduce multi-skilling in the way discussed above will not ensure collective development of expertise or the culture of conscious inquiry important in high-reliability work. This requires an open communicative climate where observations and reflections are shared. As discussed earlier, sometimes in can-do masculinist cultures such as that found in aviation, openness to inquiry can be limited. In the following transcript, for example, the controller inhibits the possibilities of increasing his own options for handling a problem, through refraining from asking a team member for help, because the norm of practice within his team does not support inquiry.

Interviewer: "Under what conditions would you ask?"

Respondent: "There have been times in the past when I've been training or I've had a question or something I couldn't quite understand and I think, '*Would he know? I'll ask.*' But '*No, I'll pick it up.*' And sometimes you do [pick it up] and sometimes you don't, until later [and] you think, '*oh God! Is that what they meant?*' And it can be little things... [But] '*I don't want to look stupid, I should know that...* Oh well, *I'll pick it up.*' I'd hoped! (Int. 40, Controller).

For this controller the process of intentional inquiry is hindered because the controller does not wish to reveal his lack of knowledge to another. In this case generating alternative courses of action was constrained because of a belief the controller held that "*he should know.*" Moreover, in the workplace under study, the divisions of labour of experienced controllers being concentrated in the Approach sectors and the inexperienced controllers concentrated in the en-route sectors also led to contested cultures, as the following controller explains in discussion about team working within a new flatter organizational hierarchy:

Interviewer: "So you don't like teams then?"

Respondent: "We're all supposed to be part of the same team, but if you've noticed, there's a huge division between the Tower and these guys over here on Approach and there's a huge division between Approach and the rest of the room. We are the prima donnas and we carry on like prima



donnas. A lot do, and now, everyone used to aspire to go onto Approach and that doesn't happen any more because of the salary compression (flatter hierarchy). So now, they find it difficult to get Approach controllers and now there's resentment from the Approach controllers to the people on the other side of The Room. Not the people individually but to the fact that they are getting paid as much as we are in some cases more and they're just out of nappies and it really has got up a lot of guy's noses. So people on the other side of The Room often get the cold shoulder and they [get] *'Go and ask your team leader'* when they're asked a question or something like that. *'You're getting paid as much as I am, go find out for yourself'*" (Int. 43, Controller).

The quote above highlights some of the impacts of organizational change on opportunities for informal learning because such changes have disrupted certain organizational cultures. This has led to resentment on the part of some group members, with negative impacts for informal learning.

### **Transparency and reciprocity: Enabling collective expertise**

The argument presented here supports Engestrom's (2004) contention that we need to rethink how expertise develops. In this paper, this is held to be important and interwoven within everyday work practice. The skills of collaborative expertise postulated by Engestrom need to be further developed so that networks of relations within organizations and their divisions of labour can support collaborative problem solving and in so doing, can raise the collective level of expertise within the professional community. Those skills include transparency and reciprocity. As this paper shows workplace learning facilitators and organizational designers need to scrutinize organizational cultures and structures to evaluate the ways in which these mediate the capacity for communicative practices to be transparent and reciprocal and thus to support continuous learning and conscious inquiry.

### **Conclusion**

At the commencement of this paper, we argued that understanding expertise needed to be revised to provide a more satisfactory account of how expertise is developed in dynamic and changing work contexts. On the one hand, traditional views of expertise have been built around three assumptions: First, that expertise is developed from a stable knowledge base; second, that it is acquired individually and internalized; and third, that it is extended unidirectionally from expert to novice.

In contemporary workplaces, each of these assumptions needs to be revised and expanded. Expertise does require building a depth of knowledge that needs to be undertaken collectively if it is to be sustained. Job content knowledge will be gained and shared through the process of learning for both the experienced other as well as the newcomer and developed dialogical inquiry. However, this will be impeded if workplace cultures do not value openness and inquiry.

New forms of expertise and of the collective production of knowledge are particularly important in high reliability workplaces as well as ones undergoing change. "They have to invent a system of reciprocal learning without which there can be no collective production of knowledge in an innovation-intensive context" (Hatchuel, Le Masson, & Weil, 2002, p.26).



Developing and improving the community of practice by “creating active and supportive climates for experiential learning ensures that experiential learning will be shared among team members that the practice community will be self improving, and that system redesign will be ongoing” (Benner, 2004, p.293). However, this will only occur in cultures that already support inquiry and learning.

If we are to enhance expertise as a continuous process, which is embedded in and supported by such communities, we need to deconstruct the value laden cultural understandings that may impede collective learning. Teasing out the tacit and implied understandings that enable or constrain learning contributes to the enhancement of learning as a continuous and iterative process, which is intertwined with and supported by organizational structures.

Moreover, individualistic approaches cannot fully account for enhanced learning in workplaces. A gap in theory exists, which fails to fully consider the context where such individuals learn and act and the reciprocal nature of the collective experiences afforded by this context, both formal and informal.

Increasing importance should be given to the high levels of skill needed in occupations of the future (Engestrom, 2004). The development of technologies which automate less complex jobs leads to the “rapidly increasing need to train students to even higher levels of expertise to continue the development of our modern society” (Ericsson 2006, p.14). The instructor role for controllers might be seen as an unintentional “innovative deviation” of the expected course of work action, enriching learning proximally, reconfiguring old knowledge by reflection on practice, and re-embedding the expert practitioner in the historical context of his own journey as learner.

Organizations of the future need to evaluate all aspects of work practice and identify strategies to systemize and enhance learning. While it is acknowledged that organizational flexibility has benefits in enabling competent handling of uncertainty and thus enhancing resilience (Hollnagel, Woods and Leveson, 2006), there remains a current belief that flexibility and change also carry risks of system failure (Grote et al., 2008; Amalberti 1999). A high level of standardization in the design of high-risk organizations such as air traffic control can be the result. How then do we attain a balance between the formal control needed to ensure safety in high-risk organizations and the flexibility, which allows the growth of expertise in order to safely engage with such uncertainty and dynamic change?

The challenges of enhancing safety outcomes in high reliability organizations have never been greater. In the context of dynamic change, systems, which do not support flexible responses, may be inherently flawed. The interrogation of organizational culture offers the key to determining how dynamic responses might be encoded in organizational standards and routines as an implicit affordance, without compromising, but rather enhancing safety culture. The results of this study suggest that opportunities for enabling the development of expertise need to be reexamined in contemporary work contexts as they may emerge from a variety of unintended sources. Better understanding these sources as well as how they are mediated by socio-cultural contexts will enable new opportunities for learning and the development of expertise to be encoded into organizational structures and cultures.

## References

- Amalberti, R. (1999) Risk Management by Regulation. Paper presented at the 19<sup>th</sup> Myron B Laver *International postgraduates course risk management*, Department of Anesthesia, University of Basel, Switzerland, March 26-27.
- Benner, P., Tanner, C., & Chesla, C. (1996). *Expertise in nursing practice: caring, clinical judgement and ethics*. New York: Springer.
- Benner, P. (2001). Creating a culture of safety and improvement: a key to reducing medical error. *American Journal of Critical Care*, 10(4), 281-284.
- Benner, P. (2004). Using the Dreyfus model of skill acquisition to describe and interpret skill acquisition and clinical judgement in nursing practice and education. *Bulletin of Science Technology and Society*, 24(3), 188-199. doi: 10.1177/0270467604265061
- Benner, P. (2005). Honouring the good behind rights and justice in health care when more than justice is needed. *American Journal of Critical Care*, 14(2), 152-156. Retrieved from <http://ajcc.aacnjournals.org/cgi/content/full/14/2/152>.
- Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. Resnick (Ed), *Essays in honour of Robert Glaser* (pp.361-363). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brookfield, S. (1995). *Becoming a critically reflective Teacher*. San Francisco: Jossey-Bass.
- Burrell, G. & Morgan, G. (1979). *Sociological paradigms and organizational analysis*. London: Heinemann Educational Books.
- Carrington, G. (2004). Supervision as a Reciprocal Learning Practice. *Educational Psychology in Practice*, 20(1), 31-42. doi: 10.1080/0266736042000180393
- Charmas, K. (2006). *Constructing grounded theory* London: Sage.
- Chi, M., Glaser, R., & Farr, M. (1988). *The nature of expertise*. Hillsdale, New York: Lawrence Erlbaum Publishers.
- Collin, K. (2006). Connecting work and learning: design engineers' learning at work. *Journal of Workplace Learning*, 18(7-8), 403-413. doi: 10.1108/13665620610692971
- Collinson, D. L. (1992). *Managing the shopfloor: subjectivity, masculinity and workplace culture*. Berlin: Walter de Gruyter & Co.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research*, London: Sage
- Denzin, N. & Lincoln, Y. (2000). (Eds.), *Handbook of qualitative research*. Thousand Oaks, CA: Sage.
- Engestrom, E., & Middleton, D. (1996). *Cognition and communication at work*. Cambridge, New York, Melbourne: Cambridge University Press.
- Engestrom, Y. (2001). Expansive learning at work: toward an activity theoretical reconceptualization. *Journal of Education and Work*, 14(1), 133-156. doi: 10.1080/13639080020028747

- Engestrom, Y. (2004). The new generation of expertise. In H. Rainbird, A. Fuller & A. Munro (Eds), *Workplace learning in context* (pp. 145-165). London, New York: Routledge.
- Eraut, M., Alderton, J., Gerald, C., & Senker, P. (1998). *Learning from other people at work*. Bristol: The Policy Press.
- Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.
- Fogarty, G., & Buikstra, E. (2008). A test of direct and indirect pathways linking safety climate, psychological health, and unsafe behaviours. *International Journal of Applied Aviation Studies*, 8(2), 199-207. Retrieved from [http://www.faa.gov/about/office\\_org/headquarters\\_offices/arc/programs/academy/journal/pdf/Winter\\_2008.pdf](http://www.faa.gov/about/office_org/headquarters_offices/arc/programs/academy/journal/pdf/Winter_2008.pdf)
- Garfinkel, H. (2007). *Studies in ethnomethodology: revised and updated*, Malden MA: Paradigm Publishers.
- Gherardi, S. (2001). From organizational learning to practice-based knowing. *Human Relations*, 54(1), 131-139. doi: 10.1177/135050840072001
- Gillett, G. (1995). The philosophical foundations of qualitative psychology. *The Psychologist* (March), 109-114.
- Grote, G., Weichbrodt, J. C., Gunter, H., Zala-Mezo, E., & Kunzle, B. (2008). Coordination in high-risk organizations: the need for flexible routines. *Cognition, Technology and Work*, 11, 17-27. doi: 10.1007/s10111-008-0119-y
- Hassard, J. (1990). Ethnomethodology and organizational research: An introduction. In J. Hassard & D. Pym (Eds.), *The theory and philosophy of organizations*. London: Routledge.
- Hatchuel, A., Le Masson, P., & Weil, B. (2002). From knowledge management to design oriented organizations. *International Social Science Journal*, 54(1), 25.
- Hollnagel, E. Woods, D., Leveson, N (eds) (2006) *Resilience Engineering: concepts and precepts*. Ashgate, London
- Hoover, A. (2008). Educational learning theories: Informing the fundamentals of Instruction. *International Journal of Applied Aviation Studies*, 8(2), 363-371.
- Hubbard, T. P. (2008). Using probability and set theory to examine illustrations of situation awareness. *International Journal of Applied Aviation Studies*, 8(1), 36-56.
- Hudson, P. (2007). Implementing a safety culture in a major multi-national. *Safety Science*, 45, 697-722.
- Konkola, R., Toumi-Grohn, T., Lambert, P., & Ludvigsen, S. (2007). Promoting learning and transfer between school and work. *Journal of Education and Work*, 20(3), 211-228. doi: 10.1080/13639080701464483
- Lave, J. (1993). Situating learning in communities of practice. In L. Resnick, J. Levine & S. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 63-82). Washington DC: American Psychological Association.

- Lave, J., & Wenger, E. (2005). Practice, person, social world. In H. Daniels (Ed.), *An introduction to Vygotsky* (pp. 149-156). Hove, New York: Routledge.
- Marginson, S (2005) Education and human capital, in P. Saunders & J. Walter (eds.) *Ideas and Influence: Social Science and Public Policy in Australia*, Sydney, NSW Press, pp. 64-84
- Mitchell, J. (2009). Flying Glass: A Qualitative Analysis of Pilot Perceptions of Automated Flight-Decks After 20 Years. *International Journal of Applied Aviation Studies*. 9(1), 13-27.
- Owen, C. (2009). Instructor beliefs and their mediation of instructor strategies. *Journal of Workplace Learning*, 21(6), 477-495. doi: 10.1108/13665620910976757
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117-175.
- Patton, M. (2002). *Evaluation and research methods*. Newbury Park: Sage.
- Pillay, H. K., & McCrindle, A. R. (2005). Distributed and relative nature of professional expertise. *Studies in Continuing Education*, 27(1), 67-88.
- Schon, D. A. (1991). *The reflective turn*. New York: Teachers College Press.
- Schulz, K.-P. (2005). Learning in complex organizations as practicing and reflecting: a model development and application from a theory of practice perspective. *Journal of Workplace Learning*, 17(7/8), 493-507.
- Suchman, L. (1996). Constituting shared workspaces. In Y. Engestrom & D. Middleton (Eds.), *Cognition and communication at work* (pp. 35-60). Cambridge, New York, Melbourne: Cambridge University Press.
- Strauss, A. & Corbin, J. (1994). Grounded theory methodology: An overview. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research*. Thousand Oaks, CA: Sage.
- Tesch, R. (1990). The mechanics of interpretational qualitative analysis. In R. Tesch (Ed.), *Qualitative research: analysis types & software tools* (pp. 113-127). New York: Falmer Press.
- Van Maanen, J. (1988). *Tales of the field: On writing ethnography*. Chicago: Chicago University Press.
- Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.
- Weick, K. E. (2001). *Making sense of the organization*. Oxford, Massachusetts: Blackwell Publishers Inc.
- Westrum, R. (1997). Social factors in safety-critical systems. In F. Redmill & J. A. Rajan (Eds.), *Human factors in safety-critical systems* (pp. 233-256): Butterworth/Heinemann.
- Yin, R. (2003). *Case study research design and methods*, 3<sup>rd</sup> Ed. Thousand Oaks, London: Sage.
- Zimmerman, B. J. (2006). Development and adaptation of expertise: the role of self-regulatory processes and beliefs. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise*.

## **APPENDIX A**

### **The Research Questions Guiding the Study**

The overall research question was “in what ways do structures and cultures enhance and inhibit learning in the workplace?” The following research sub-questions were also addressed in this study.

- What structures and cultures can be identified within the workplace studied?
- In what ways do organizational structures and cultures enhance and inhibit learning in the workplace?
- In what ways do organizational changes, such as the introduction of complex technology, influence workplace learning?
- How might workplaces be designed to create possibilities for practices of continuous learning and the development of educative work environments?

## APPENDIX B

### Examination of Work-Based Learning and Instruction

1. How long have you been in your existing position (e.g., OJTI)?
2. If Instructor - Why did you become one?
3. How many trainees have you had? Over what time?
4. I am trying to develop an understanding of how on-the-job training occurs in air traffic control. Think back to your last trainee (or current one). Can you describe what happened: For example,
  - (a) When did you “take on” this person? Did you have a choice? How did you find out you were getting a trainee? Was there any contact (between you and the trainee) beforehand? What did you know about this person before they arrived?
  - (b) How did their training progress? Were there any times when they progressed quickly or slowly? Why do you think that was the case?
  - (c) What indicators did you use to know that the trainee was progressing? (i.e., What did you look for?). What kind of written documentation was kept? Where might I find it?
  - (d) Was the trainee or any aspect of their learning difficult? Why do you think this was?
  - (e) Throughout the trainee’s learning period, what contact did you have with - your team leader; the training annexe; others regarding the trainee? Can you describe examples of such contact/discussions? (Probe purpose, who initiated, resolutions).
  - (f) Is what you have described a typical OJTI experience or is this one unusual? Why?

## ***A Two-Group Experiment to Measure Simulator-Based Upset Recovery Training Transfer***

Rodney O. Rogers

*Department of Aeronautical Science*  
[rogers\[at\]erau.edu](mailto:rogers[at]erau.edu)

Albert Boquet

*Chair, Department of Human Factors*  
[boque007\[at\]erau.edu](mailto:boque007[at]erau.edu)

Cass Howell

*Chair, Department of Aeronautical Science*  
[howellc\[at\]erau.edu](mailto:howellc[at]erau.edu)

*Embry-Riddle Aeronautical University*  
*Daytona Beach, Florida 32114*  
*USA*

Charles DeJohn

*Federal Aviation Administration*  
*Civil Aerospace Medical Institute*  
*Aerospace Medical Research Team*  
*USA*  
[charles.dejohn\[at\]faa.gov](mailto:charles.dejohn[at]faa.gov)

### *Abstract*

*Although air transport upset recovery training is typically simulator-based, there is little evidence to suggest that such training actually improves a pilot's ability to recover an actual airplane from a serious upset. We report on a research experiment to evaluate simulator-based upset recovery training transfer. Participants were trained in upset recovery maneuvering using low-cost desktop flight simulation, and then subjected to serious in-flight upsets in a general aviation aerobatic airplane. Their performance in upset-recovery maneuvering was compared with the performance of untrained pilots subjected to the same upsets. Statistical analysis of data collected during flight testing suggests that simulator-based training combined with classroom instruction improves a pilot's ability to recover an airplane from an upset. We summarize prior related research, describe the experiment, analyze and interpret flight-test data, and explain the implications of our research with respect to federally mandated upset recovery training requirements.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents



## **A Two-group Experiment to Measure Simulator-based Upset Recovery Training Transfer**

An upset occurs when an airplane enters an unexpected attitude that threatens loss of control (LOC) and subsequent ground impact. It is well known that from 1991 onward LOC has been the major cause of air transport accidents worldwide and of general aviation airplane accidents in the US and Australia (Rogers, Boquet, Howell, & DeJohn, 2009). What follows describes an FAA funded research experiment to evaluate transfer of upset-recovery training conducted using low-cost flight simulation. We assessed training effectiveness by means of in-flight upset-recovery testing in a general aviation airplane. In what follows, we

1. Summarize relevant prior research.
2. Describe the experiment.
3. Present and interpret the experimental results.
4. Discuss the importance of good training.
5. Explain what our research implies about federally mandated upset recovery training.

### **Prior Research<sup>1</sup>**

We have found only a few research articles related to the transfer of simulator-based upset-recovery training. Several reports result from research at the Calspan In-Flight Upset Recovery Training Program in Roswell, N. M. A second set of articles discusses human factors considerations in upset-recovery training. A third group summarizes two articles related to training transfer when upset maneuvering is taught using low-cost flight simulation.

### **Calspan Related Research**

Calspan provides in-flight simulator-based upset-recovery training in a variable stability Learjet 25 modified to simulate the control characteristics of an air transport airplane. The Calspan Lear can simulate various accident scenarios that in the past have resulted in air transport upsets leading to uncontrolled crashes.

Gawron (2004) used Calspan's Learjet to test five groups of airline pilots with varying degrees of upset-recovery training and/or aerobatics experience on a series of eight upsets, hypothesizing that pilots with more training and/or experience would outperform those with less. However, she found no significant difference among the performances of the five groups.

Kochan (2005) used the Calspan Lear to examine the roles of *domain knowledge* and *judgment* in upset-recovery proficiency. Domain knowledge is specific knowledge about upset-recovery procedures. Judgment is the ability to analyze

---

<sup>1</sup> Subsections 1.1 and 1.2 are reproduced verbatim from Rogers *et al.* (2009), 1-2. A more detailed description of the same articles may be found in Rogers, Boquet, Howell, and DeJohn (2007), 2-4.

and learn from an in-flight upset-recovery experience. She tested four groups of participants on a series of three in-flight upsets. Statistical analyses revealed that judgment was a significant factor in successful upset recovery, especially when a pilot has low domain knowledge, i.e., when a pilot is not trained to proficiency in upset recovery.

Kochan and Priest (2005) studied the effect of upset-recovery training in the Lear. They measured pre- and post-training pilot performance in recovering from a series of upsets. Statistical analysis indicated “a strong positive influence of the [Calspan program] on a pilot’s ability to respond to an in-flight upset.”

Kochan, Breiter, Hilscher, and Priest (2005) surveyed retention of knowledge in Calspan trained pilots. Although participant in retrospect “rated their ability to recover from loss-of-control situations as being greatly improved by the training,” most were unable to recall various specific details about upset-recovery maneuvering taught during their training.

### **Human Factors Considerations in Upset-Recovery Training**

A number of papers examine the “surprise” or “startle” factor in aviation, an effect that can hinder a pilot’s ability to respond appropriately to an emergency situation such as an upset. Kochan, Breiter, and Jentsch (2004), found pilots often miss cues that might lead to avoiding an emergency that later arrives as a surprise. In a follow-on paper (Kochan, Breiter, & Jentsch, 2005), the researchers develop “a conceptual framework for the study of unexpected events in aviation. Kochan, Priest, and Moskal (2005a, 2005b) use a model for the “cognitive process of surprise” [Based on Kochan (2005)] to study “how an unexpected event can escalate to a loss-of-control situation.” They conclude that in-flight [as opposed to ground-based] simulator training may be necessary to teach pilots to deal adequately with their perceptual *biases* in processing information during a surprise upset. In a related paper, Kochan (2006) argues that a pilot’s response to unexpected events can be improved through *cognitive flexibility training* (to discourage formulaic and encourage flexible responses to surprise events), *adaptive expertise training* (to reinforce modified or new responses to surprise based on responses learned in previous expert training), and *metacognitive training* (to teach pilots how to evaluate their mental processes in responding to surprise).

### **Low-Cost Simulation**

Roessingh (2005) studied training transfer from low-fidelity ground-based flight simulators to control of an actual airplane during aerobatic flight. Two experimental groups received ground-based instruction in aerobatic maneuvering using desktop flight simulators. The simulator syllabus was the same for both groups, but one experimental group’s simulator training was enhanced with a more “realistic layout of stick, rudder pedals, and throttle.” Then the two experimental groups and a control group received five hours of in-flight aerobatic training. Data collected during subsequent testing revealed no significant difference in the aerobatic maneuvering of experimental and control group pilots.

In a predecessor to the experiment described herein, Rogers, Boquet, Howell, & DeJohn (2007) trained participants in upset recovery using low-cost desktop

simulation, and then subjected them to a series of serious upsets in an aerobatic Beech Bonanza airplane. Trained participants outperformed untrained control group participants in a variety of dependent variable such as thrust manipulation, G force control, and roll responses; however, in the most important discriminator—altitude loss—there was no significant difference between the two groups. The authors argued that shortcomings in training and testing procedures negatively influenced the experimental results, conjecturing that increased training transfer would result if the experiment were repeated with improved approaches to training and testing.

## The Research Experiment

### Experiment Design

Our research hypothesized that upset-recovery training in low-cost desktop flight simulators develops flying skills that improve a pilot's ability to recover a real airplane from a serious upset in a real airplane. To test this hypothesis, we trained a group of participant pilots in upset-recovery maneuvering using Microsoft Flight Simulator (MFS). Then we subjected them to a series of four upset situations in an aerobatic airplane and collected data on their performance in recovering the airplane to straight and level flight. We also subjected a group of untrained participants to the same series of four upsets. Participants in our experiment were self-selected student pilots at Embry-Riddle Aeronautical University (ERAU) in Daytona Beach, Florida. All held a current instrument rating (implying a private pilot certificate as a minimum) and completed an academic course in basic aerodynamics for pilots. None has prior aerobatic experience or upset-recovery training beyond that required for FAA flight certificates and ratings.

As reflected in Table 1, our experiment is a 2 x 4 repeated measures factorial. The first independent variable is degree of training and has two levels—trained and untrained. Trained participants received ten hours of classroom and ten hours of MFS upset-recovery training.<sup>2</sup> Untrained participants—control group pilots—received no classroom or simulator training. The second independent variable is upset attitude. It has four levels corresponding to the four upsets each participant is subjected to during flight testing.<sup>3</sup>

### Dependent Variables

We defined a good upset recovery as one where a pilot respects aircraft operating limitations while returning the aircraft to straight and level flight with the minimum possible loss of altitude. Minimum altitude loss will result from:

- Prompt and correct control and throttle inputs in response to an upset situation.
- A high roll rate toward an upright attitude to orient the lift vector toward the sky.

---

<sup>2</sup> Training materials may be viewed at <http://faculty.erau.edu/rogersr/as471>.

<sup>3</sup> The upsets were ordered from least to most difficult to present the same learning curve to each participant. We had hoped to be able to demonstrate that trained pilots learned more quickly than untrained pilots, but the data did not support this conclusion.

- Use of appropriate G forces (unloaded during low-speed or inverted rolls; high Gs in dive pullouts while avoiding accelerated stalls).

The dependent variables in our experiment—shown in Table 2—are designed to measure these factors.

Table 1

*The 2 x 4 Factorial Design*

2 x 4 Factorial		Upset Attitude (Repeated Measure)			
		Nose-high Upright	Nose-low Upright	Nose-high Inverted	Nose-low Inverted
Training	10 Hours Classroom/ Simulator (Trained Group)	Trained	Trained	Trained	Trained
	None (Control Group)	Untrained	Untrained	Untrained	Untrained

Table 2

*Dependent Variables*

Dependent Measure
Altitude Loss in Feet: Negative Value = Altitude Gain
Maximum G Force in Dive Pullout
Minimum G Force Unloading during Rolls
Time to First Throttle Response in Seconds
Time to First Roll Response in Seconds
Time to Recover in Seconds

**Upset Attitudes**

We categorized upset attitudes as nose-high or nose-low and as upright or inverted. An inverted attitude is one where the bank angle exceeds 90°. Table 3 specifies specific the attitude, thrust setting, and kinetic energy level for each of the four upsets. Nose-high initial airspeeds were set 12 MPH above  $V_S$  for the Decathlon, while nose lose airspeeds reflect a maximum safe value based on the Decathlon's red line speed  $V_{NE}$  of 200 mph. The 180° roll attitude for inverted upsets was chosen because it simplified the demanding safety pilot task of positioning the aircraft accurately from the Decathlon back seat, which has no instrumentation. Accurate positioning was critical to the success of the experiment because other than small deviations from a prescribed upset attitude or energy level significantly affect the potential minimum altitude loss for an upset.

Table 3

*Levels of the Upset Attitude Independent Variable*

Upset	Pitch	Bank	Airspeed	Thrust
Nose-high Upright	60° Nose-high	45° Left Wing Down	65 MPH	Idle
Nose-low Upright	45° Nose-low	70° Right Wing Down	130 MPH	Full
Nose-high Inverted	60° Nose-high	180° (Inverted, Wings Level)	65 MPH	Idle
Nose-low Inverted	20° Nose-low	180° (Inverted, Wings Level)	110 KIAS	Full

**Data Collection**

To collect data, we installed a battery-operated video camera focused on the Decathlon's instrument panel. A high-resolution palm-size video recorder captured the camera's output and cockpit voice communications. Two factors prevented our installing a more sophisticated data recording system. One was the significant cost. The other is a prohibition against invasive instrumentation in an Embry-Riddle training aircraft. Figure 1 presents a screen capture of a video recorded during flight testing. We also installed an Appareo GAU-1000 AHARS data recorder, an inexpensive battery operated GPS-based system capable of recording aircraft position, altitude, airspeed, attitude (pitch and bank), G forces (x, y, and z), yaw angles ( $\beta$ ), and similar parameters. However, only G force data from it proved reliable in aerobatic attitudes.



Figure 1. Sample Decathlon Video Recorder Output

## Results

### Data

We collected complete data sets for 25 trained participants and for 26 control group participants. Average flight time for trained and control group participants respectively was 201.2 and 160.5 hours. Six trained pilots and eight control group pilots experienced unsuccessful recoveries during the nose-low inverted upset. In every case, the safety pilot took control in dive pullout to avoid exceeding the Decathlon's redline speed  $V_{NE}$ . We excluded data for these 14 upsets from our statistical analysis because they reflect safety pilot input to the airplane's flight controls. We also failed to obtain data for one trained participant for the nose-high inverted upset due to air sickness.

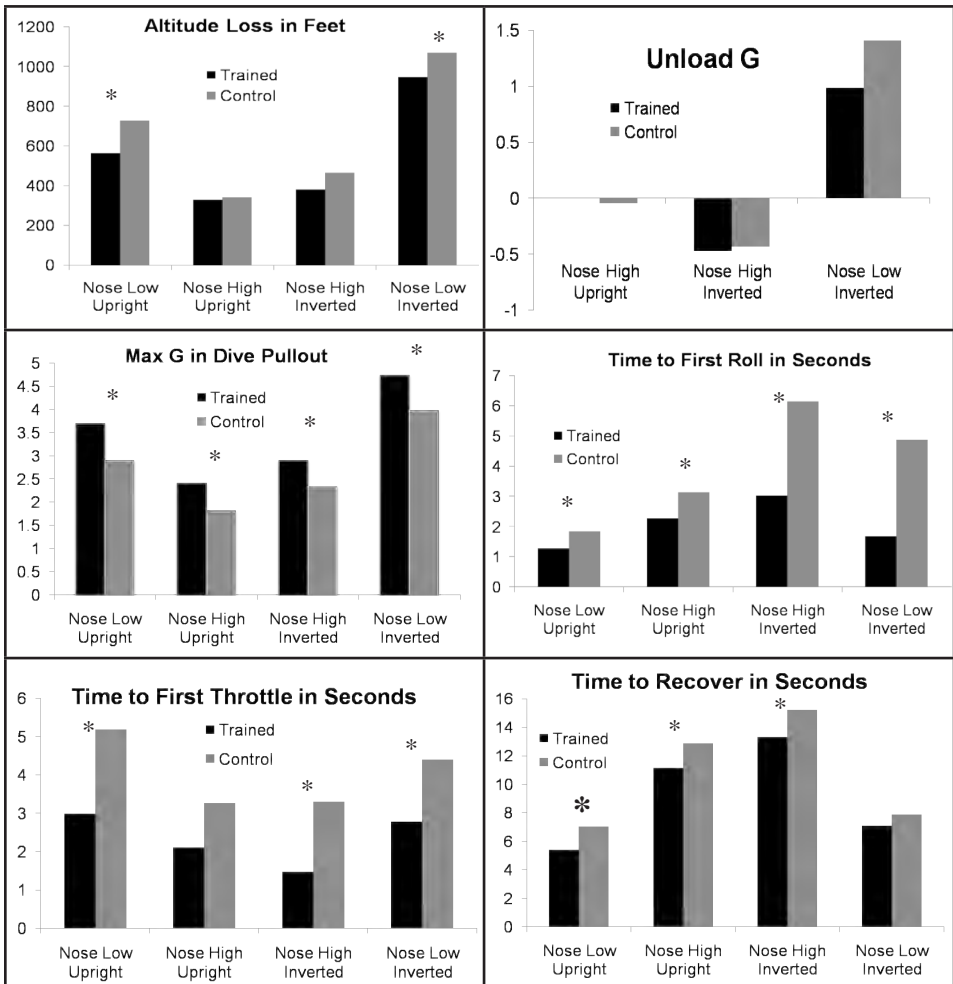


Figure 2. Results in Graphical Format (\* = Significant Effect)

For each of the four upsets, Table 4 presents usable-data averages and *standard deviations* for the six dependent variables. **Bold** indicates a significant difference between groups as reported in Subsection 3.2. Figure 2 presents the data from Table 4 in graphical format.

Table 4.

*Dependent Measures Means and Standard Deviations (Bold = Significant Difference)*

Upset	Nose-Low Upright		Nose-High Upright		Nose-Low Inverted		Nose-High Inverted	
	Trained	Control	Trained	Control	Trained	Control	Trained	Control
Altitude Loss In Feet	<b>565.20</b>	<b>728.46</b>	331.20	340.38	<b>948.95</b>	<b>1069.41</b>	382.08	464.62
	<b>75.28</b>	<b>169.51</b>	225.56	184.75	<b>167.03</b>	<b>139.08</b>	200.65	169.59
Min Unload G in Rolls	Not Applicable †		0.00	-0.04	0.99	1.41	-0.47	-0.43
			0.12	0.15	0.86	.63	.28	.26
Max G in Dive Pullout	<b>3.70</b>	<b>2.90</b>	<b>2.41</b>	<b>1.82</b>	<b>4.74</b>	<b>3.98</b>	<b>2.90</b>	<b>2.34</b>
	<b>0.64</b>	<b>0.49</b>	<b>0.90</b>	<b>0.30</b>	<b>0.62</b>	<b>0.50</b>	<b>0.84</b>	<b>0.45</b>
Seconds to First Throttle	<b>3.0</b>	<b>5.19</b>	2.12	3.27	<b>2.79</b>	<b>4.41</b>	<b>1.48</b>	<b>3.31</b>
	<b>1.66</b>	<b>2.43</b>	1.62	2.97	<b>1.78</b>	<b>2.81</b>	<b>.68</b>	<b>3.21</b>
Seconds to First Roll	<b>1.28</b>	<b>1.85</b>	<b>2.28</b>	<b>3.15</b>	<b>1.68</b>	<b>4.88</b>	<b>3.04</b>	<b>6.15</b>
	<b>.46</b>	<b>.68</b>	<b>.89</b>	<b>1.38</b>	<b>.67</b>	<b>3.30</b>	<b>1.30</b>	<b>2.98</b>
Seconds to Recover	<b>5.40</b>	<b>7.04</b>	<b>11.16</b>	<b>12.88</b>	7.11	7.88	<b>13.33</b>	<b>15.23</b>
	<b>1.38</b>	<b>1.64</b>	<b>1.43</b>	<b>2.98</b>	1.29	.99	<b>1.74</b>	<b>2.27</b>

† Not applicable because trained pilots were taught to use rolling pullouts from upright dives.

## Analysis

To compare trained and control group pilot performance, we conducted one-way MANOVAs for each of the four upsets using the dependent measures in Table 4. Two factors motivated our decisions to forego a more traditional 2 x 4 mixed-model analysis. First, because we eliminated data from unsuccessfully recoveries in the nose-low inverted upset, a mixed-model analysis would have substantially reduced the sample size. Second, the nature of the upset data themselves argues against the direct comparisons that characterize repeated measures MANOVA. For example, a nose-high recovery may lead to an altitude gain whereas nose-low recoveries invariably result in significant altitude losses. Rather than compare “apples to oranges,” we opted for a more direct and operationally more relevant approach to data analysis. The Wilks’ Lambda values resulting from the MANOVAs are shown in Table 5; they reflect a significant difference between the two groups in each of the four upsets.<sup>4</sup>

<sup>4</sup> While we understand that performing multiple one-way analyses increases the family-wise error rate, the low computed alphas together with the magnitude of effect for each of the analyses provides confidence in our results while maintaining acceptable type 1 risks (below .05).



Table 5

*Multivariate Wilks' Lambda Values and Group Sizes for Each Upset*

Group Size	Upset			
	Nose-Low Upright	Nose-High Upright	Nose-Low Inverted	Nose-High Inverted
Trained	n=25	n=25	n=19	n=24
Control	n=26	n=26	n=17	n=26
Combined	n=51	n=51	n=36	n=50
Wilks' Lambda Value	F (5,45) =9.59 p = .0001 $\eta^2 = .052$	F(6,44) = 4.47 p = .001 $\eta^2 = .38$	F (6,29) =9.11 p = .0001 $\eta^2 = .653$	F (6,43) =10.26 p = .0001 $\eta^2 = .60$

Since data for all four upsets indicated significant differences between trained and control group pilot performance, we then conducted ANOVAs for each of the paired dependent variables in Table 4. Table 6 presents F values associated with significant effects indicated by **Bold Text** in Table 4. Each entry in Table 6 reveals superior performance by trained pilots compared to control group pilots.

Statistical analysis confirms our hypothesis that low-cost simulator-based upset-recovery training improves pilot performance in recovering an airplane from a serious upset. Trained pilots lost less altitude than control group pilots because they initiated rolls toward a wings level upright attitude sooner and applied more Gs in dive pullouts than untrained pilots, both critical factors in minimizing altitude loss. In addition, trained pilots also applied throttle more promptly than untrained pilots. These differences in turn resulted in a quicker return to straight and level flight. G unloading in rolls was the only dependent measure where trained and control group performance did not differ statistically. Excluding altitude loss, trained pilots were statistically superior to control group pilots in 14 of 19 categories, or 73.7 % of the time. Including altitude loss, trained pilot performance exceeded untrained pilot performance in 16 of 23 categories, or 69.6% of the time.

Table 6

Significant Effect *F* Values as Determined by Univariate Analysis

Dependent Measure	Upset			
	Nose-Low Upright	Nose-High Upright	Nose-Low Inverted	Nose-High Inverted
Altitude Loss	F(1,49) = 19.48 p = .0001		F(1,34) = 5.45 p = .03	
Minimum Unload G in Rolls	Not Applicable †			
Maximum G Load in Dive Pullout	F(1,49) = 25.52 p = .0001	F(1,49) = 10.11 p = .003	F(1,34) = 16.02 p = .0001	F(1, 48) = 8.912 p = .004
Seconds to First Throttle	F(1,49) = 14.02 p = .0001		F(1,34) = 4.38 p = .04	F(1,48) = 7.46 p = .009
Seconds to First Roll	F(1,49) = 12.19 p = .001	F(1,49) = 7.18 p = .01	F(1,34) = 17.16 p = .0001	F(1,48) = 22.29 p = .0001
Seconds to Recover	F(1,49) = 14.82 p = .0001	F(1,49) = 6.83 p = .012		F(1,48) = 10.90 p = .002

† Not applicable because trained pilots were taught to use rolling pullouts from upright dives.

### An Unanswered Question

The training transfer we observed resulted from both classroom and simulator-based training. An open question is how much classroom training affected trained pilot upset maneuvering performance as opposed to simulator-based training. It would be interesting to repeat our experiment with the simulator training omitted. We believe—non-intuitively, perhaps, but based on our experiences training two sets of participants—that pilots taught upset recovery in the classroom only might still outperform untrained pilots in upset maneuvering. We currently are searching for money to fund such an expansion of our experiment.

### Importance of Good Training Methodology

Our experiment significantly improves the results of an earlier iteration of the same experiment, reported in the article described at the end of Subsection 1.3. We attribute the improvement in large part to better training methods learned through hard experience during the earlier experiment.

For example, in the first experiment we trained participants to perform the “canonical” knife-edge high bank angle recovery from nose high upsets, only to learn in flight testing that this recovery does not result in minimum altitude loss in airplanes like the Decathlon or the aerobatic Beech Bonanza. In high thrust-to-weight ratio airplanes such as these, far less altitude loss results from unloading the airplane at full thrust while using full aileron and rudder to roll wings-level upright as the nose falls toward above the horizon.<sup>5</sup> Done correctly, this Immelman-like maneuver results in the airplane’s nose reaching the horizon wings-level at an airspeed still well below  $V_{S_0}$ . At that point, it remains merely to continue allowing the nose to fall at 0 G until flying airspeed is regained, at which point application of a low G force will bring the nose smoothly to the horizon with a small altitude loss. Occasionally, in fact, the maneuver results in an overall altitude gain, since the aircraft is climbing—trading airspeed for altitude—as long as the nose remains above the horizon. Trained pilot altitude loss improved considerably when we taught this recovery in the research described herein.

We also learned how—and how not—to teach aircraft attitude analysis to novice pilots. In the initial experiment, we taught participants to use the Bonanza’s non-tumbling attitude indicator as the primary information source for categorizing upset attitudes. While this ability is indispensable for air transport pilots in flight where weather or darkness may obscure the horizon, perfecting it proved very difficult for inexperienced student pilots using only a simulator. For the research described in this paper, we taught participants to use VFR information only—front and side outside-the-cockpit views—to determine aircraft attitude. The result was quicker attitude analysis and a decreased time to initial throttle and control input responses, with a resultant improvement in altitude loss. Figure 3 depicts the MFS window configuration we used to teach pilots how to analyze aircraft attitude using out-the-window views. The two small windows bottom left and right are views outside the left and right windows of the Decathlon. (Figure 4 shows the window configuration we used to teach aerobatic maneuvering.)

MFS responses to control stick inputs are realistic near the middle of the Decathlon flight envelop ( $V$ - $n$  diagram). Near the envelop’s available G line [ $\alpha_{CRIT}$ ,  $(C_L)_{MAX}$ ], however, responses to control inputs tend to differ from the Decathlon’s behavior in actual flight situations. For example, if a Decathlon pilot inadvertently stalls the airplane with an aileron down, a departure from controlled flight may result. However, MFS does not simulate departures from controlled flight realistically. MFS also responds inaccurately during accelerated stalls in low-speed dive pullouts. To recover the actual airplane from such a stall, it suffices to momentarily relax back stick pressure a small amount, resulting in a slightly reduced G force. To recover the simulated Decathlon from the same accelerated stall situation, a pilot must unload completely to 0 Gs for a second or two before reapplying G to continue the dive pullout. In such scenarios, the potential for negative training is significant and must be countered by effective classroom and flight-simulator instruction about the limitations of the simulator’s aerodynamic model.

While we are aware of no research to support the assertion, we believe excellent upset recovery training given on a less sophisticated simulator is probably more effective than average training given on a more sophisticated simulator.

---

<sup>5</sup> The Decathlon will of course roll with aileron alone in this situation, but the roll rate is substantially increased if both aileron and rudder are used.



Figure 3. MFS Window Configuration for VFR Simulated Upset Recovery Maneuvering



Figure 4. MFS Window Configuration for Aerobatic Maneuvering

## Implications with Respect to Upset Recovery Training<sup>6</sup>

Our research establishes a statistical relationship between MFS-based upset-recovery training and improved all-attitude maneuvering skills in an actual airplane. However, it also seems to imply the limitations of low-cost flight simulation in teaching upset maneuvering. Consider, for example, the fact that trained pilots lost significantly less altitude than untrained pilots in both nose-low upsets but in neither nose-high upset. Why would recovery from nose-high upsets be more challenging to teach and more difficult to learn than recovery from nose-low? There are at least two answers.

First, MFS limitations may explain why trained pilots were less proficient in nose-high upset maneuvering than in nose-low. Nose-high low-kinetic-energy maneuvering occurs in the extended flight envelop where—we have seen—MFS control responses tend to be inaccurate, whereas maneuvering during high energy nose-low upsets occurs near the middle of the Decathlon flight envelop where MFS control responses are reasonably accurate. When training in a flight simulator that responds inaccurately to control inputs in such situations, it is difficult to prepare pilots to handle them adequately in a real airplane.

Second—and far more important, we think—general aviation pilots are inexperienced in nose-high low-kinetic-energy maneuvering. Required control inputs when maneuvering in high-air-speed nose-low dive recoveries differ in degree only from control inputs general aviation pilots routinely use in non-aerobatic upright flight. By contrast, during nose-high low-air-speed upset maneuvering, the proper control inputs differ in kind from what general aviation pilots are typically accustomed. As an example, efficient rolling in nose-high upsets requires a pilot to use large aileron and rudder inputs at maximum thrust while maintaining zero G at airspeeds approaching 0 mph. If significant positive or negative G is applied, a stall and departure from controlled flight is likely to result. However, using elevator to maintain 0 G while applying large rudder and aileron inputs to roll at very low airspeeds is something general aviation pilots never experience. Not surprisingly, then, imperfect participant pilot control inputs not infrequently resulted in a nose-high upset progressing into low-speed, steep nose-low upset. Whenever this occurred, accelerated stalls and departures from controlled flight were common during the low-speed dive pullout. Pulling out of a low-kinetic-energy steep dive at the stall buffet is also an experience unknown to general aviation pilots.<sup>7</sup>

Perhaps pilot behaviors necessary to perform nose-high, low-air-speed precision aerobatic maneuvers can be rehearsed using MFS and similar low-cost flight simulators but can only be perfected in a real airplane. More important, even when trained pilots significantly bettered control group pilots in altitude loss, their performance was far from optimal. For each of the four upsets, Table 7 presents Phase 2 average altitude losses for trained and control group pilots. The bottom row of Table 7 reflects the minimum altitude losses that we observed for each upset during safety pilot training. There is a large disparity between research participant altitude losses and the far smaller altitude losses achievable by experienced pilots.

---

<sup>6</sup> Adapted from Rogers *et al.* (2009), 11 and Rogers *et al.* (2007), 16.

<sup>7</sup> As it was to the pilots of Colgan Air Flight 3407 which crashed near Buffalo, NY on March 25, 1009 after an unsuccessful attempt to recovery from a what would appear to be a recoverable nose-low, low-air-speed upset.



Low-cost flight simulator training clearly improves a pilot's ability to recover an airplane from a serious upset. Just as clearly, however, it is prelude and complement to—not a substitute for—all-attitude maneuvering experience in a real airplane.

Table 7.

*Altitude Losses to Nearest Foot for the Four Upsets (Bold = Significant Effect)*

Data Source	Altitude Loss in Feet			
	Nose-Low Upright	Nose-High Upright	Nose-Low Inverted	Nose-High Inverted
Trained Pilot Average	<b>565</b>	331	<b>949</b>	382
Control Group Pilot Average	<b>728</b>	340	<b>1069</b>	465
Observed Minimums during Safety Pilot Training <sup>*</sup>	220	-50	465	-30

<sup>\*</sup> The values in the last row are not losses produced by optimal recoveries, a subject we have not pursued systematically. They merely reflect the minimum recorded altitude loss during safety pilot training with experienced pilots maneuvering the Decathlon out of an upset.

For altitude loss as well as other dependent variables, our data reflect only a modest difference in upset maneuvering performance between trained and untrained participants. It may well be the case that pilots need all-attitude flight experience in a real airplane to hone simulator-taught upset-recovery skills to an acceptably high level. It appears that during an initial upset-recovery experience, low-cost simulator training improves a pilot's ability to recovery only to a very limited extent. Simulator shortcomings—for example, unrealistic control feedback, inaccurate accelerated stall responses, and the inability to replicate the positive and negative G forces that characterize all-attitude flight—limit a trainer's ability to prepare a pilot mentally and emotionally for a real world upset. As a result, any subsequent initial experience in a real upsets may appear strange and disquieting. In such a circumstance, a pilot easily loses situational awareness and instinctively resorts to old control input habits. Long reinforced patterns of behavior and the significant stress of a serious upset tend to inhibit the application of new and relatively unfamiliar piloting skills developed during simulator-based training.

Our research findings seem to call in question the implicit assumption that airline simulator-based upset-recovery training programs impart flying skills sufficient to make it probable that typical line pilots can recover an airliner from a serious upset. It is true that airline pilots on average are considerably more experienced than our research participants, hence may benefit more from any kind of upset-recovery training. However, air transport pilots experience consists of hours of flying straight and level punctuated by occasional excursions into very small

bank and pitch angles.<sup>8</sup> Airline pilots typically receive only about four hours of classroom-based upset-recovery training and perhaps an hour of simulator training, in comparison to ten hours of each for our trained participants. Moreover, the primary advantage of a Level-D simulator over low-cost desktop flight simulation is limited to cockpit verisimilitude and realistic control forces. Typically, the motion of a Level-D simulator is disabled to avoid stressing the mechanism unnecessarily in upset-recovery maneuver. In any event, Level-D simulator motion is in no way realistic in all-attitude maneuvering, and neither Level-D nor low-cost desktop simulators can replicate the G forces that characterize upset-recovery maneuvering in a real airplane.

Thirty years ago, U.S. airline pilots typically came from military flight backgrounds where training afforded them extensive experience in all-attitude flight maneuvering. For these pilots, there were no *unusual* attitudes, only *unexpected* attitudes. By contrast, most air transport pilots flying today lack a military background and have never experienced the extreme pitch and bank angles and high G forces associated with severe airplane upsets. Indeed, most have never even been upside-down in an airplane. Informal conversations with current airline pilots suggest that while virtually all regard the upset training they receive as useful, a significant number also perceive it as a *pro forma* approach to a serious safety problem—better than nothing but far from what would be desirable if training costs were not a paramount consideration. In short, it seems unlikely that airline upset training is a completely acceptable substitute for upset-recovery maneuvering experience in a real airplane.

Upsets are known to be a primary cause of commercial air transport accidents. Passenger and aircrew safety considerations mandate that air transport pilots be able to recover from the infrequent but potentially catastrophic upsets, which inevitably will occur from time to time. Our research implies that all-attitude maneuvering experience in a real airplane may be required to make recovery probable with a reasonably small altitude loss. Since aerobatic experience cannot be obtained legally in transport type aircraft, perhaps the FAA should consider making aerobatic experience in a light airplane part of the requirement for a commercial pilot license. This requirement would ensure that air transport pilots would be better prepared to receive transport-type upset recovery maneuvering training in Level D simulators.

---

<sup>8</sup> The captain of Colgan Air Flight 3407 had 3379 hours of flight time and the copilot 2220 hours, but were still unable to recover their aircraft from a nose-down low-kinetic-energy upset.



## References

- Gawron, V. (2004). *Aircraft upset training evaluation report (revision 20)*. National Aeronautics and Space Administration (NASA). (NAS2-99070, *Airplane upset training evaluation report, February 25, 2002*. National Aeronautics and Space Administration (NASA) NAS2-99070 is an earlier version of this research report.)
- Kochan, J. A. (2005). *The role of domain expertise and judgment in dealing with unexpected events*. Ph.D. Dissertation, Department of Psychology, University of Central Florida, Orlando Florida.
- Kochan, J.A. (2006). Human factors aspects of unexpected events as precursors to unwanted outcomes. *18<sup>th</sup> Annual European Aviation Safety Seminar*. Flight Safety Foundation and European Regions Airline Association, Athens, Greece.
- Kochan, J.A., Breiter, E. G, and Jentsch, F. (2004). Surprise and unexpectedness in flying: Database reviews and analysis. *Proceedings of the Human Factors and Ergonomics Society 48<sup>th</sup> Annual Meeting*, New Orleans, LA.
- Kochan, J.A., Breiter, E. G., and Jentsch, F (2005). Surprise and unexpectedness in flying: Database factors and Features. *Proceedings of the 13<sup>th</sup> International Symposium on Aviation Psychology*, Oklahoma City, OK.
- Kochan, J. A., Breiter, E., Hilscher, M., and Priest, J. E. (2005). Pilots' perception and retention of in-flight upset recovery training: evidence for review and practice. *Proceedings of the Human Factors and Ergonomics Society 49<sup>th</sup> Annual Meeting*, Orlando, FL.
- Kochan, J.A. and Priest, J.E. (2005). Program update and prospects for in-flight upset recovery training. *Proceedings of the 13<sup>th</sup> International Symposium on Aviation Psychology*, Oklahoma City, OK.
- Kochan, J.A., Priest, J.E., and Moskal, M. (2005a). Human factors aspect of upset recovery training. *17<sup>th</sup> Annual European Aviation Safety Seminar*, Flight Safety Foundation and European Regions Airline Association, Warsaw, Poland.
- Kochan, J.A., Priest, J.E., and Moskal, M. (2005b). The application of human factors principles to upset recovery training. *50<sup>th</sup> Annual Corporate Aviation Safety Seminar*, Flight Safety Foundation and National Business Association, Orlando, FL.
- Roessingh, J. J. M. (2005). Transfer of manual flying Skills from PC-based simulation to actual flight—comparison of in-flight data and instructor ratings. *The International Journal of Aviation Psychology*, 15 (1), 67-90.
- Rogers, R.O., Boquet, A., Howell, C, and DeJohn, C (2009). An experiment to evaluate transfer of low-cost simulator-based upset-recovery training. *FAA Technical Report DOT/FAA/AM09/05*, Office of Aerospace Medicine, Washington, DC.
- Rogers, R.O., Boquet, A., Howell, C, and DeJohn, C (2007). Preliminary results of an experiment to evaluate transfer of low-cost, simulator-based airplane upset-recovery training. *FAA Technical Report DOT/FAA/AM07/28*, Office of Aerospace Medicine, Washington, DC.

## ***Effects of Video Weather Training Products on General Aviation Pilot Weather Knowledge and Flight Behavior into Adverse Weather***

William Knecht

And

Jerry Ball

*Civil Aerospace Medical Institute,  
Oklahoma City, OK*

### ***Abstract***

*This study examined whether brief video weather training products can significantly affect pilot weather knowledge and flight behavior in the face of potential instrument meteorological conditions (IMC). Fifty general aviation (GA) pilots comprised three groups. In Phase 1, two groups received a 90-minute video weather training product while one served as control. Weather knowledge was tested pre- and post-video. All pilots then flew a flight simulator mission featuring rising terrain with descending cloud bases. Three months later, pilots returned for a Phase 2 knowledge retest and to fly a similar mission. After Phase 1, a 3-variable logistic regression model significantly predicted 83.3% of flight completions. However, in Phase 2, this model was non-significant and little residual effect of video training was evident for weather knowledge, flight behavior, or flight safety. We conclude that weather training requires systematic lengthy study and practice.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

# Effects of Video Weather Training Products on General Aviation Pilot Weather Knowledge and Flight Behavior into Adverse Weather

## Background

The term “adverse weather” involves multiple factors such as restricted visibility due to low cloud ceilings, fog, rain, snow, thunderstorms, or airframe icing. Adverse weather is a perennial concern to general aviation (GA). Analyses of GA accidents from the 1970s-2000s show that, despite a relatively low incidence rate for weather-related accidents (4-5%, depending on data source and classification scheme), their fatality rate is 3-4 times higher than for other GA accident causes (Bazargan, 2005; Bud, Mengert, Ransom, & Stearns, 1997; NTSB, 1989; NTSB, 2005). This is largely because weather accidents often involve flight into terrain or loss of control, which typically results in a high percentage of fatalities.

Training is classically cited as a way to minimize hazards of flying, including weather. Yet, the body of actual research concerning measured effectiveness of weather-related training in GA is small and often involves the difficult task of trying to correlate the implementation of training methods with subsequent reductions in accidents or accident rates (Adams & Ericsson, 1992).

Formal logic asserts that correlation is necessary, but not sufficient, to demonstrate causation. Hence, we are never sure that training increases pilot skill and results in safer behavior. We merely assume it. Yet, a large body of research in perceptual, behavioral, and educational psychology shows that acquisition and retention of learning is often anything but assumable (Ellis, Semb, & Cole, 1998; Goldstein, 1999; Mackintosh, 1974; Semb & Ellis, 1994).

Characteristically, training is not permanent. New learning starts at some maximum level, after which it decays with time (assuming it is not refreshed). Thus, the amount of initial learning plus the rate of decay are two crucial parameters of knowledge retention. A third is how well knowledge is transferred from one domain to another: for example, from the classroom to the real world (Perkins, 1992). Finally, measuring “cognition in the wild” is often a very different set of circumstances from measuring in a carefully controlled laboratory setting (Hutchins, 1995). This makes real-world assessment of training a challenge for researchers.

In the real world, non-instrument rated pilots are supposed to fly by visual flight rules (VFR). VFR pilots learn they are supposed to avoid weather. However, they sometimes attempt a flight when weather is a factor along their route, either as forecast or unknown. Believing the weather is safely flyable, the VFR pilot is actually ill-prepared to deal with an encounter since practical weather skill training is usually minimal or absent from the private pilot syllabus. Similarly, a newly minted GA instrument pilot may know intellectually to avoid thunderstorms and icing when flying in clouds, but has little practical knowledge and skill to allow him or her to proceed safely.

## Purpose of This Research

Civil Aerospace Medical Institute (CAMI) researchers were tasked by the FAA Flight Standards division (AFS-810) to explore the following issues:

1. Do video weather training products significantly affect pilot weather knowledge and flight behavior in the face of potential instrument meteorological conditions (IMC)?
  - a. If so, what are the immediate effects?
  - b. Do these effects persist over time?
2. How are modern Web-based weather products used during preflight briefing?
3. Do local<sup>1</sup> pilots differ appreciably from non-local pilots in either weather knowledge or weather-related flight behavior?

The focus of the current report will be issue #1. Issues 2-3 addressed general issues of flight simulation methodology critical to both CAMI and the human factors flight simulation research community at large and are reported in Knecht, Ball, & Lenz (2010a, b).

## Method

This research was conducted in two phases. Phase 1 examined data collected from January to July, 2008. Phase 2 was similar in approach, examined data collected from August to September, 2008, and constituted the second half of a longitudinal study designed to test retention of learning.

## Participants

In Phase 1, 50 GA pilot volunteers participated with informed consent. The group mean age was 41.0 (median = 39, SD = 17.5), mean flight hours was 1314 (median = 268, SD=2709). A few high-hour pilots elevated the flight hours mean, generating a statistical concern that will be addressed later.

We recruited local pilots from a list of pilots who had participated in previous studies and by advertising in local flight schools. Non-local pilots saw an advertisement in *Flying* magazine, which resulted in more than 350 responses. Figure 1 shows demographics for eventual participants' states of residence and flying experience.

---

<sup>1</sup> A "local" pilot was defined as one living in Oklahoma at the time of the study. For the most part, this meant long-term OK residents (n=18). However, there were four instances of pilots living in OK whose state-of-legal residence was not OK because they were attending local flight schools.

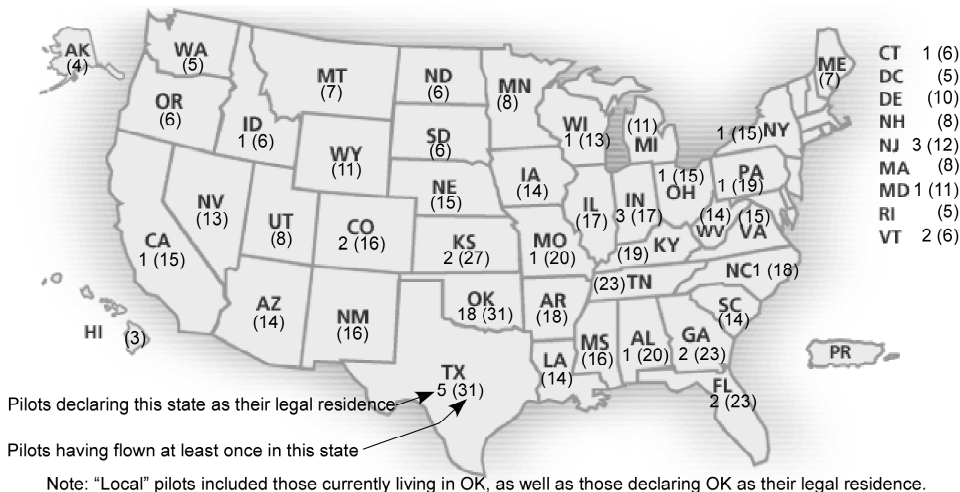


Figure 1. Pilot demographics.

## Research design, assignment to groups, and order of treatments

In measuring how training can change a person, changes can occur in what we *know* (cognitive knowledge) and/or how we *behave* (here, flight behavior). Therefore, this experiment attempted to measure both.

Table 1 shows the 3x2x2 mixed design with its three primary independent variables of training product, instrument rating, and residence. These were analyzed as between-groups comparisons. Conversely, weather knowledge was analyzed as within-groups comparison in which each pilot served as his or her own statistical control over repeated administrations of different (but equivalent) test forms.

Two variants of weather direction were used as distractors to help divert attention from the fact that the flight scenario would be essentially repeated in Phase 2 ("Flight scenario" column). Variant 45 had weather approaching from 45° (aeronautical coordinates, 0° = North, increasing clockwise), while Variant 135 had weather approaching from 135°. However, weather direction was only a distractor, not a variable of interest.

## Apparatus

**Weather training products/control materials.** We selected two well-known video weather training products from a list of candidate products. Given the impossibility of knowing product quality a priori, two publicly prominent products were chosen. The authors of these products graciously provided them, at no cost, on condition of confidentiality; therefore, their wishes for confidentiality shall be respected in this report.

Training product 1 focused mainly on the aeronautical decision-making aspects of weather flight. It offered systematic, mnemonic risk factor checklists applicable to specific factors such as the weather in question, internal pilot factors

affecting performance (e.g., skill, health, fatigue), and factors external to the pilot that could affect risk-taking (e.g., passengers needing to arrive at their destination by a certain time). After each video lecture session, it presented hypothetical flight scenarios and asked the student to evaluate these based on the lecture content presented so far.

Table 1

*Experimental design, Phases 1 and 2.*

Independent variables			Dependent variables
Primary		Secondary	
Weather training product	Instrument rating	Locality of residence	<i>Weather knowledge score</i>
Video product 1	Instrument-rated (IR)	Local	Web preflight data
		Non-local	Pages viewed
	Private (non-IR)	Local	Pageview duration
		Non-local	<i>Flight simulator data</i>
Video product 2	Instrument-rated (IR)	Local	Flight duration
		Non-local	Minimum distance to ABQ
	Private (non-IR)	Local	Cloud clearance
		Non-local	Minutes in IMC
Control (non-weather video)	Instrument-rated (IR)	Local	Minutes < 500' fr cloud base
		Non-local	Minutes < 500' AGL
	Private (non-IR)	Local	Initial takeoff decision (Y/N)
		Non-local	

Training product 2 focused largely on the recognition of different cloud types, visibility conditions, horizon recognition, and terrain clearance. Still pictures of common weather types were shown, after which pilots were queried as to their appropriateness for VFR flight. This technique has been used in research (Wiggins & O'Hare, 2003). Sample exercises showed still pictures of a weather situation as seen aloft, asking what recognition factors were problematic, and then asking for a go/no-go decision on VFR flight. A second section presented details of an accident scenario, asking pilots to decide the primary cause. A third section began by discussing factors involved in deciding whether or not to divert because of weather. It then presented a detailed weather flight scenario involving a number of possible alternate landing sites, asking which was most appropriate. Finally, it presented a

list of several potential flights with preflight briefing details of each, next it asked for a go/divert decision after presentation of a still photo of in-flight weather, and finally it asked for a justified choice of alternate, if diversion was chosen.

The third video group—the Control group—received an FAA-produced video on aviation physiology, having nothing whatsoever to do with weather.

**Timer.** To examine whether time spent viewing the training product might be a useful statistical covariate, we wrote a timer utility. This was activated by each pilot at the beginning of training and turned off afterward to record elapsed viewing time.

**Weather Knowledge Tests.** We constructed three parallel forms of a 30-question general weather knowledge test. These were matched on item difficulty, using questions and proportion-correct data provided by FAA's Airman Testing Standards Branch (AFS-630). One-third of the questions on each test were taken from private pilot tests; two-thirds came from instrument rating tests (FAA, 2008). This was not expected to pose a problem since pre- minus post-treatment change scores were to be analyzed, which are immune from overall test difficulty as long as the tests are neither impossibly difficult nor trivially easy (i.e., do not suffer from either ceiling or floor effects). The administration order of the parallel forms was counter-balanced across pilots to control for any unintended variations in difficulty across forms.

Tests were administered on a laptop computer using software we wrote in Microsoft *Visual Studio 2005*<sup>TM</sup>. The program recorded each question, each response, response-correct or incorrect, time spent per response, overall percent correct, and total elapsed time. Figure 2a shows a screenshot of a sample question.

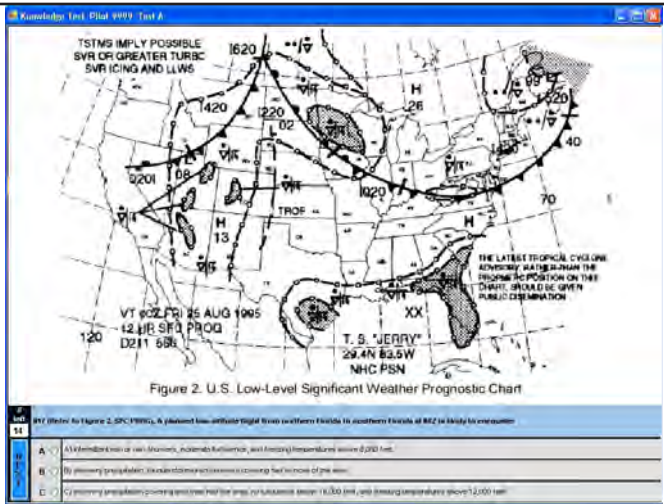
**Preflight weather-briefing materials.** Briefing materials included a verbal description of the flight mission, plus standard DFW (Dallas-Fort Worth) and ABQ (Albuquerque) sectional charts.

To simulate Internet weather briefing, we wrote a part-task emulation of the NOAA/NWS Web weather-briefing site <http://aviationweather.gov>, also in Microsoft *Visual Studio 2005*. This emulation automatically recorded which pages were viewed and the view duration of each page. Figure 2b illustrates a sample page. The Appendix shows screenshots of all pages.

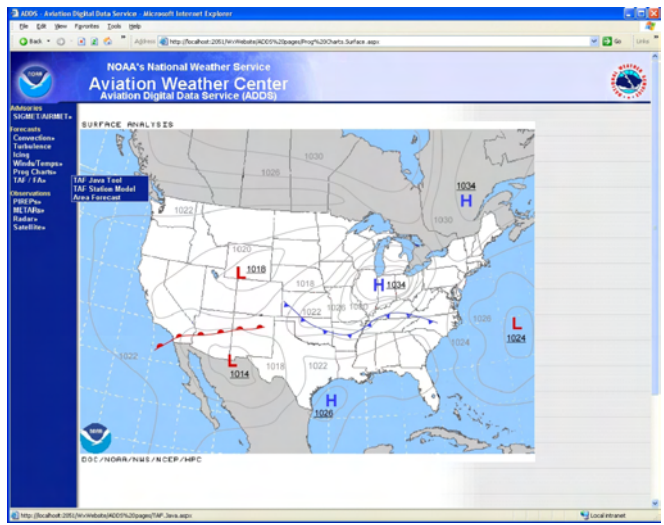
### **Advanced General Aviation Research Simulator (AGARS)**

Figures 3a and 3b show the AGARS—a real-time, fixed-based flight simulator configured as a Piper Malibu for this experiment. Choice of the Malibu configuration was intended to both “level the playing field” and increase overall task difficulty since none of the participants had ever flown this aircraft.





a



b

Figure 2. Sample screenshots from a) the Weather Question Tester program, b) the <http://aviationweather.gov> emulation.

Equipped with a high-resolution visual system with a 150° field of view, AGARS allows precise control of meteorological conditions. It continuously captures up to 150 variables at 30Hz for a 4-hour mission, including up to 85 programmable non-routine events. It is equipped with an experimenter operating station (EOS) and an ATC workstation. During the course of a flight scenario, the EOS allows the experimenter to visually monitor the cockpit and simulation environment. In addition to digital recordings of the flight data, a digital camera records a global view of the cockpit and pilot.

To enable in-flight AWOS weather updates, we wrote a control panel capable of triggering prerecorded METAR information (Figure 3c). Pilots could tune the cockpit radio to one of a set of frequencies, alerting the experimenter to activate the AWOS control panel, which triggered the corresponding pre-recorded METAR into the pilot's headphones.

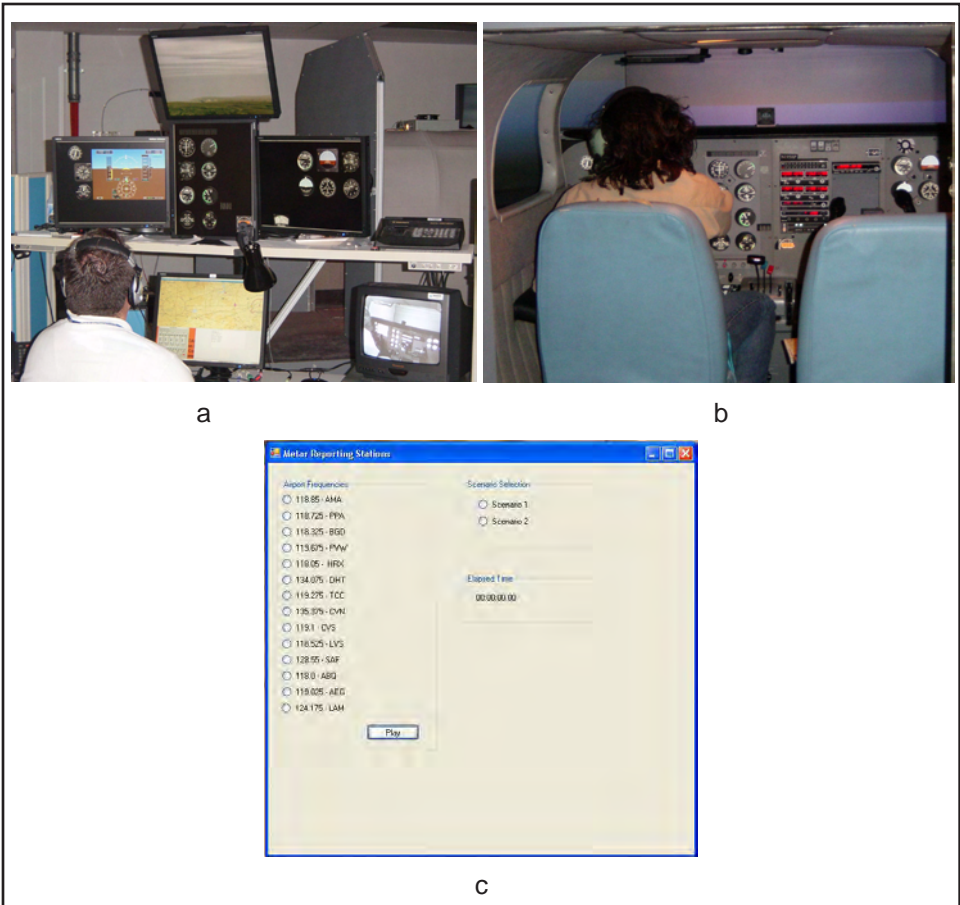


Figure 3. a & b) the CAMI Advanced General Aviation Research Simulator (AGARS), c) AWOS emulator. Photos used by permission of participant.

To emulate the Flight Service Station (FSS) providing air traffic control services such as flight following, vectors-to-destination, and weather, one of us (Ball) served as a pseudo-FSS briefer during the flight phase.

### Procedure

Upon arriving at the simulator lab, we asked pilots to plan an east-to-west, VFR flight from Amarillo, TX (AMA) to Albuquerque, NM (ABQ). This route takes approximately 90 minutes to fly in the Malibu at high speed cruise. Pilots planned the route utilizing two VORs (VHF OmniRange Navigation System) and an ADF (Automatic Direction Finder). Pilots could access the Web-based weather emulation ad lib on a stand-alone PC during preflight planning. Upon finishing their flight

plan, pilots took the post-weather knowledge test. Next, we offered a 15-minute convenience break, after which each pilot had a 30 to 40-minute familiarization session with AGARS. Specific training was provided on the usage of the autopilot, horizontal situation indicator (HSI), and the flight parameters and characteristics of the Malibu aircraft (e.g., maximum/stall speeds, associated power settings). Familiarization typically took about 60 minutes. Due to unfamiliarity with the simulator and the complexity of the Piper Malibu, pilots were allowed to ask for assistance with flight settings at anytime during the course of the flight scenario.

The route consisted of gradually rising terrain during the first two-thirds of the flight, followed by a dramatic elevation change during the last one-third. During the flight, weather deteriorated. Initially, visibility was set at 8 nautical miles and gradually decreased to 5 miles by the time the pilots had flown approximately two-thirds of the route. Concomitantly, cloud ceilings were lowered from 4500 feet AGL to 3500 feet AGL across the same stretch of terrain, gradually squeezing pilots closer to the ground. Figure 4 illustrates this with two sample flight profiles illustrating a) brief IMC penetration and prolonged scud running and b) violation of ground clearance over the mountains near ABQ.

The terrain issues, coupled with marginal visual meteorological conditions (VMC) and potential rapidly changing barometric pressure, resulted in a potentially dangerous flying situation with hazardous encounters throughout the course of the flight.

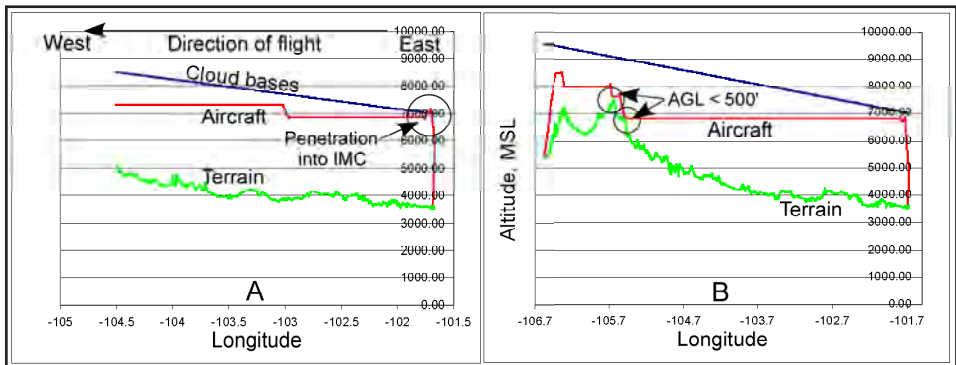


Figure 4. Two pilots' flight profiles. Since the flight was basically east-to-west, the x-axis is drawn as degrees longitude by feet altitude-above-mean sea level (MSL) on the y-axis.

## Results (Phase 1)

### Preliminary examination of data

Since this was a preliminary study, a decision was made to leave data uncorrected for multiple comparisons. The data distributions were, however, checked for statistical normality to determine the appropriate subsequent statistical tests. Weather knowledge pre- and post-test scores appeared normal (2-tailed Shapiro-Wilk  $p_{pre-test} = .297$ ,  $p_{post-test} = .786$ , NS). However, flight hours, showed serious non-normalities due to the presence of a few very high-hour pilots. The Web-emulation data also appeared non-normal due to a considerable number of pilots forgetting

to close out the final page. Of the flight simulator data, only flight duration passed the Kolmogorov-Smirnov normality test. Finally, although we instructed pilots to use the timer utility to monitor how long they studied their training product, compliance proved low. Therefore, study time was not used as a covariate in subsequent analysis.

### First-order Data Relations

Table 2 shows correlations between key variables. Statistically significant correlations are highlighted in gray. *P*-values approaching .05 significance are also included for the sake of interest. Spearman correlations ( $r_s$ ) are nonparametric, being based on rank order. Point-biserial correlations ( $r_{pb}$ ) are used when one variable is dichotomous, the other continuous. However,  $r_{pb}$  is still a mean-based statistic, therefore not purely nonparametric even though dichotomous variables are non-normal by definition. As such,  $r_{pb}$  may be subject to higher Type I error when the continuous distribution is non-normal, hence, some caution is appropriate during interpretation.

Table 2

*Correlations between key variables.*

Variable 1	Variable 2									
	Instrument rating (1=instrument rated) <sup>1</sup>	Locality of residence (1=Local) <sup>1</sup>	Pilot age <sup>2</sup>	Pilot flight hours <sup>2</sup>	Ave. wx Knowledge <sup>2</sup>	Web pre-flight duration <sup>2</sup>	Flight duration <sup>2</sup>	Minimum dist to ABQ <sup>2</sup>	Minutes scud running <sup>2</sup>	Minutes in IMC <sup>2</sup>
Instrument Rating	1.0									
State of Residence	.3	1.0								
Pilot Age	.523 <sup>1</sup> (.0001)	.3	1.0							
Pilot Flight Hours	.401 <sup>1</sup> (.004)	.3	.757 ( $<.001$ )	1.0						
Ave. Wx Knowledge	.233 (.057)	.271	-.035	.086	1.0					
Web Preflight Duration	.020	<b>-.348</b> (.013)	<b>.417</b> (.003)	.227	.204	1.0				
Flight Duration	-.039	.042	<b>-.423</b> (.002)	-.270	-.010	-.222	1.0			
Minimum Dist to ABQ	.013	.013	<b>.422</b> (.002)	<b>.303</b> (.032)	.029	.242	<b>-.936</b> ( $<.001$ )	1.0		
Minutes scud running	-.013	-.012	.051	.107	-.054	.027	.013	-.042	1.0	

Minutes in IMC	-0.020	-0.005	-0.089	-0.084	-.123	-.124	.028	-.035	.676 ( <i>&lt;.001</i> )	1.0
Minutes < 500' AGL	-.281 (.048)	.144	-.167	-.289 (.041)	-.229	-.039	.379 (.007)	-.384 (.006)	-.095	-.174

<sup>1</sup> $r_{pb}$  = Point-biserial correlation; <sup>2</sup> $r_s$  = Spearman rho correlation; Low p-values are in parentheses (all others are non-significant (NS)); <sup>3</sup> No correlation run because sample had been partitioned for these factors.

**Trivial and marginal relations.** Five of the largest correlations are statistically significant but expected. These cells are marked in the lightest shade of gray. Additionally, given the 50 correlations computed, 2-3 can be expected to appear “significant” near  $p = .05$  merely due to chance. Such marginal cells are highlighted in a darker gray, but the text is not boldfaced.

**Non-trivial relations.** Table 2 highlights non-trivial relations in a dark shade of gray with boldfaced text. They are:

1. Locality of residence x Web preflight duration ( $r_{pb} = -.348$ )
2. Pilot age x Web preflight duration ( $r_s = .417$ )
3. Pilot age x Flight duration/ Minimum distance to ABQ ( $r_s = -.423/.422$ )
4. Flight duration/minimum distance to ABQ x Minutes < 500' AGL ( $r_s = .379/-.384$ )

Correlation 1 (-.348) implies that local Oklahoma pilots tended to spend slightly less time using the Web preflight briefing tool than non-Oklahoma pilots did ( $\bar{x} = 20.31$  v. 13.66 min). The effect size was modest, accounting for  $r^2_{pb} = 12\%$  of the variance. Correlation 2 (.417) implies that older pilots tended to spend somewhat more time using the Web tool than younger pilots did. Effect size was 18%. Correlations 3 (-.423/.422) imply that older pilots tended to have somewhat shorter flights (and, hence, to end up farther away from ABQ). Effect size was about 18%. Finally, Correlations 4 (.379/-.384) represent the flight scenario's tendency to “squeeze” pilots between clouds and terrain near the destination ABQ.

## Specific Effects

### Effect of the weather training products on GA pilot weather knowledge.

Viewing a weather-training product did not seem to significantly improve pilots' weather knowledge test scores. Repeated measures analysis of variance (ANOVA) for post-test-pre-test score gain x training product interaction yielded a non-significant  $p_F = .734$ .

### Relation between pilot weather knowledge and subsequent flight safety.

Pilots with higher weather knowledge did not appear to be safer pilots. As Table 2 showed, average weather knowledge ((pre-+post-test score)/2) did not correlate significantly with any flight behavior variables. Spearman correlations between weather knowledge scores and flight duration, minimum distance to ABQ, minutes scud running, minutes in IMC, and minutes < 500' AGL all ranged from  $-.229 \leq r_s \leq .029$ , all NS.

Furthermore, weather knowledge did not seem strongly influenced by age, flight hours, or instrument rating ( $r_s = -.035, .051, r_{pb} = .233$  [respectively], all NS). Although instrument rated pilots showed a slightly higher average knowledge score (67.6%) than did non-instrument rated pilots (60.8%), this just missed the statistical criterion for reliability (1-tailed  $p_t = .057$ , NS).<sup>2</sup>

**Effect of Web preflight briefing time on subsequent flight safety.** Pilots who spent more time on their Web-based weather briefing did not seem to be safer pilots. Spearman correlations of Web preflight duration with flight duration, minimum distance to ABQ, minutes scud running, minutes in IMC, and minutes < 500' AGL ranged from  $-.222 \leq r_s \leq .242$  respectively, all NS.

**Takeoff hesitancy.** We told pilots that the best way to give good flight data was to treat this mission as if it were a real flight. Given those instructions, 12 of the 50 pilots initially stated that having to fly this mission VFR, they would choose not to even take off. This was perhaps predictable, given the weather and being scrutinized by FAA officials at an FAA facility. Therefore, to overcome any reservations they might understandably have about being scrutinized, all “hesitators” who declined to take off were explicitly asked to take off and fly at least briefly. All complied. If this methodology troubles some readers, recall the Milgram experiments (1974) on obedience-to-authority.

Locality of residence had no apparent effect on takeoff hesitancy—18% of local (Oklahoma) pilots hesitated versus 32% of non-local (non-Oklahoma) pilots (2-tailed  $p_{\chi^2} = .251$ , NS). Directionality of effect was consistent with intuition, since locals would be more likely to be familiar with the terrain.

Instrument rating did not seem to matter (15% hesitancy for instrument rated v. 33% for non-instrument rated, 2-tailed  $p_{\chi^2} = .138$ , NS). Here, too, directionality of effect was in the anticipated direction, since one would expect somewhat greater confidence from instrument rated pilots.

Despite the confidence-building tendencies often associated with experience, neither age nor flight hours seemed to affect hesitancy (2-tailed Mann-Whitney U,  $p_U = .146, .625$  respectively, NS). At first inspection, the cause of this takeoff hesitancy appeared mysterious.

**Effect of takeoff hesitancy on subsequent flight safety.** The 12 hesitators did not end up flying noticeably safer than the remaining 38 pilots. There were no significant differences between hesitators and non-hesitators for minutes spent in IMC, minutes scud running, or minutes < 500' AGL (2-tailed Mann-Whitney  $p_U = .102, .147, .498$  respectively, all NS). However, hesitators did seem to continue their conservatism into their flight, making significantly briefer flights ( $p_U = .002$ ) with consequently less penetration into the marginal weather close to ABQ ( $p_U < .001$ ).

**Effect of the weather training products on takeoff hesitancy.** The weather training products appeared associated with takeoff hesitancy. Table 3 shows the number of pilots who initially hesitated versus the values expected by chance (in

---

<sup>2</sup> This was based on average combined pre-test and post-test scores ( $[(\text{pre-test score} + \text{post-test score})/2]$ ).



parentheses). The Yates-corrected  $p_{\chi^2}$  is .034, implying that the training groups differed. However, there was a statistical issue clouding the results: Half the cells had expected frequencies  $< 5$ , which violated the convention of  $\leq 20\%$  allowable.

Given that statistical caveat, *if* this was indeed a reliable effect, pairwise tests of odds-ratios implied that the unusual group was the Control, where 17 of 18 pilots showed no hesitancy to take off.

Table 3

*Takeoff hesitancy*

		Trg Prod 1	Trg Prod 2	Control
Initial takeoff decision	Yes	12 (12.2)	9 (12.2)	17 (13.7)
	No	4 (3.8)	7 (3.8)	1 (4.3)
Pairwise odds-ratios, 1-tailed $p$		← .152 →		
			← .004 →	
		← .037 →		

In other words, studying a weather training product may have made pilots more hesitant to take off into deteriorating weather. However, *cognitive priming* is an alternate hypothesis that we will consider later in the Discussion section.

**Effect of the weather training products on subsequent flight safety.** Overall, the weather training products did not appear significantly associated with subsequent flight safety.

The Control group showed significantly less takeoff hesitancy, as we just saw. It also displayed greater flight duration and, consequently, lower minimum distance to ABQ (Kruskal-Wallis  $p_{KW} = .007, .005$  respectively). Follow-up pairwise Mann-Whitney U tests implied that the Control group was significantly different from both weather training products ( $p_{U-TRG1 \times CONTROL} = .011, .004$  respectively and  $p_{U-TRG2 \times CONTROL} = .004, .005$  respectively), although the two weather products themselves did not differ significantly ( $p_U = .867, 1.0$  respectively, NS). Now—because the maximum hazard of this flight lay near the destination—we might be tempted to conclude that the longer flights of the Control group should predict greater risk exposure. As Table 2 showed, this was supported by a moderate correlation ( $.379, p = .007$ ) between flight duration and minutes  $< 500'$  AGL.

However, there were no significant overall differences between the three training groups for subsequent minutes spent in IMC, minutes scud running, or minutes  $< 500'$  AGL ( $p_{KW} = .245, .158, .812$  [respectively], all NS). Even though the Control group showed less hesitancy and longer flight duration, and even though longer flight duration correlated significantly with minutes  $< 500'$  AGL, the net effect of the weather training videos on subsequent flight safety seemed nonsignificant.

So, how can there be no significant differences in flight safety among the three training groups? If seeing the weather training video related to takeoff hesitancy, and takeoff hesitancy related to flight duration, and flight duration related to min-



utes spent < 500' AGL—how could weather video not relate to minutes spent < 500' AGL?

The answer lies in the nature of causation versus correlation. If each factor perfectly *caused* the next factor in the chain, then the first factor would perfectly predict the final factor. But, if each factor only *partially predicts* the next factor, then the overall correlation between the first and last factors can theoretically be as low as 0.

## Modeling flight behavior

**Cluster analysis.** One of the most interesting questions here is: “*What differentiated pilots who chose to complete the flight through deteriorating weather from pilots who chose not to complete the flight?*” To investigate this we constructed models—simplifications that still hopefully captured essential relations.

Cluster analysis is one approach to modeling. It starts with a set of measurements (“variables”) taken on individuals (“cases”—here, individual pilots). It then explores the relations between variables by combining individual cases into groups (“clusters”). The end goal is to group cases so that those within the same cluster are more similar to each other than they are to cases from different clusters. “Similarity” is operationalized by calculating a “mathematical distance” between cases. Once complete, it becomes the analyst’s job to interpret what each cluster means in logical and practical terms.

Here, we used a “TwoStep” cluster analysis procedure (SPSS, 2001) to group pilots based on demographic characteristics as well as behavioral responses to the simulated flight scenario. Schwarz’s Bayesian Criterion (BIC) was the clustering criterion and log-likelihood the distance measure. In the first step, sequential clustering calculated the BIC for each cluster within a specified range and used that to estimate the initial number of clusters. In the second step, the estimate of clusters was reduced by finding the largest increase in “mathematical distance” between the two closest clusters using an agglomerative hierarchical clustering method (SPSS, 2001).

SPSS’s “TwoStep” cluster analysis works with both categorical and continuous variables when using the log-likelihood method. Assumptions of normality often tend to be relaxed in cluster analysis so nonparametric follow-up tests were used to examine individual relations (see below).

We selected candidate variables based upon logic and prior results of correlational analysis (e.g., Table 2). The initial *categorical* candidate variables were weather training product, pilot’s instrument rating, locality of residence, initial go/no-go takeoff decision, whether or not a preflight weather call was made just prior to takeoff, whether or not penetration into IMC occurred during the flight, and final flight decision. Final flight decision was quantified by collapsing all possible categories into a binary (2-category) categorical variable *To ABQ* (“Did a given pilot complete the entire flight to ABQ, yes or no?”). *Continuous* candidate variables were pilot age, total flight hours, minimum final distance from the destination, time spent in IMC, time spent less than 500’ below cloud ceiling, time spent at less than

500' AGL, and number of en route weather updates. Note that weather knowledge scores were excluded for prior non-significance (see Table 2).

**Results of the cluster analysis.** Compared to Cluster 1 pilots (n=32), Cluster 2 pilots (n=16) tended to be

1. younger, with (Age)
2. lower flight hours, (FH)
3. closer final minimum distance to ABQ, (MDABQ)
4. more minutes flying less than 500' AGL, (M<500AGL)
5. usually did *not* receive a weather training product, (Trg Product)
6. greater % "Go" responses for takeoff (100%), (Takeoff [TO] Decision)
7. less likely to recheck weather just before takeoff, (Recheck, Y=Yes)
8. greater % flew all the way to ABQ (100%). (To ABQ)

Interestingly, 100% of the Cluster 2 pilots flew direct to ABQ through the near-by mountain pass, whereas only one pilot in Cluster 1 did so. In contrast, over 84% of Cluster 1 pilots decided to divert or return to the departure airport (AMA). Of the four Cluster 1s who did fly to ABQ, three flew completely around the troublesome mountain range. Together, these results support the notion of Cluster 2 pilots as greater risk-takers.

In contrast, there were no significant differences between Cluster 1 and Cluster 2 for instrument rating, locality of residence, number of en route weather updates, minutes spent in IMC, or minutes scud running.

**Binary logistic regression analysis.** As a second modeling approach, we used stepwise forward likelihood-ratio binary logistic regression. Requiring no assumptions about the frequency distributions of the predictor variables, logistic regression takes a set of candidate variables, categorical or continuous, and selects only those demonstrating significant orthogonal (uncorrelated) ability to help predict a binary outcome (Tabachnick & Fidell, 2001).

Specifically, we wanted to predict which pilots would risk flying completely through the deteriorating weather (*To ABQ* = Yes/No). Table 4 shows the smallest set of variables capable of doing that reliably.

Note that *TO Decision* reflects "takeoff hesitancy" as discussed earlier and that training product is broken out into its three groups. Negative B-weights mean that a *positive* value for the independent variable subsequently related to a *reduced* groupwise tendency to fly all the way to ABQ<sup>3</sup> For example, pilots hesitant to take off (*TO Decision* = 1) subsequently showed a reduced tendency to fly all

---

<sup>3</sup> B-weight functions similarly to a correlation coefficient, except that it is not restricted to  $-1 \leq r \leq 1$ . The Control group has no B-weight because SPSS essentially used it as an "invisible baseline," in relation to which the B-weights Trg Prod1 and Trg Prod2 could then be compared. For example,  $B_{TP1} = -3.08$ , which is less than 0, means that TrgProd1 pilots were less likely to complete the flight to ABQ than were Control group pilots.

the way to ABQ. Similarly, pilots receiving either weather training product subsequently showed reduced tendency to fly all the way to ABQ compared to the Control group.

Table 4

*Binary logistic regression for To ABQ*

	B	<i>p</i> <sub>if term removed</sub>
Age	- 0.081	.002
TO decision	- 21.20	.016
Control <sup>14</sup>		
Trg Prod 1	- 3.08	.006
Trg Prod 2	- 2.53	
Constant	4.64	

Nagelkerke R<sup>2</sup> = .640

In practical terms, this is a moderate-strength model, accounting for 64.0% of the explainable (Nagelkerke) variance in the data. It implies that pilot age may work in combination with an instinctive reaction to a weather situation to affect ultimate continuation into adverse weather. Impulsivity may be further reduced by the presentation of a training product. This contrasts somewhat with the null conclusion reached earlier about training product, so we will revisit that theme in the Discussion section.

Table 5 compares the prediction success rate for completed flight to ABQ made by logistic regression (boldface) versus cluster analysis (italics, in parentheses). Grey cells represent successful predictions.

Table 5

*Success rate for **binary logistic regression** versus (cluster analysis)*

Observed <i>To ABQ</i>	Predicted <i>To ABQ</i>		
	Did not make it to ABQ	Made it to ABQ	% correct
Did not make it to ABQ	<b>26 (27)</b>	<b>4 (4)</b>	86.7 (87.1)
Made it to ABQ	<b>4 (0)</b>	<b>14 (16)</b>	77.8 (81.3)
Overall % correct	Base logistic prediction rate = 62.5%		<b>83.3 (91.5)</b>

This shows that a simplified logistic model containing only pilot age, initial takeoff decision, and training product correctly predicted 83.3% of these pilots' overall decisions whether or not to fly through the deteriorating weather all the way to ABQ. The model was slightly better at predicting those who did not make it to ABQ (86.7% correct) than it was at those who did (77.8% correct).

Overall, this 3-variable model produced a gain of about 21% from the base rate predicted by a constant only (62.5%,  $p=.000004$ ). Compare this to the 8-variable cluster model's correct predictions of 91.5% versus a "complete" 15-variable logistic regression (not shown) where 100% of all cases were predicted correctly. However, note that the "complete" model was vastly overfitted, meaning it contained too many predictors given the number of cases. A case/predictor ratio of  $\geq 10/1$  is typically a rule of thumb in regression analysis, implying that our models should arguably be limited to  $48/10 = 4-5$  predictors. This shows that modeling involves a tradeoff. Simpler models, while offering somewhat less accurate predictions, compensate with greater reliability.

Finally, recall our earlier statements that completing the entire flight to ABQ did not always reflect dangerous behavior (as measured, for example, by time spent  $< 500'$  AGL). In fact, we cross-checked the logistic regression model against regular linear regression<sup>4</sup> by substituting the original binary outcome variable *To ABQ* with the continuous flight risk outcome variable *Minutes < 500' AGL*. This showed no effect of the "FH+TO Decision+Training Product" model on actual flight risk ( $p = .612$ , NS). So, again, we need to consider the distinction between mere flying and the subsequent hazard of flying.

### Summary of Phase 1 Results

Using simple univariate modeling, merely seeing a 90-minute weather video training product all by itself produced no significant changes in either weather *knowledge* or ultimate *flight safety*. The training did, however, affect some aspects of *flight behavior*, which might arguably have an *indirect* effect on flight safety. Compared to the Control group, both training products seemed to induce *takeoff hesitancy*. "Hesitators" flew into uncertain weather only after direct encouragement by the experimenter. In contrast, 17 of 18 Control group pilots took off without any encouragement.

Subsequently, hesitators continued their conservatism, tending to make significantly *shorter flights* than non-hesitators. Since the bulk of the scenario danger lay near the flight's destination, we might be tempted to imagine a chain of events—that the training product induced takeoff hesitancy, which induced shorter flights, which led to lower groupwise risk exposure. However, these separate events, though individually related, did *not* result in a statistically significant causal event chain from beginning to end.

In contrast, multivariate modeling showed that a *combination* of

- higher pilot age,

---

<sup>4</sup> However, bear in mind that this was not technically a reliable analysis because flight hours and minutes  $< 500'$  AGL were both severely non-normal distributions, which violates the assumptions of linear regression.

- receiving either weather training product, and
- takeoff hesitancy

significantly predicted 26 of 30 diversions (86.7%) from deteriorating weather and 14 of 18 successful flight completions (full penetration into the weather, 77.8%).

## Results (Phase 2)

### Participants and Attrition

Six Phase 1 pilots were unable to continue in Phase 2. Fortunately, no statistically significant changes were found between training groups for key demographic factors of age, flight hours, or certificate type.

Because the flight situation was essentially the same from Phase 1 to 2, the issue of learning had to be addressed.<sup>5</sup> Pilots might fly better in Phase 2 simply because they had learned the aircraft and the physical terrain. Therefore, two methods were used to distract pilots from the similarities between Phases 1 and 2. First, the approach direction for Phase 2 weather was made symmetrical to, and counterbalanced with, whatever each pilot had experienced during Phase 1. Second, a primary flight display (PFD) was introduced (Figure 5). Each pilot received a 20-30 minute PFD training session involving a short flight from an airport out to a VOR at a requested altitude followed by a return to the departure airport for landing.

### Data Normality, Phase 2

Acceptable Shapiro-Wilk normality was found only for weather knowledge test scores and flight duration. Hence, most data were again analyzed non-parametrically.

### Specific Effects

#### **Effect of the weather training products on GA pilot weather knowledge.**

As in Phase 1, there were no significant changes in weather knowledge over time. Repeated measures ANOVA showed neither significant weather knowledge effects ( $p_F = .396$ ) nor training product effect ( $p = .908$ ). Adding instrument rating and pilot's locality of residence failed to enhance the analysis (smallest  $p = .389$ ). Table 6 shows means (with 95% confidence interval in parentheses).

---

<sup>5</sup> Two reasons dictated not using 2 different, counterbalanced routes. First, the AGARS scenery database did not encompass the entire continental U.S. Second, even if it had, pilots would have easily spotted the essential flight features (i.e., slowly rising terrain terminated by hills, with deteriorating visibility and ceiling along the way).



Trg Prod 2	1414	1021	924	1120
Control	1117	831	873	940
Column Ave.	1339	962	927	1076

Could this explain why we failed to see any significant knowledge gain for any of the treatment groups? Unfortunately, in Phase 1, we saw that ET did not predict weather knowledge score. On average, pilots who spent *more* time on the test actually got *lower* average scores ( $r_{\text{Spearman}} = -.187$ ). But, because it was not statistically significant, this explains why ET was useless as a covariate in ANOVA.

We speculate that the pre-post drop in ET probably occurred because 1) on first testing, all pilots were unfamiliar with the software and 2) over half took their first test in the undisturbed privacy of their hotel rooms or homes, with little time pressure. In contrast, all remaining tests occurred in the laboratory where pilots were under supervision often with a plane to catch that afternoon, *and* none could fly the simulator until after their knowledge test was finished. This does not necessarily mean they were less attentive in the later sessions. It may merely mean they were highly motivated to get the test over with and get into the simulator and likely concentrated more intently on the last two tests.

**Flight behavior.** Figure 6 shows Phase 1 flight paths (left column) and Phase 2 flight paths (right column), grouped by training product (rows 1-3). Non-instrument-rated pilots' paths are shown as black lines; instrument-rated pilots' as white lines.

Figure 6 shows us mainly how similar the flight patterns manifested across training groups, instrument rating, and Phases 1 versus 2. Most pilots who did complete the flight picked their way relatively directly, straight through the mountain passes east of ABQ, even though this required great care to simultaneously maintain adequate cloud and ground clearance. Interestingly, the one Phase 1 Control group instrument-rated pilot (lower-left box, white path) seen to flank the weather by flying north and later south in Phase 2 was the same pilot. He reported that, after being successful the first time, he was simply trying the same strategy, just going the opposite direction.



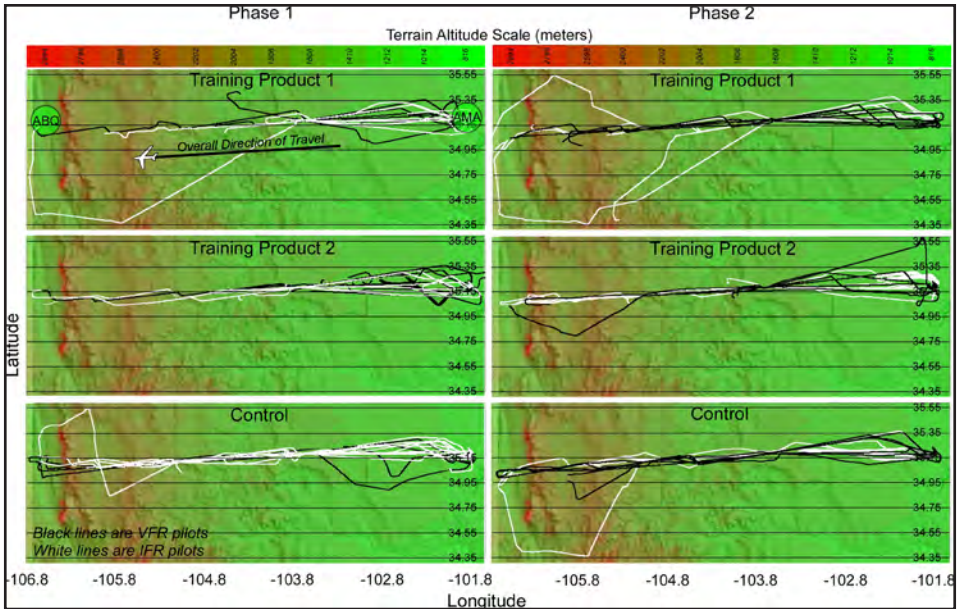


Figure 6. Top-down, flight-profile views for Phases 1 and 2, laid over the terrain map. Terrain slowly rose as pilots flew east to west, squeezing them between clouds and ground, especially near the north-south ridge just before ABQ. Digital elevation data were obtained from National Geophysical Data Center (2008) and drawn by *Mathematica* (2008).

We can analyze flight pattern consistency—if a pilot did or did not make it all the way to ABQ during Phase 1, did he or she do the same thing in Phase 2? Table 8 shows the 2x2 consistency matrix for 42 pilots.<sup>6</sup>

Table 8

*Consistency of flight decisions from Phase 1 to 2.*

Phase 1		Phase 2	
		No	Yes
Made it to ABQ	No	22 (16.7)	6 (11.3)
	Yes	3 (8.3)	11 (5.7)

1<sup>st</sup> number is actual n (2<sup>nd</sup> is expected n).  $p_{Fisher} = .0007$  (2-tailed)

<sup>6</sup> 44 completed both Phases 1 and 2, but 2 experienced controlled flight into terrain and so were excluded from this analysis

Most pilots behaved more consistently than expected by chance (note the two gray-highlighted cells, 22+11=33 of 42 = 78.6%,  $p_{\text{Fisher's Exact Test}} = .0007$ ). Pilots tended to repeat whatever flight decision they made the first time (e.g., if they flew all the way to ABQ in Phase 1, they generally did so in Phase 2 as well).

**AGARS intercorrelations.** Table 9 is the Phase-2 equivalent of Table 2.

Table 9

*Correlations between key Phase 2 variables.*

Variable 1	Variable 2									
	Instrument rating (1=instrument rated) <sup>1</sup>	Locality of residence (1=Local) <sup>1</sup>	Pilot age <sup>2</sup>	Pilot flight hours <sup>2</sup>	Ave. wx Knowledge <sup>2</sup>	Web pre-flight duration <sup>2</sup>	Flight duration <sup>2</sup>	Minimum dist to ABQ <sup>2</sup>	Minutes scud running <sup>2</sup>	Minutes in IMC <sup>2</sup>
Instrument Rating	1.0									
State of Residence	.3	1.0								
Pilot Age	.382 (.010)	-.275	1.0							
Pilot Flight Hours	.379 (.011)	.060	.736 (<.001)	1.0						
Ave. Wx Knowledge <sup>4</sup>	.247	-.255	-.066	.154	1.0					
Web Preflight Duration	-.058	-.421 (.004)	.475 (.001)	.371 (.013)	.225	1.0				
Flight Duration	.130	-.281	-.107	.043	.070	.109	1.0			
Minimum Dist to ABQ	-.083	.215	.231	.030	-.064	-.023	-.908 (<.001)	1.0		
Minutes scud running	.077	.206	-.049	-.037	-.036	.038	.090	-.194	1.0	

Minutes in IMC	.088	-.202	-.148	-.235	-.260	-.076	.138	-.248	.234	1.0
Minutes < 500' AGL	-.147	.034	-.139	-.053	-.175	.049	<b>.373</b> (.013)	<b>-.422</b> (.004)	-.033	<b>.446</b> (.002)

<sup>1</sup> $r_{pb}$  = Point-biserial correlation; <sup>2</sup> $r_s$  = Spearman rho correlation; Low p-values are in parentheses (all others are non-significant [NS]); <sup>3</sup> No correlation run because sample had been partitioned for these factors.

**Trivial relations.** As in Phase 1, five of the largest correlations are statistically significant, but trivial. These cells are marked in a light shade of gray.

**Non-trivial relations.** These are boldfaced and marked in a darker shade of gray.

1. Web preflight duration x Locality of resi-dence (rpb = -.421)
2. Web preflight duration x Pilot age (rs = .475)
3. Web preflight duration x Pilot flight hours (rs = .368)
4. Minutes < 500' AGL x Flight duration (rs = .373)
5. Minutes < 500' AGL x Minimum distance to ABQ (rs = -.422)
6. Minutes < 500' AGL x Minutes in IMC (rs = .446)

Correlations 1-3 imply that non-Oklahoma pilots, older pilots, and higher flight-hour pilots tended to spend slightly more time using the Web preflight briefing tool. Effect sizes were no more than modest, accounting for r2pb = 18, 23, and 14% of the variance, respectively. Correlations 4-6 imply that pilots who flew longer, got closer to ABQ, and those who spent more time in IMC tended to spend slightly more time < 500' AGL. Effect sizes were also modest, accounting for r2pb = 14, 18, and 20% of the variance, respectively.

**Durability of non-trivial relations.** A “durable” relation is one that remains statistically significant across both Phases 1 and 2. Table 10 shows these.

Table 10

*Durable, non-trivial relations between Phases 1 and 2 variables.*

Variable 1	Variable 2	
	Locality of Residence (1=Local) <sup>1</sup>	Pilot Age <sup>2</sup>
Web Preflight Duration	-.348 (.013) Ph1	.417 (.003) Ph1
	-.421 (.004) Ph2	.475 (.001) Ph2

<sup>1</sup> $r_{pb}$  = Point-biserial. <sup>2</sup> $r_s$  = Spearman rho. P-values are in parentheses.

Here, both local pilots and younger pilots spent slightly less time on their Web weather preflight briefing. Arguably, local pilots were more familiar with local terrain and weather patterns. Speculatively, older pilots could have been either slightly more careful briefers or might have simply been a bit less familiar with Web-based briefing, especially *aviationweather.gov*.

**Effect of pilot weather knowledge on subsequent flight safety.** As in Phase 1, weather knowledge (as measured by our questions) did not significantly predict flight safety. As Table 9 showed, average weather knowledge<sup>7</sup> did not relate significantly to any Phase 2 flight behavior variables.<sup>8</sup> Also as in Phase 1, neither was it significantly related to age, flight hours, locality of residence, or instrument rating.

**Effect of Web preflight briefing time on subsequent flight safety.** As in Phase 1, spending more time on the Phase 2 Web-based weather briefing did not lead to significantly safer flight. Table 9 reveals non-significant Spearman correlations of Web preflight duration with flight duration, minimum distance to ABQ, minutes scud running, minutes in IMC, and minutes < 500' AGL, with values ranging from  $-.076 \leq r_s \leq .109$ .

Phase 2 gave a more sophisticated sense of how pilots used the Web-emulation for self-briefing. It might seem logical to explore relations between, say, individual page view durations and our flight safety variables. However, there was arguably too much variability in the data to be able to do this confidently at this time.<sup>9</sup>

**Effect of the weather training products on takeoff hesitancy.** In Phase 1, the weather training products appeared to induce takeoff hesitancy. Twelve of 50 pilots initially stated that having to fly this mission VFR, they would choose not to even take off. In Phase 2, seven of the 44 returning pilots made the same decision.

Overall, this decrease was not significant ( $p_{\chi^2} = .330$ , NS). Phase 1 and Phase 2 decisions remained significantly correlated<sup>10</sup> ( $r_s = .323$ ,  $p = .032$ ), meaning that most pilots repeated their Phase 1 initial takeoff decision. Table 11 shows that

---

<sup>7</sup> (pre-+post-test+final score)/3.

<sup>8</sup> Spearman correlations with flight duration, minimum distance to ABQ, minutes scud running, minutes in IMC, and minutes < 500' AGL ranged from  $-.260 \leq r_s \leq .07$ , all NS.

<sup>9</sup> First, the variation in numbers of page views-per-dependent variable (DV) was enormous (range 1-92, mean 44.8, SD 28.0), meaning that correlations and models would either be based on wildly different numbers of cases or would be saddled with huge numbers of zero values. Second, the frequency distributions for the 18 Web pages' durations were, without exception, unacceptably non-normal for parametric techniques. Even excluding non-zero values, all Shapiro-Wilk  $p$ s were  $\leq .001$  except Collaborative Convective Forecast Product (CCFP) = .011 and Convective Outlook-Wind = .183 (but, which was based only on  $n=3$ ). Currently, there is no widely accepted method of nonparametric multiple regression. In short, we would mistrust the results.

<sup>10</sup> The correlation used here was  $\phi$  (phi), which measures the relation between two dichotomous variables (in this case, Phase 1 hesitancy yes/no vs. Phase 2 hesitancy yes/no).

4+30=34 pilots repeated their decision, while 3+7=10 (23%) reversed their decision. Notably, 7 of those 10 (70%) were pilots who formerly did *not* want to take off, who now *did* want to take off, even though the flight situation was essentially identical to Phase 1.

Table 12 compares numbers of Phase 1 and Phase 2 pilots who initially hesitated versus the values expected by chance (in parentheses). In Phase 1, we saw *lack of hesitancy* significantly concentrated in the Control group. In Phase 2, it became more uniform across groups ( $p_{\chi^2} = .554$ , NS).

Table 11

*Takeoff decision across phase for 44 pilots who participated in both Phase 1 and 2 (“Yes” means “Would take off.”)*

		Phase 2	
		No	Yes
Phase 1	No	4 (1.8)	7 (9.3)
	Yes	3 (5.3)	30 (27.8)

1<sup>st</sup> number is actual n (2<sup>nd</sup> is expected n)

Table 12

*Takeoff hesitancy, Phase 1 vs. 2.*

	Phase 1			Phase 2			
	Trg Prod 1	Trg Prod 2	Control	Trg Prod 1	Trg Prod 2	Control	
Initial takeoff decision	Yes	12 (12.2)	9 (12.2)	17 (13.7)	12 (12.6)	12 (12.6)	13 (11.8)
	No	4 (3.8)	7 (3.8)	1 (4.3)	3 (2.4)	3 (2.4)	1 (2.2)
Pairwise odds-ratios, 1-tailed <i>p</i>	← .152 →			← .004 →			NS
	← .037 →						

All this implies that training product-induced takeoff hesitancy did not persist over time. This is important because Phase 1 evidence for effect of training products was sparse and rested on the effect of our 3-variable model (consisting of training product + pilot age + takeoff hesitancy) to predict whether or not pilots would complete the entire flight to ABQ. Were hesitancy to cease to exert an effect, that might obviate that model.

**Effect of takeoff hesitancy on subsequent flight safety.** The seven Phase 2 hesitators did not end up flying more safely than the remaining 37 pilots. No significant differences between hesitators and non-hesitators were seen for minutes spent in IMC, minutes scud running, or minutes < 500' AGL (2-tailed Mann-Whitney  $p_U = .268, .089, .950$  respectively, all NS). In Phase 1, hesitators seemed to continue their conservatism into their flight, making significantly briefer flights, with consequently less penetration into the marginal weather close to ABQ. This was not true in Phase 2 ( $p_U = .550, .450$ , respectively, NS).

**Effect of the weather training products on subsequent flight safety.** When all was said and done, we wanted to know if viewing a weather training product would significantly affect flight safety. In Phase 1, there was *indirect* indication of this. Seeing a weather training video related to takeoff hesitancy, which related to flight duration, which related to minutes spent < 500' AGL—although seeing the weather video did not significantly *directly* relate to minutes spent < 500' AGL (nor to scud running or time spent in IMC).

However, in Phase 2, the same indirect correlational chain did not seem at work. Nor could any direct relation be seen between training product and subsequent flight safety, as measured by flight duration, minimum distance to ABQ, minutes in IMC, minutes scud running, or minutes < 500' AGL ( $.154 < p_{Kruskal-Wallis} < .768$ , NS).

### Modeling Flight Behavior

To investigate why some pilots chose to complete the flight through deteriorating weather while others did not, we again constructed models. The bivariate correlations in Table 9 represent the simplest possible kind of logical model. However, real-world events are rarely well-explained by a single factor. So, to try to get at multi-factor relations, we turned to multivariate modeling.

**Cluster analysis.** In Phase 1, a subset of the candidate variables formed two significant similarity clusters. Table 13 reiterates these.

In Phase 2, however, a repeat cluster analysis failed to find any variables related sufficiently to sort the pilots into even two clusters. The logical significance of this will become apparent shortly.

Table 13

#### *Phase 1 variables contributing significantly to clustering*

Continuous	Categorical
Age	Training product
Flight hours	Takeoff hesitancy
Final minimum distance to ABQ	Wx recheck just before takeoff
Minutes flying < 500' AGL	Flew all the way to ABQ
Clustering	

Cluster 1 tendencies	Cluster 2 tendencies
Younger	Older
Lower flight hr	Higher flight hr
Closer final minimum dist to ABQ	Farther final minimum dist to ABQ
More minutes at < 500' AGL	Fewer minutes at < 500' AGL
Control group (no wx trg product)	Received a wx training product
Initial takeoff response was to fly	Initial takeoff response was to not fly
No wx check just before takeoff	Wx check just before takeoff
Flew all the way to ABQ	Diverted before ABQ

**Binary logistic regression analysis.** We next retested the Phase 1 3-variable logistic regression model for durability. In Phase 1, a model based on training product, age, and takeoff decision<sup>11</sup> was able to predict 64% of the explainable (Nagelkerke) variance in *ToABQ* ( $p=.000004$ ) and correctly predicted 83.3% of the cases (diversion vs. continuation on to ABQ), compared to a baseline prediction rate of 62.5%.<sup>12</sup>

In retesting this model with Phase 2 data, however, the identical model predicted only 19.9% of the Nagelkerke variance and 72.1% of the cases, compared to its baseline rate of 60.5%. This was not a significant improvement ( $p=.145$ , NS) over an educated guess (i.e., the baseline, “constant-only” model).

This performance degradation of the Phase 1 model was certainly not simply due to the raw number of pilots who actually made it all the way to ABQ (18 of 48<sup>13</sup> in Phase 1 vs. 17 of 43 in Phase 2). Nor was it due to pilot age (because pilots were only 3-4 months older than they were during Phase 1).

We speculate that the Phase 1 model collapsed due to inconsistencies in takeoff decision from Phase 1 to Phase 2. Whatever coherent effect, or “signal,” the weather training products may have engendered in Phase 1 dissolved amongst the “noise” of individual variation in Phase 2. Table 9 showed the 2x2 consistency matrix. Ten pilots (3+7) reversed their Phase 1 takeoff decisions in Phase 2. While Fisher’s Exact Test gives this probability at  $p=.054$ —technically non-significant—it is arguably close enough to suspect that we have the culprit that disabled the Phase 1 3-variable model.

<sup>11</sup> Note that takeoff decision reflects “takeoff hesitancy” as discussed earlier.

<sup>12</sup> The baseline rate is the model’s ability to predict an outcome by chance alone, given only knowledge of average group behavior. For instance, if 50% of Americans voted Democratic, that would be the baseline rate, assuming we had no other knowledge about individual voters. However, if we knew individual voters’ incomes, educational levels, ethnicities, genders, religious preferences, and job categories, we might expect to predict individual votes at greater than a 50% success rate. The primary purpose of multivariate modeling is to maximize that kind of additional predictability.

<sup>13</sup> The original Phase 1 N=50, with 1 missing data, 1 eliminated for CFIT → 48. The original Phase 2 N=44, with 1 eliminated for CFIT → 43.



After this admittedly exhausting analysis of the details, we now turn to a summary discussion of the overall experiment.

### Discussion

The primary purpose of this research was to investigate the effects of video weather training products, namely a) their immediate effects on pilot weather knowledge and flight behavior in the face of potential instrument meteorological conditions, and b) whether these effects persist across time.

Fifty GA pilots participated in a study designed to examine both pilot weather knowledge and flight behavior. Pilots took a weather knowledge pre-test, followed by exposure either to one of two weather training videos (the Experimental groups), or to a video having nothing to do with weather (the Control group). They then took a knowledge post-test to measure knowledge gain induced by the training product. Next, they planned for, and flew, a simulated flight mission through deteriorating weather from Amarillo, TX, to Albuquerque, NM (ABQ). Numerical flight data were collected and flight behaviors noted.

Figures 7 and 8 graphically depict significant correlations between variables. Single-headed arrows imply directional causation (e.g. instrument rating could conceivably cause minutes < 500' AGL to vary, but the reverse would not be true). Double-headed arrows make no assumption about what might cause what.

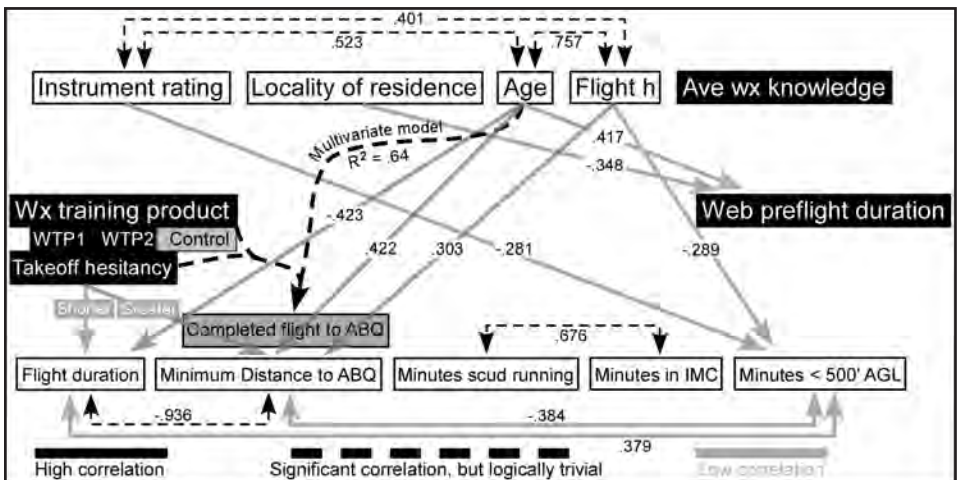


Figure 7. Phase 1 univariate and multivariate correlational structure.

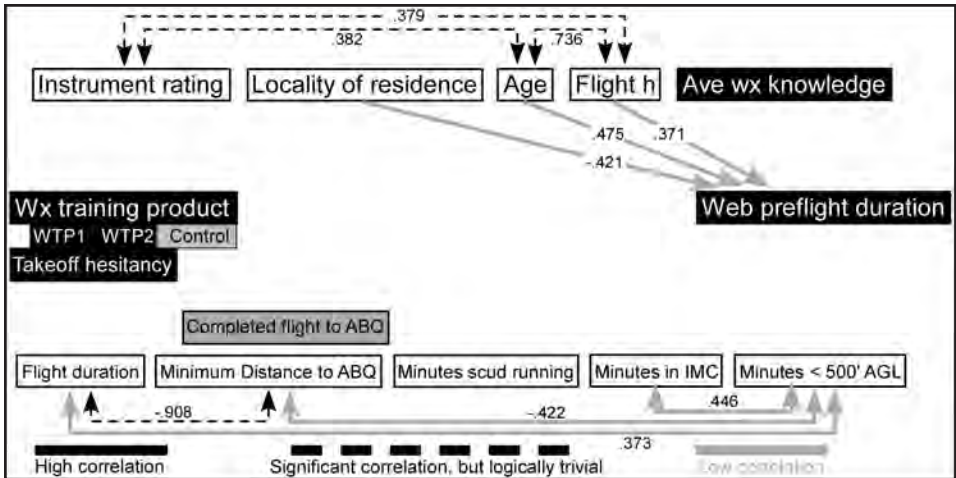


Figure 8. Phase 2 correlational structure.

In Phase 1, only a limited number of significant training effects were seen. For one, there was a tendency for pilots who viewed either of the two weather videos to hesitate taking off into the marginal weather. These “hesitators” flew only after encouragement. In contrast, 17 of 18 control group pilots took off without any encouragement. Subsequently, the hesitators continued their conservatism, making shorter average flights than non-hesitators.

Since the bulk of the scenario’s danger lay near the flight’s destination, we might speculate that watching a weather training video induced *takeoff hesitancy*, which then induced shorter flights, which then led to lower groupwise risk exposure. There was correlational evidence to support each individual link of this chain. However, the overall chain of logic was not simple. Ultimately, no beginning-state variables (e.g., pilot age, flight hours) ended up *directly* correlating with end-state flight-risk variables (scud running, time spent in IMC, or time spent at < 500’ AGL). Therefore, no simple model based on video training product alone could be ultimately shown to modulate flight risk.

So, we turned to multivariate modeling. A binary logistic regression model selected from multiple candidate variables (continuous and discrete) to find the best combination capable of explaining the variation in a single, discrete dependent variable—whether or not a pilot completed the entire flight to ABQ. This model assumed that the farther one flew into the deteriorating weather, the greater the overall risk.

In Phase 1, this led to a significant 3-variable model of weather-related risk taking based on

- Video training product
- Pilot age
- Takeoff hesitancy

This model implied that, while a brief video weather training product *alone* might not significantly affect risk taking, it might do so if combined with other factors. Specifically, relatively older pilots with more conservative behavioral (flying) tendencies may have been *sensitized*—but less by a specific training product than by the more general fact of being exposed to “something safety-related.” In technical terms, older, more cautious pilots in the two experimental groups might be more susceptible to cognitive priming than those in the control group. In lay terms, experimental-group pilots may have suspected that the FAA had an agenda about weather and that they had better act cautiously.

However interesting this model, it did not hold up in Phase 2. There, pilots were under less pressure: They had fewer tests to take; they were already familiar with equipment and procedure, and now fully understood that the FAA researchers were benign. They could relax and act naturally. They desensitized. And, when they did, it became hard to discern any significant differences between the weather training groups and the Control group.

Note that this does not mean that weather training products are ineffectual. It merely means we failed to demonstrate effect. The fact remains that weather is complex and we cannot expect one 90-minute training session to change much. Just as we cannot build an entire house from a single brick—no matter how good the brick—we probably cannot radically and permanently alter pilot behavior in a lasting way from a single brief training session.

## Findings common to both studies

**Durable relations.** A few non-trivial durable relations persisted from Phase 1 to 2. There was a slight tendency for older pilots, and for Oklahoma pilots, to spend a bit more time on their Web-based preflight briefing. We can speculate that older pilots were likely to be slightly less familiar with Web-based preflight briefing, and that local pilots tended to know the terrain better.

**Consistency of flight behavior.** Faced with a similar situation later in time, pilots tend to repeat what they did earlier. Table 8 showed that in Phase 2, more than 78% of pilots made the same ultimate choice about either diverting or continuing on to the destination that they first made in Phase 1. Table 11 showed that 77% of initial yes/no takeoff decisions remained the same.

## Acknowledgments

Research question 1 was conceived and promoted by Scott Shappell and William Krebs. The authors gratefully acknowledge the invaluable assistance of Stanley Roberts, Manager, AFS-630, for providing item difficulties to FAA test questions. Janine King, Sally Glasgow, and Suzanne Thomas of Xyant Technology, Inc., and Tammy Harris, FAA, provided key scheduling and subject payment support. The authors extend grateful thanks to both the individuals who provided the training products and to the pilot participants, without whom this study would not have been possible. Finally, we thank our sponsor, Mike Lenz, FAA Headquarters, Washington, DC, for his support and intellectual guidance throughout this project.

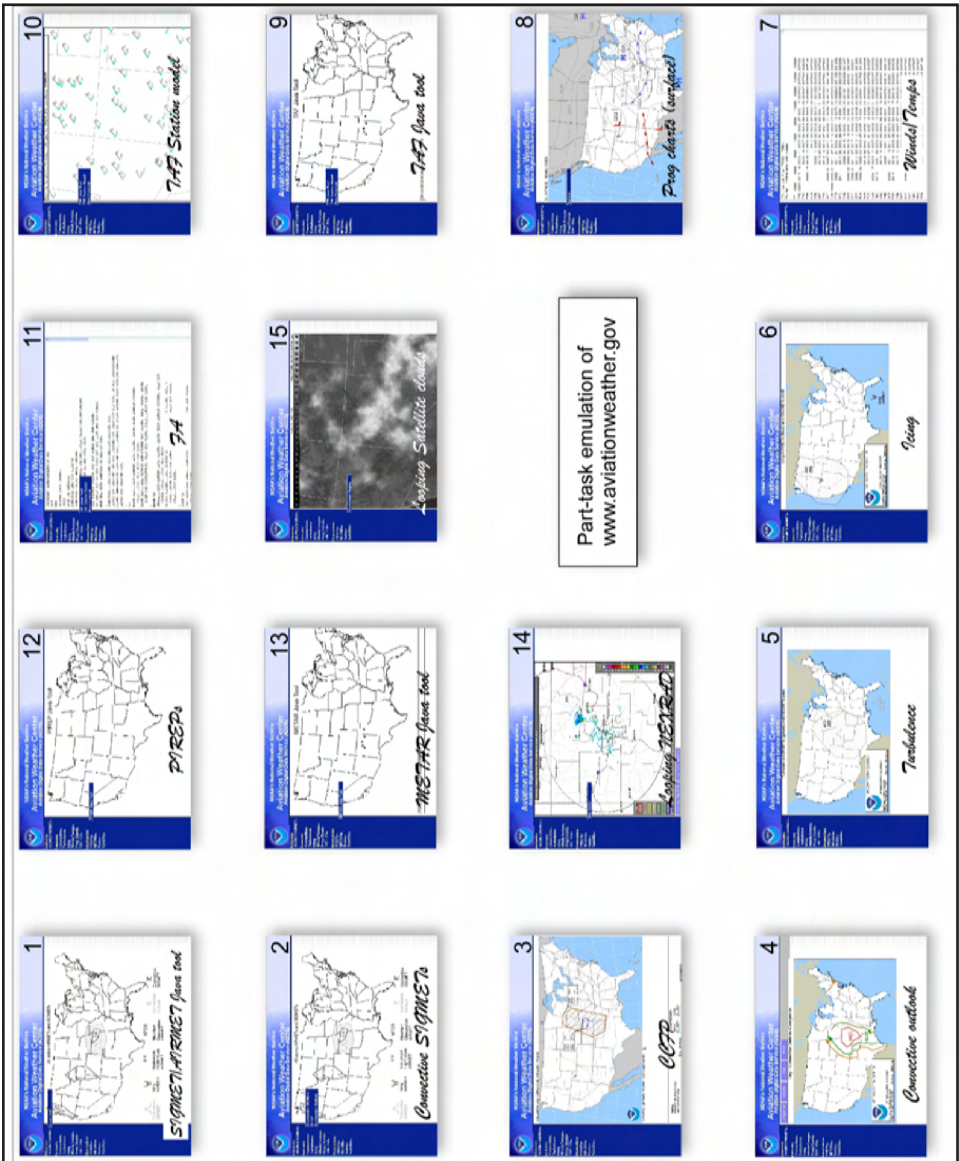
## References

- Adams, R.J., & Ericsson, K.A. (1992). Introduction to cognitive processes of expert pilots. (Technical Report DOT/FAA/RD-92/12). Washington, DC: Federal Aviation Administration, Research and Development Service.
- Bazargan, M. (2005). Analysis of national top 10 first causes of fatal accidents of general aviation between 1983 and 2002. Part 1-national-all aircraft types. (Technical Report 61022). Daytona Beach, FL: Embry-Riddle Aeronautical University, Center for General Aviation Research.
- Bud, M.J., Mengert, P., Ransom, S., & Stearns, M.D. (1997). General aviation accidents, 1983-1994: Identification of factors related to controlled flight into terrain (CFIT) accidents. (Technical Report DOT/FAA/AAR 100-97-2). Washington, DC: Federal Aviation Administration, Office of Aviation Research.
- Ellis, J.A., Semb, G.B., & Cole, B. (1998). Very long-term memory for information taught in school. *Contemporary Educational Psychology*, 23, 419-33.
- FAA (2008). 2008 airmen knowledge tests. Downloaded 1 Jun, 2009 from [http://www.faa.gov/data\\_research/aviation\\_data\\_statistics/test\\_statistics/media/2008/annual/2008\\_Annual\\_Statistics\\_for\\_Airmen\\_Knowledge\\_Testing.pdf](http://www.faa.gov/data_research/aviation_data_statistics/test_statistics/media/2008/annual/2008_Annual_Statistics_for_Airmen_Knowledge_Testing.pdf).
- Goldstein, E.B. (1999). *Sensation & Perception* (5<sup>th</sup> Ed.). New York: Brooks/Cole.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: M.I.T. Press.
- Knecht, W.R., Ball, J., & Lenz, M. (2010a). *Effects of video weather training products, Web-based preflight weather briefing, and local versus non-local pilots on general aviation pilot weather knowledge and flight behavior, Phase 1*. (Technical Report DOT/FAA/AM-10/1). Washington, DC. Office of Aerospace Medicine.
- Knecht, W.R., Ball, J., & Lenz, M. (2010b). *Effects of video weather training products, Web-based preflight weather briefing, and local versus non-local pilots on general aviation pilot weather knowledge and flight behavior, Phase 2*. (Technical Report DOT/FAA/AM-10/6). Washington, DC. Office of Aerospace Medicine.
- Mackintosh, N.J. (1974). *The psychology of animal learning*. New York: Academic Press.
- Milgram, S. (1974). *Obedience to Authority: An Experimental View*. New York: Harper and Row.
- National Geophysical Data Center (2008). *NOAA-NGDC-MGG-GLOBE Custom Data Selection Page*. Digital elevation data downloaded Oct. 23, 2008 from <http://www.ngdc.noaa.gov/cgi-bin/mgg/ff/nph-newform.pl/mgg/topo/custom-datacd>.
- NTSB. (1989). *Safety report: General aviation accidents involving visual flight rules flight into instrument meteorological conditions* (Technical Report NTSB/SR-89/01). Washington, DC: National Transportation Safety Board.
- NTSB. (2005). *Risk factors associated with weather-related general aviation accidents* (Technical Report NTSB/SS-05/01). Washington, DC: National Transportation Safety Board.
- Perkins, D.M. (1992). *Transfer of learning*. Downloaded 25 Nov, 2008 from <http://learnweb.harvard.edu/alps/thinking/docs/traencyn.htm>.

- Semb, G.B., & Ellis, J.A. (1994). Knowledge taught in schools: What is remembered? *Review of Educational Research*, 64, 253-86.
- Statistical Package for the Social Sciences. (2001). *The SPSS two-step cluster component*. SPSS Technical Report TSCWP-0101.
- Tabachnick, B.G. & Fidell, L.S. (2001). Using multivariate statistics (4<sup>th</sup> Ed.). Boston: Allyn & Bacon.
- Wiggins, M. & O'Hare, D. (2003). Expert and novice pilot perceptions of static in-flight images of weather. *International Journal of Aviation Psychology*, 13, 173-87.
- Wolfram Research (2008). *Mathematica V7.0*.

# APPENDIX

## Web preflight briefing screenshots







## ***MD-11 Landing Incidents/Accidents 1992 – 2009:***

### ***A Textual Analysis of NTSB and FAA Reports***

John T. Cocklin

*Government Information Librarian  
Dartmouth College*

#### ***Abstract***

*In 2009, a MD-11 crashed at Narita, Japan with the loss of two lives. This was the third MD-11 rollover, a rare occurrence for a wide-bodied airplane. It was also the latest in a line of landing events with similar characteristics. After the 1997 rollover of a MD-11, the 2000 rollover of a China Airlines MD-11, and other aircraft landing events, the NTSB asked FAA to sponsor a comparative study of stability and control characteristics of large transport category airplanes. This paper provides the results of a textual analysis of MD-11 incident/accident public reports from 1992-2009. Three distinct types of landing events were discovered. The majority were type one with shared characteristics to the rollover accidents. Type two involved runway overruns with none of the defining variables found in type one. Type three involved the main landing gear and, as with type two, had no apparent relation to the variables found in type one. The paper discusses the MD-11's relaxed stability design and related pitch sensitivity. The paper also reviews possible improvements in incident/accident reporting.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

## **MD-11 Landing Incidents/Accidents 1992 – 2009:**

### **A Textual Analysis of NTSB and FAA Reports**

A wide-bodied airplane rollover on landing is a rare occurrence. It has happened three times to the MD-11 in the midst of a series of MD-11 landing events. The first landing rollover occurred to a freighter landing at Newark, New Jersey in 1997 (FedEx 14), and a passenger MD-11 rolled over in 1999 at Hong Kong with the loss of three lives (China Airlines 642). While the number of MD-11 landing events slowed after 2000, they continued, and in 2009 an MD-11 rolled over on landing at Narita, Japan with the loss of two lives (FedEx 80). In response to FedEx 14 and other landing events involving MD-11s, DC-10s, and B757/767s, the National Transportation Board (NTSB) in July 2000 recommended that the Federal Aviation Administration (FAA) sponsor a National Aeronautics and Space Administration (NASA) study of the “stability and control characteristics of widely used, large transport category airplanes.” The report was to have two goals:

1. Identify undesirable characteristics that may develop during the landing phase in the presence of adverse combinations of pilot control inputs, airplane center of gravity position, atmospheric conditions, and other factors.
2. Compare overall qualitative and quantitative stability and control characteristics on an objective basis (Safety Recommendation A-00-100. National Transportation Safety Board [NTSB], 2000a; NTSB, 2000b).

The study was never initiated. Instead, the FAA cited two previous comprehensive landing studies that reviewed multiple variables but did not attempt any comparison between aircraft models (Approach and Landing Joint Safety Analysis Team [JSAT], 1999; Khatwa, 1999). “Based on the reports...the FAA does not believe that basic research based on past accident reports will identify any undesirable landing phase combinations that are directly related to stability and control characteristics.” The NTSB classified the safety recommendation “Closed-Acceptable Alternate Action” (NTSB, 2001b).

The MD-11 does have a high hull loss accident rate per million departures, 2.57, when compared with other airplane types including its two direct competitors, the Airbus A340 at 1.05 and the Boeing B777 at 0.24 (Boeing, 2009).

However, no objective comparison of wide-bodied airplane landing events has ever been completed, and no objective comparisons can be made. This paper attempts to partially address the NTSB recommendation by analyzing on an objective basis through the textual analysis of incident and accident reports the stability and control characteristics of the MD-11 during the landing phase of publicly reported events since 1992. The most significant MD-11 accident involving loss of life was not included in this analysis due to the unique nature of that accident (Swissair 111 with 229 fatalities). According to the Transportation Safety Board of Canada, a fire started via a short involving the in-flight entertainment system wiring installed after the airplane was delivered to Swissair (Transportation Safety Board of Canada [TSB], 2003).

# Accident Rates by Airplane Type

## Hull Loss Accidents – Worldwide Commercial Jet Fleet – 1959 Through 2008

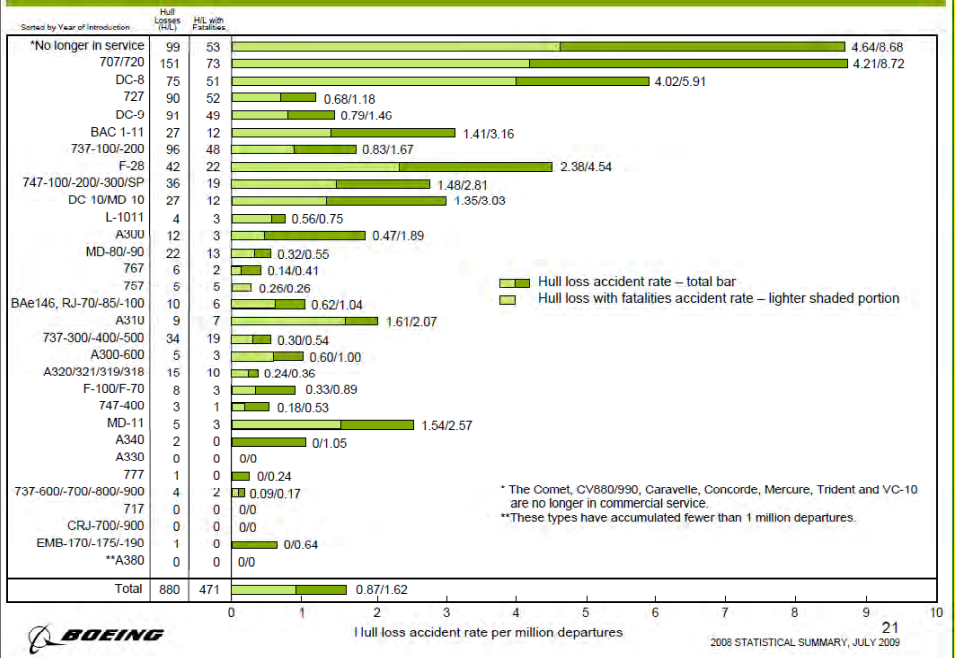


Figure 1. Boeing (2009).

### Methodology

By examining the publicly available incident/accident reports on MD-11 landing events contained in NTSB and FAA databases, it was theorized that patterns would emerge. The earliest incident contained in the databases was a Delta flight (number not recorded) from August 2, 1992. The latest one analyzed for this paper was a Saudi Arabian flight (number not recorded) from June 9, 2009. Two Lufthansa Cargo landing events (one a hull loss accident) occurred after this date and were not included in the analysis for this paper.

NTSB Aviation Incident/Accident Reports, Factual Reports, Probable Cause Reports, and FAA Reports were analyzed and compared with a list of variables (NTSB Aviation Accident Database and Synopses; FAA Accident/Incident Data System [AIDS]). If the incident/accident report explicitly described that variable, for example a go-around, then the event received a score of one for that variable. If the variable was not explicitly described then the event received a score of zero for that variable. The two exceptions of this scoring were firm or hard landings (1 for firm, 2 for hard) and landing gear damage (1 for damaged, 2 for destroyed). In some cases it was tempting to infer a conclusion based on the available information. For example, in the Eva Air flight from November 20, 2001, the NTSB report states "Hard landing...with nose gear hitting ground" (NTSB, 2001a). Based on similar events, a researcher might be tempted to add high sink rate or pitch-up on spoiler deployment as possibilities for this event. However, the report states only a hard landing, and that is what was counted. High sink rate and pitch up were given 0's for that event.

Other sources, including the Aviation Safety Network database, and reports from the Hong Kong and Ireland investigation boards were used to provide more detailed information on certain events when that information was not available from either the FAA or the NTSB (Aviation Safety Network Database <http://aviation-safety.net/database/>; Civil Aviation Department Hong Kong [CAD], 2004; Irish Air Accident Investigation Unit [AAIU], 2001). FedEx 80 is still under investigation, but video and news reports from that event were used to provide information on the event when the information was clear and without question (Aviation Week & Space Technology, 2009; "FedEx Narita Japan," 2009).

In addition to descriptive information such as date, time, damage, and weather, categories based on the three MD-11 hull loss accidents were created. The accident reports of the first two and the video of the third accident helped to inform these categories. The following categories were initially used:

- oscillation
- bounce
- high sink rate
- go-around
- hard landing
- landing gear damage
- sheared wing
- rollover
- pilot error

As anticipated, it became clear during the analysis that other categories needed to be added.

- 1<sup>st</sup> or 2<sup>nd</sup> landing = This variable helped determine if the landing event precipitated the go-around, if damage was partially or totally caused by the go-around maneuver, or if the damage was caused on the second landing attempt.
- tailstrike = Early on it became apparent that a large percentage of the landing events involved tailstrikes.
- firm landing = Initially only hard landings were noted. However, some reports labeled the landings as firm (based usually on eyewitness accounts) so this second variable was created.
- pitch up = The MD-11 has a tendency to pitch up on landing. While the cause of this pitch up has never been clearly diagnosed, enough of the incident/accident reports mentioned pitch up to warrant the inclusion of this variable.
- pilot reported wind shear = Specifically, this variable was created to indicate when a pilot reported wind shear but there was no indication of actual wind shear either from other pilots or from wind shear instrumentation.

Through the course of the analysis, three distinct types of landing events emerged.

- Type one = Landings involving some combination of a firm/hard landing, tailstrike, and/or high sink rate (21 events reported).
- Type two = Landings involving runway overruns with no indication of the type one defining variables. (7 events reported – three of them under wet or rainy conditions.)
- Type three = A group of events involving the main landing gear with no indication of the defining type one variables (4 events reported).

This paper focuses on type one events. No explicit link was found between this first broad type and the other two.

UPS 6971 suffered a nose gear collapse on landing. A minor bounce of the nose gear occurred, but otherwise the flight encountered none of the other variables typical for a type one event. It was included, however, because the characteristics of the event otherwise fell into this type. In addition, the aircraft had encountered a hard landing on its previous landing with a different flight crew making it of further interest to this study.

As one commonly discovers in textual analysis, terminology was used inconsistently among incident/accident reports. Specifically, in some cases the terms firm and hard landing were used interchangeably. The definition for each was unclear and was often determined by the perception of the airplane occupants. That is why an explicit mention was required. The explicit rule in some ways makes this analysis possible, and it is believed, leads to useful conclusions. However, some nuances are lost. As an example, this researcher's opinion is that oscillations occurred in more events than listed, but it was only explicitly mentioned or seen in three events.

One major caveat has to be clearly outlined. As the seeming severity of the event decreased, less of an investigation was completed and less information was given in the incident/accident report. In addition, several of these events occurred outside the United States. While the NTSB was made a party to or notified of some of the investigations, occasionally a fuller report was not issued by that country's investigation body. In one case where it was, an Eva Air flight from November 20, 2001, the accident occurred in Taiwan, and the final accident report was written in Chinese (Taiwan Aviation Safety Council [ASC], 2001). Translation of this report into English was beyond the resources available for this research project.

Of the twenty-one events falling into type one, the information for eight were deemed incomplete by the author. This gave the researcher two options – drop the eight from the analysis or include them with this caveat. The latter course was chosen. After careful review, it was determined that enough information was available to include them in the analysis. One of these is still under investigation (FedEx 80). For the others, the major factors were outlined in preliminary reports.

## Results

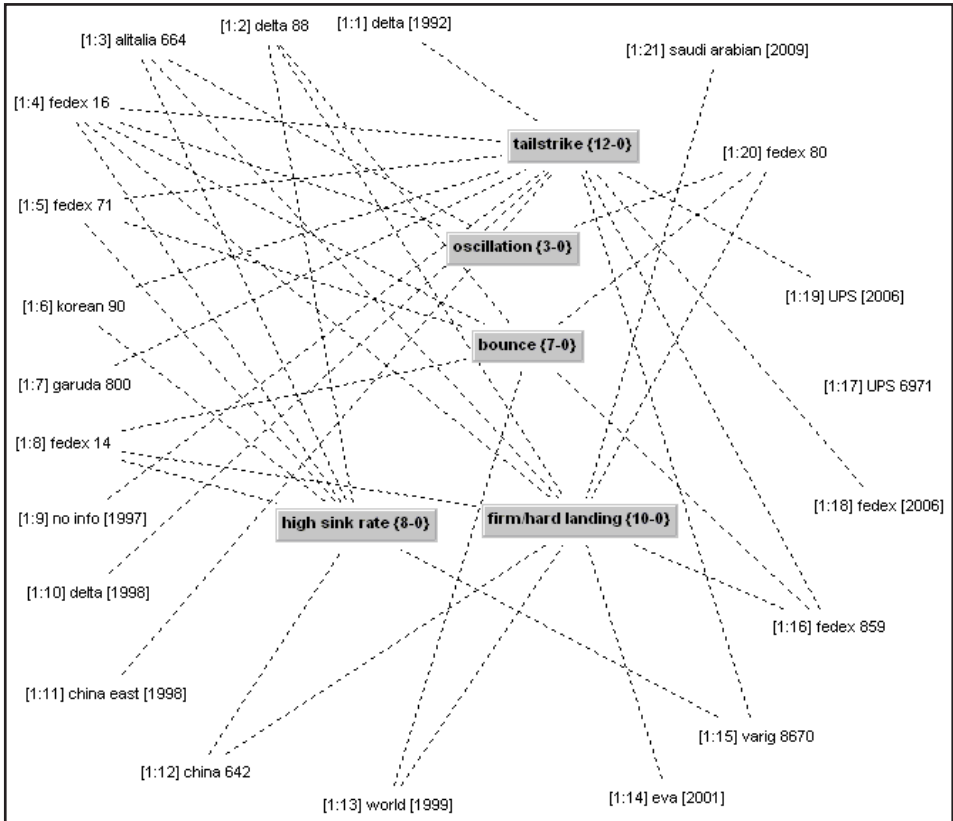


Figure 2. MD-11 type one landing incidents/accidents in chronological order (counter-clockwise starting from top). 1992 - 1999 on left; 2000 - 2009 on right.

Table 1

### Type one Landing Incidents/Accidents - Variables

MD-11 Type One Landing Incidents/Accidents - Variables 1992-2009		
variables	total (21)	hull loss (3)
visual conditions	15	2
pilot error	13	2
tailstrike	12	0
high sink rate	8	2
bounce	7	2
hard landing	7	2
fedex	6	2

pitch up	6	0
go-around	5	0
wind gusts 30+	3	2
oscillation	3	1
firm landing	3	1
sheared wing	3	3
rollover	3	3
reported wind shear by pilot	3	1

In the methodology section above it was noted that the landing events fell into three broad categories.

- Type one = Landings involving some combination of a high sink rate, firm/hard landing, and/or tailstrike (21 events reported).
- Type two = Runway overruns with no indication of the type one defining variables. (7 events reported – three of them under wet or rainy conditions.)
- Type three = A small group of events involving the main landing gear but with no indication of the type one defining variables (4 events reported). Three involved the center main landing gear and failure of the drag brace. The first type three occurrence in the databases, from September, 16, 1992, was actually a McDonnell Douglas test flight trying to determine a cause for two previous such incidents. McDonnell Douglas did issue a Service Bulletin with modifications, but two of the three occurred after the Service Bulletin and with the modifications. The latest one in the databases was dated November 7, 2006.

The majority of MD-11 landing events between 1992 and 2009, including the three hull loss accidents, were type one. The available information gave no readily apparent connection between the events in Types Two and Three and those in type one. For these reasons, type one is the focus of this paper and the following analysis.

The events primarily occurred in good weather conditions and fifteen were under visual conditions. Three occurred with gusts of 30+ knots and two of these three were hull loss events. The pilots in one of these were attempting a landing during Severe Tropical Storm Sam (China Airlines 642).

Thirteen involved some level of pilot error. These thirteen exactly correlate with the thirteen deemed “complete” meaning that these events warranted fuller investigations from the NTSB, CAD, and AAIU and thus fuller reports.

Tailstrikes played a role in twelve of the events. In three tailstrikes, pilots reported experiencing wind shear, when no other pilots (either in the vicinity or landing previously) reported wind shear and no indication of it was reflected in the instrumentation. Two of these reports explicitly stated excessive sink rates on landing (FedEx 71 and FedEx 14). The MD-11’s propensity to pitch up at ground spoiler deployment (a trait inherited from the DC-10 from which it was derived) played a role in at least six tailstrikes. However, the fundamental problem underlying the tailstrikes was outlined in a June 1996 FedEx Tail Strike Awareness



Training Instructors Guide. "One consistent factor in every landing tail strike to date [underlined in original] has been an excessive descent rate with an increasing attitude rate prior to the initial touchdown" (NTSB, 2000a, Appendix F, page 133).

This Guide was part of FedEx's Tail Strike Awareness Training Program unveiled in response to the FedEx 71 incident of 1996. The direct impact of this training program in its first four years was negligible – 8 of the 12 tailstrikes occurred after 1996. However, when combined with other developments discussed below the impact on tailstrikes after 1999 was clearer with only 4 of the 12 occurring between 2000 and 2009.

Ten of the type one landings were considered firm or hard landings (three firm, seven hard). These led to six bounces and three oscillations. FedEx 71 was unique in reporting a bounce without a firm or hard landing – all other instances of a bounce were tied to either a firm or hard landing. FedEx 71 did, however, experience a high rate of descent. Once the MD-11 started bouncing or oscillating on landing a potentially dangerous situation presented itself as when FedEx 14 and China Airlines 642 rolled over. The FedEx 14 accident report, with information from the China Airlines investigation, was released July 25, 2000. Among the NTSB recommendations were the following.

- Improved pilot training that would include structural failure awareness, a syllabus for landing simulator training, and a pro-active emphasis on go-arounds.
- An upgrade to Flight Controller Computer software (FCC-908) software on all MD-11's within one year (NTSB, 2000a; NTSB, 2000b).

These recommendations had an impact. Between 1992 and 1999, MD-11 type one landing events averaged 1.6 per year with a run of two in 1996, three in 1997, and two in 1998 and 1999. After 1999 the number of type one landing events per year fell to an average of 0.8 and there were stretches when none were reported to the FAA or NTSB (2000, 2002-2003, and 2007-2008).

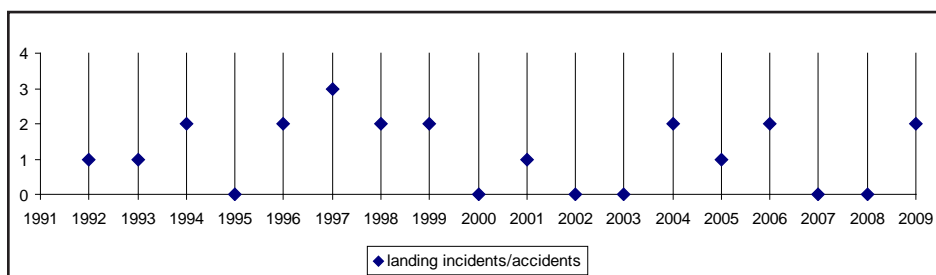


Figure 2. MD-11 Type One Landing Incidents/Accidents by Year (1992-2009).

Table 2

*MD-11 Type One Landing Incidents/Accidents - Descriptions (1992 - 2009).*

year	month	date	carrier	flight	reg	serial	passenger/ freight
1992	8	2	delta	N/A	N803DE	48474	p
1993	4	30	delta	88	N803DE	48474	p
1994	8	19	alitalia	664	IDUPO	48429	p
1994	11	4	fedex	16	N611FE	48604	f
1996	5	16	fedex	71	N614FE	48528	f
1996	5	25	korean airlines	90	HL737	48409	f
1997	6	21	garuda indonesia	800	EICDK	48501	p
1997	7	31	fedex	14	N611FE	48604	f
1997	10	24	N/A	N/A	N275WA	48631	f
1998	11	11	delta	N/A	N801DE	48565	p
1998	12	25	china eastern	N/A	B2174	N/A	p
1999	8	23	china airlines	642	N/A	N/A	p
1999	9	18	world airways	N/A	N272WA	48437	p
2001	11	20	eva	N/A	B16101	N/A	p
2004	1	26	varig	8670	PPVTF	N/A	p
2004	9	19	fedex	859	N601FE	48401	f
2005	6	7	UPS	6971	N250UP	48745	f
2006	9	14	fedex	N/A	N623FE	N/A	f
2006	11	14	UPS	N/A	N282UP	N/A	f
2009	3	23	fedex	80	N526FE	N/A	f
2009	6	9	saudi arabian	N/A	HZ-ANB	N/A	f

*Gray = Rollover.*

Table 3

*MD-11 Type Two Landing Incidents/Accidents - Descriptions (1992 - 2009).*

Year	Month	Date	Carrier	Flight	Reg	Serial	Passenger/ Freight
1996	11	6	world airways	201	N271W	48518	p
1998	6	14	japan air lines	N/A	JA8580	48571	p
1999	8	8	china eastern	N/A	N/A	N/A	p
1999	10	17	fedex	N/A	N582FE	N/A	f
2002	2	3	delta	N/A	N803DE	N/A	p
2003	5	30	gemini	N/A	N703GC	48411	f
2006	2	11	world airways	N/A	272WA	N/A	p

Table 4

*MD-11 Type three Landing Incidents/Accidents - Descriptions (1992 - 2009).*

year	month	date	carrier	flight	reg	serial	passenger/ freight
1992	9	16	McDonnell Douglas	N/A	N90187	48437	p
1993	2	14	"foreign air carrier"	N/A	N/A	N/A	p
1995	10	5	fedex	1700	N606FE	48602	f
2006	11	7	fedex	N/A	N584FE	N/A	f

### Discussion - Relaxed Stability

In addition to providing information about the variables involved in each event, the incident/accident reports also provided information about the MD-11's relaxed stability design and its impact on MD-11 pitch sensitivity. In the early 1970's McDonnell conducted research on the combination of fly-by-wire technology and relaxed stability to improve fighter performance for the Air Force Flight Dynamics Laboratory (Berger, 1973). By 1987, engineers on the Douglas side of the company were incorporating relaxed stability into their MD-11 design. To reduce both drag and weight they designed the MD-11 horizontal tail 30% smaller in area than the DC-10 tail from which it was derived. One way they accomplished this was by relaxing the static longitudinal stability to increase the aft limit of the center of gravity (CG). A 2,000 gallon fuel tank was also placed in the horizontal stabilizer to help maintain an aft CG. To decrease pilot workload and enhance stability in manual flight at aft CG's, a longitudinal stability augmentation system (LSAS) was integrated into the Flight Control Computer (FCC). The 30% reduction in horizon-

tal tail area resulted in a 2% reduction in total aircraft drag, 2,000 pounds of weight savings, and a net fuel-burn reduction of almost 3% (Field, 1987).

The NTSB first investigated MD-11 relaxed stability in connection with a series of high altitude oscillations. The NTSB found, and Douglas test pilots confirmed, that the control column forces required for longitudinal control of the MD-11 in high altitude manual flight were lighter than those required for other “transport category” airplanes with which the pilots might have been familiar. This pitch sensitivity led to pilot induced oscillations (PIO’s) in the cases of China Airlines 61-012 - 1992, China Eastern 583 - 1993, and eight similar events of which the NTSB was aware. The NTSB recommended a review of the MD-11’s longitudinal stability and related pilot training (NTSB, 1993, pp. 30-32 and 45-50; NTSB, 1994, pp. 19-23 and 26-31). In December 1995 McDonnell Douglas engineers responded by designing a software upgrade for the LSAS, FCC-907, which introduced a pitch rate damper (PRD) control law. As they were addressing a high altitude, high speed problem they decided to leave low altitude flight characteristics unchanged and had the PRD phase in/phase out between 15,000 and 20,000 feet (NTSB, 2000a).

The continued frequency and increased severity of MD-11 tailstrikes led to larger NTSB investigations, and McDonnell Douglas/Boeing engineers turned their attention to MD-11 landings. They developed software upgrade FCC-908 with a PRD that remained active at 30% strength from approximately 17,500 feet to ground level. The NTSB felt this upgrade would “render the airplane less susceptible to [pilot] overcontrol in pitch” as seen in such events as FedEx 14. The software also included features to help prevent the pitch-up that would occasionally occur when the spoilers deployed upon landing, a problem which dated back to the MD-11’s predecessor the DC-10. FCC-908 became FAA certified in May 2000, and in July 2000 the NTSB recommended all carriers load their MD-11 fleets with the new software within one year (NTSB, 2000a; NTSB, 2000b). After the FedEx 71 landing event in 1996, FedEx had implemented a tailstrike training awareness program with emphasis on maintaining proper sink rates, recovering from bounced landings, and flying low level go-arounds. In 2000 the NTSB expanded this to a recommendation for increased and improved landing training for pilots in general (NTSB, 2000a). These improvements in training when combined with the flight control software upgrades had an impact. Though MD-11 landing events with type one variables continued, their rate per year decreased after 2000 (see Results above).

## Conclusion

Textual analysis of MD-11 landing incident/accident reports between 1992 and 2009 revealed three distinct types. Type one, the most prevalent, involved some combination of a high sink rate, firm/hard landing, and/or tailstrike. Of all the variables, the twelve tailstrikes garnered the most attention from investigators and the manufacturer. However, a review of the reports clearly indicates that tailstrikes were the symptom of the MD-11’s propensity to incur an excessive sink rate on landing. FedEx outlined this connection in their 1996 Tail Strike Awareness Training Instructors Guide (NTSB, 2000a, Appendix F, page 133). Two design issues contributed to the flying and landing characteristics of the MD-11.

1. Relaxed stability from its comparatively aft center of gravity and small horizontal stabilizer (NTSB, 1993; NTSB, 2000a, pp.49-50 and 62-63).

2. Pitch up on spoiler deployment, a problem carried over from its DC-10 predecessor (NTSB, 2000a).

Other airplanes, such as the MD-11's direct competitor the B777, were also designed with relaxed stability. However, pilots encountered higher pitch sensitivity with the MD-11 than with other airplanes. The NTSB first mentioned this in reports involving MD-11 upsets at higher altitudes but later extended this concern to pilot "overcontrol in pitch" which occurred during the FedEx 14 landing accident. The resulting NTSB recommendations to improve landing training programs and the MD-11 flight controller computer software (FCC-908) had an impact. Though MD-11 landing events continued, their rate per year decreased after 1999.

FedEx accounted for two of the three rollovers and five of the twenty-one type one events, the largest of any single carrier. This might simply be a statistical association as FedEx has flown more MD-11s (58 as of March 2009) longer (since 1991) than other carriers. Further research is required. A comparison of FedEx maintenance, operation, and training procedures with other MD-11 freighter fleets would be informative. UPS' fleet of MD-11's had two type one events. Lufthansa Cargo MD-11's were involved in two landing events, one a hull loss accident, but these occurred after July 9, 2009 and were not included in the analysis for this paper. (NTSB and FAA databases; Bundesstelle für Flugunfalluntersuchung [BFU] <http://www.bfu-web.de/>).

Textual analysis also revealed potential areas of improvement for the public reporting of aviation incidents/accidents.

1. Standardized terminology. There needs to be further standardization of language among and between investigators and pilots. For example, hard and firm were sometimes used interchangeably for the same landing event. In some cases a reported bounce was really an oscillation and vice versa. Even the terms incident and accident, though defined by the FAA and NTSB, were used inconsistently in the field. Common definitions combined with discipline in reporting would allow for improved discovery of patterns.
2. Minimum information requirements for each incident/accident. What might seem common or routine in one event could actually be a part of a larger and pattern. Currently a bare amount of information is reported for some events, even those incurring substantial damage. For example, the significance of MD-11 tailstrikes is outlined in this paper. Here is an example of an abbreviated FAA report on one such incident.

We checked Captains certificate and medical for currency and presence. (First Officer had left the airport prior to our checking certificate or medical...[person unnamed due to privacy concerns] called and asked one of the other inspectors what we were going to question the Captain about and said they had procedures and wanted to know if the company could get the information. We agreed. We checked the aircraft logbook entries, and left the area. This incident is closed (Federal Aviation Administration [FAA], 2006, Narrative)

Not only does this abbreviated report fail to provide specifics on the structural damage to the airplane, the minimum found in other reports and one of the reasons for post-incident inspections. It also fails to mention any possible causation for the incident.

3. Language requirements for major accident reports. In 2001 Eva (flight number not indicated) suffered a hard landing with structural damage in Taiwan. The 97 page ASC of Taiwan report is only available in Chinese limiting its usefulness internationally. Currently, the International Civil Aviation Organization (ICAO) has six official languages. One option would be to have major reports translated into all six languages. However, this would be costly as evidenced by ICAO's ongoing efforts to reduce translation costs for its own publications (ICAO, 2004a). Another possibility would be to provide English versions of major accident reports to the public. Of the 188 contracting ICAO nations, over 100 use English exclusively as their official language of correspondence (ICAO, 2004a). ICAO requires English be made available when a flight crew is unable to communicate in the language of a station on the ground. Officials at ICAO have written that these provisions "do not in any way limit the use of a national, regional, or local language but recognize the practical requirement for English to be available for the many pilots who do not speak the national language of a particular State" (ICAO, 2004b). It would be reasonable for ICAO to study the extension of this policy to major accident reports by making them available in English for investigators, regulators, and manufacturers who do not read the national language of a particular State.

## References

- Approach and Landing Joint Safety Analysis Team (1999). *Approach and landing: Results and analysis*. Washington, DC: Author.
- Berger, R.L., Hess, J.R., & Anderson, D.C. (1973). *Compatibility of maneuver load control and relaxed static stability applied to military aircraft*. Dayton, OH: Air Force Flight Dynamics Laboratory.
- Boeing (2009). *Statistical summary of commercial jet airplane accidents worldwide operations 1959-2008*. Seattle, WA: Author.
- Civil Aviation Department Hong Kong (CAD) (2004). *Report on the accident to Boeing MD11 B-150 at Hong Kong International Airport on 22 August 1999* (Aircraft Accident Report 1/2004). Hong Kong: Author.
- Federal Aviation Administration (FAA) Accident/Incident Data System [AIDS]. [http://www.asias.faa.gov/portal/page/portal/ASIAS\\_PAGES/ASIAS\\_HOME/DATAINFO](http://www.asias.faa.gov/portal/page/portal/ASIAS_PAGES/ASIAS_HOME/DATAINFO)
- Federal Aviation Administration (FAA) (2006). *ASIAS Brief Report [UPS MD-11 282UP 14 November 2006] (Report No. 20061114025179C)*. Washington, DC: Author.
- FedEx Narita Japan March 23 2009. Video posted to <http://www.youtube.com/watch?v=yWZyPp8SiPc>
- Field, G.C. (1987). MD-11 design: Evolution, not revolution. In *AIAA '87. AIAA/AHS/ASEE Aircraft Design, Systems and Operations Meeting*, September 14-16, 1987, St. Louis, MO.
- International Civil Aviation Organization (ICAO) (2004a) *Increasing ICAO efficiency* (Assembly – 35<sup>th</sup> Session – Working Paper A35-WP/208 EX/84 24/09/04). Montreal, Canada: Author.
- International Civil Aviation Organization (ICAO) (2004b). *Manual on the implementation of ICAO language proficiency requirements* (Doc 9835 AN/453). Montreal, Canada: Author.
- Irish Aviation Accident Investigation Unit (AAIU) (2001). *Report [World Airways MD-11 18 September 1999]* (AAIU Report No. 2001-006). Dublin, Ireland: Author.
- Khatwa, R., Helmreich, R. (1999). Flight Safety Foundation Approach-and-Landing Accident Reduction Task Force. Analysis of critical factors during approach and landing in accidents and normal flight: Data acquisition and analysis working group final report. *Society of Automotive Engineers (SAE) 1999 Transactions Journal of Aerospace*, 108, 1172-1276.
- National Transportation Safety Board (NTSB) (1993). *Inadvertent in-flight slat deployment China Eastern Airlines Flight 583 950 Nautical Miles South of Shemya, Alaska April 6, 1993* (Aircraft Accident Report NTSB/AAR-93/07). Washington, DC: Author.
- NTSB (1994). *In-flight turbulence encounter and loss of portions of the elevators China Airlines Flight 61-012 McDonnell Douglas MD-11p Taiwan registration B-150 about 20 miles east of Japan December 7, 1992* (Aircraft Accident Report NTSB/AAR-94/02). Washington, DC: Author.



- NTSB (2000a). *Crash During Landing Federal Express, Inc. McDonnell Douglas MD-11, N611FE, Newark International Airport, Newark, New Jersey, July 31, 1997* (Aircraft Accident Report NTSB/AAR-00/02). Washington, DC:Author.
- NTSB (2000b). [*Safety Recommendations A-00-92 through -103. Letter to Federal Aviation Administration (FAA), August 25, 2000.*] Washington, DC:Author.
- NTSB (2001a). *Factual Report [Eva Air BR316 MD-11 B-16101 November 20, 2001]*. Washington, DC: Author.
- NTSB (2001b). [*Safety recommendation history for A-00-100 FAA.*] Washington, DC: Author.
- NTSB (n.d.)Accident Database and Synopses. Retrieved from <http://www.ntsbgov/ntsb/query.asp>
- Taiwan Aviation Safety Council (ASC) (2001). [*Final Aviation Accident Report. Eva Air BR316 MD-11 B-16101 November 20, 2001*] (ASC-AAR-02-12-001). Taipei, Taiwan: Author.
- Transportation Safety Board of Canada (TSB) (2003). *In-Flight Fire Leading to Collision with Water Swissair Transport Limited McDonnell Douglas MD-11 HB-IWF Peggy's Cove, Nova Scotia 5 nm SW 2 September 1998* (Aviation Investigation Report number A98H0003). Quebec, Canada: Author.
- Wind shear linked to FedEx's Narita loss. (2009, March 30). *Aviation Week & Space Technology* 170(13).



## ***Developing Next Generation Research Competencies Through Collaborative Student Design and Advanced Manufacturing Projects***

Sergey Dubikovskiy, Timothy D. Ropp,

and

Thomas Lesczynski

*Aviation Technology Department  
Purdue University  
1401 Aviation Drive  
West Lafayette, IN, 47907-2015  
U.S.A.*

[sdubikov\[at\]purdue.edu](mailto:sdubikov[at]purdue.edu)

[tropp\[at\]purdue.edu](mailto:tropp[at]purdue.edu)

[tlesczyn\[at\]purdue.edu](mailto:tlesczyn[at]purdue.edu)

### ***Abstract***

*Students can remain in passive learning states even during participative, active learning projects and risk missing deeper learning opportunities. Faculty at Purdue University's Department of Aviation Technology, Aeronautical Engineering Technology (AET) program incorporated additional outcomes of forward thinking design and innovative research considerations into a traditional design-test-build project involving two senior level courses. Using a collaborative design approach, two AET classes were chartered to produce a component for actual use within aircraft maintenance laboratories. Additionally, students were challenged to focus on extended research goals and utility beyond just production of the component itself. As a result, student teams collaboratively produced a robust, real-use tool that is being used in other 'Hangar of the Future' research and development activities within the department. Some of those same students are now assisting in larger next generation research activities using this same component.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

## **Developing Next Generation Research Competencies Through Collaborative Student Design and Advanced Manufacturing Projects**

The Bureau of Labor Statistics (2008), US Department of Labor, predicts the anticipated rate of growth of employment of engineering technicians in the areas of aviation maintenance and aerospace engineering will be 10 percent between 2006 and 2016, compared to 7 percent national average for all occupations. Implications of this growth include a significant need to close an identified performance gap between aerospace engineering technicians and engineers. Increased complexity of technologies and reliance on an adaptive learning organization has forced employers to seek a workforce with skills and abilities beyond individual technical competency. This implies better-prepared graduates who have exposure to complex design projects and a blended experience of working with multidisciplinary teams. This is often accomplished at the university level using team design projects.

Without an innovative vision or use case that looks beyond just a laboratory project's desired design product, students easily fall back into passive learning roles even when participating in hands-on projects. Students tend to take on design-test-build projects without much thought beyond meeting initial design sketches or preset instructions. Aside from less than optimal project results, of greater concern is the missed opportunity to practice critical skills like team communication, planning, and deep problem solving which are among the rigid requirements of industry (Samuel, 2005). The result is often missed opportunity to learn how thoughtful design of components for innovative use or "an eye for the future" can be leveraged to help create a learning culture. This ability can significantly impact an organization's future success and viability.

A more thorough understanding of engineering technology disciplines along with the ability to envision forward looking design parameters and component capabilities through use of assertive communication and collaborative team work are core skills that technology based industries like aviation now demand (Shull, 2005; Bouckley, 2006; Ropp & Stanley, 2006). In the fall semester of 2008, faculty from two Aviation Engineering Technology courses that closely resemble real world industry environments experimented with an intra-course collaborative student design project, with the goal of challenging and stretching students' visionary and innovative research design skills.

### **Considerations for Course Selection**

Two senior level courses (AT 402-Aircraft Airworthiness Assurance and AT 408-Advanced Aircraft Manufacturing Processes) were used to successfully produce a usable aircraft equipment stand with the added design consideration of "forward fit" use in ongoing research and development activities for maintaining the next generation of modern aircraft. AT 402 is an intensive senior level capstone course designed to challenge students' ability to incorporate and practice leadership and management competencies with their technical degree skill sets in a realistic large aircraft maintenance environment. The laboratory portion of the course utilizes Purdue University's two large transport category aircraft, a Boeing 737 and Boeing 727, to simulate a large scale aircraft maintenance operation. Both aircraft have fully functional engines and systems. The aircraft are excellent platforms for practicing industry standard maintenance procedures, performing aircraft systems operations, and as research test beds for advanced aircraft maintenance systems development. Senior maintenance technology students function as operations

managers tasked with researching, planning, and implementing a large aircraft production maintenance operation, while managing junior classmen who function as working technical crews on the aircraft.

AT 408 students integrate baseline technical skills with larger problem solving skills and processes involved in design and manufacture of complex aircraft related parts and assemblies, including structural joint design, and aircraft components which play a critical role in advanced manufacturing and flight safety in industry (Vlasman, Dubikovsky, Schwartzkopf, & Vallade, 2008). The course is almost entirely problem-based, challenging students to perform research and to design products to specific requirements. Students must follow all stages of the design process, including project cost assessment, establishing timelines, and producing process sheet and work instructions. This structure incorporates recognized benefit of making students active participants in their own learning (Massa, 2008). The course simulates an independent business enterprise technique, where the students are "hired" to perform tasks starting with design of a product to manufacturing, to assembly of a final component (Dubikovsky, 2007). Pre-requisite courses already teach students critical logic behind procedures, essential for active learning (Shakirova, 2007), and through this preparation students are freed to take on creating and testing procedures themselves.

AT 402 and AT 408 were considered excellent candidates for pairing and integrating research and innovation concepts into a collaborative hands-on design project. Several of the students in AT 408 would be transitioning into AT 402 in succeeding semesters, subsequently using the very component they would be creating in both aircraft maintenance laboratory work and in research applications.

### Method

Students from AT 402 represented a production maintenance operation, delivering a requirements presentation report to the AT408 class who acted as a design and manufacturing engineering team as shown in Figure 1.

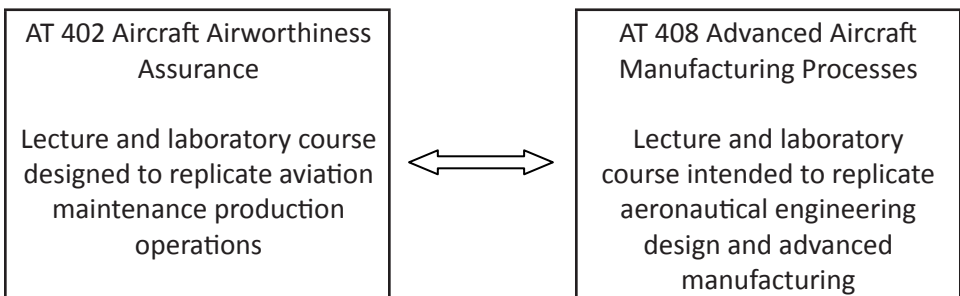


Figure 1. Course pairing for collaborative design project

Pairing AT 408 with AT 402 allowed realistic interactions similar to those found in the aviation and aerospace industries between maintenance and engineering design and support organizations. The student teams were tasked with developing a mobile aircraft parts/equipment storage cart. Capability requirements included storing components from large transport category aircraft along with future adaptability for the addition of computerized electronic parts tracking and data transmission components incorporated onto the cart structure.

A project kickoff meeting between the AT 402 and AT 408 student teams was conducted discussing details of design, manufacture, cost, and delivery estimates. Through a needs assessment exercise, students produced one of three required prototype design sketches for a proposed “Smart Cart” aircraft maintenance stand shown in Figure 2.

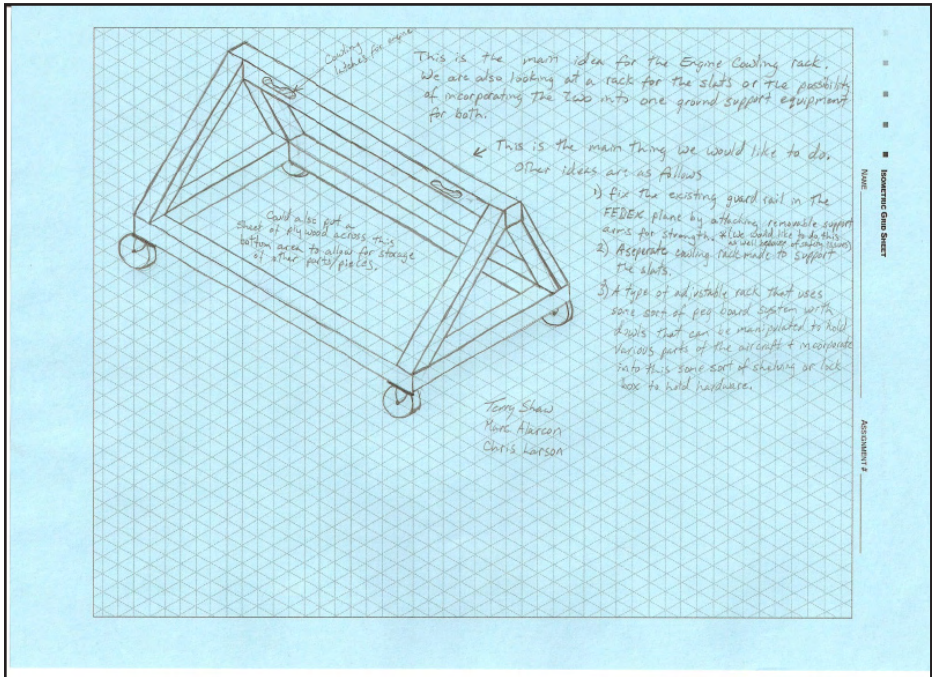


Figure 2. “Smart Cart” prototype design sketch

Subsequent follow-up design meetings between AT 402 and AT 408 teams were held throughout the design and advanced manufacturing process including a detailed design and use case risk assessment conducted by the student team.

### Producing the “Smart Cart”

Incorporated into the basic design features was a forward-looking goal of incorporating the final designed product into the Aviation Technology department’s larger *Hangar of the Future* research initiatives. As a research test bed component, the cart would be assimilated into a network of smart tools and systems development for streamlining maintenance and engineering work on modern aircraft. The project’s name was aptly titled “Smart Cart”.

Specifically, in addition the cart’s robust utility design requirements including indoor and outdoor all weather use and room for multiple large aircraft parts like engine cowlings, hydraulic pumps, and landing gear doors, faculty challenged the student teams to incorporate additional capabilities that included structural design consideration for accommodating Radio Frequency Identification (RFID) and computer equipment attach points. The resulting prototype cart in use with RFID test equipment attached is shown in Figure 3.





Figure 3. "Smart Cart" prototype aircraft maintenance stand

### **Transforming a routine design project into a research platform**

Using the smart cart as their platform, faculty and students in AT 402 began constructing research hypothesis and test case scenarios for applying RFID technology directly to the cart. The cart could then recognize when new aircraft parts were placed on it, identify the part, and transmit enhanced data to the human technician in a useful format, improving speed and accuracy of repair, tracking, and reinstallation.

Students designed initial experiments and test case scenarios around these capability targets, beginning with the more basic tasks of parts recognition. Equipment used included a passive RFID button, a bat, a low profile waterproof disk, and simple I.D. cards, which were programmed with specific aircraft components that were tagged for the research testing purposes.

A basic 125 kHz RFID reader was attached to the Smart Cart. The students learned to write and change messages and identification information to each tagged part. Early testing included testing the distance the reader could detect tags, amount of data each tag could hold, as well as how different aircraft parts physically held tags. Another difficulty encountered was a problem with RFID tag signals being limited or blocked through metal structures of aircraft parts. While the RFID antenna could read tags through thin paper material (such as pages of a book) and clothing, testing revealed tags had limited capability with metal objects such as an engine cowl or the smart cart the reader could detect and read the tags on top of the object, but not through the object.

Some significant discoveries included limitations of distance. The farthest distance reading capability of a part on the cart was 3 inches from the reader. This



was problematic in that it required a part to be placed close to the scanning RFID antenna, or, a stronger RFID system. Future testing is planned with a more complex multiple antenna array system and ultimately create a cart mounted RFID inventory system that could be moved into the marketplace and used by industry.

While the innovative use of RFID technology integration presented a long-term challenge to the student researchers, this multi-semester, multi-team student project was considered a success precisely because of those challenges. Students were forced to seek breakthrough answers or alternative solutions to issues such as RFID tag range limitations, impedance through aircraft metal structures, and transferring part data into an existing online laboratory database system.

Instead of passively stopping to wait for the instructor's direction, students had to address these additional issues, some of which they did not overcome due to time limitations of the semester. As a result, students were able to identify next step research and design targets to be passed on to future designers. Students began to quickly understand the complexity, time, and cost considerations related to design and implementation of automation and innovative design, two key components for maintaining modern aircraft.

### **Conclusion**

Problems students encountered in the combined design, production, research, and testing phases of this project served as valuable learning opportunities. Experimentation integrating technology directly onto their own designed product showed the nature and challenges of research, while forcing them to generate alternative testing strategies to overcome barriers in the technology as well as among team members. Additional benefits include a solid student designed and manufactured departmental equipment base for future use in advancing technical research in aerospace and aviation fields. Research and development partnerships between industry and universities achieve greater impact as students have the opportunity to acquire skills and experience by solving real world challenges.

## References

- Bouckley, S. (2006). Filling workforce skill gaps. *Manufacturing Engineering*, 137(2), 18-19. Bureau of Labor Statistics, U.S. Department of Labor (2008). *Occupational outlook handbook, 2008-09 edition, Engineering Technicians*, Retrieved March 23, 2009 from <http://www.bls.gov/oco/ocos112.htm>, accessed on March 23, 2009.
- Dubikovsky, S. (2007). AT408, Inc .trainee (student) orientation manual. *AT408 Course syllabus*. Purdue University, West Lafayette, IN.
- Massa, N. M. (Winter, 2008). Problem-Based Learning (PBL): A real-world antidote to the standards and testing regime. *New England Journal of Higher Education*, 22 (4), 19-20.
- Ropp, T.D. & Stanley, D. (2006). Developing learning outcomes to fit industry metrics. *Aviation Technician Education Council Journal (ATEC)*, 27(2), 37-39.
- Samuel, L. (2005). Communication skills are keu (i.e. key) to career success. *Water, Environment & Technology*, 17(2), 72.
- Shakirova, D. M. (September 2007). Technology for the shaping of college students' and upper-grade students' critical thinking, *Russian Education and Society*, 9 (9), 42–52.
- Shull, P. J. (2005). Collaborative learning and peer assessment to enhance student performance. *Journal of Engineering Technology*, 22(1), 10-15.
- Vlasman B., Dubikovsky S., Schwartzkopf J. & Vallade D. (2008). Proceedings from 2008 IAJC-IJME International Conference on Engineering and Technology. Learning aerospace design and manufacturing processes to demonstrate ABET program objectives. Nashville, Tennessee – November 17-19, 2008



## ***The Efficacy of Flight Attendant/Pilot Communication in a Post-9/11 Environment: Viewed from Both Sides of the Fortress Door***

Lori J. Brown and William G. Rantz

*Western Michigan University  
College of Aviation  
237 North Helmer Road  
Battle Creek, Michigan  
[Lori.Brown\[at\]wmich.edu](mailto:Lori.Brown[at]wmich.edu)*

### ***Abstract***

*The diminutive amount of research on flight attendant/pilot communication has shown gaps that impede effective communication and coordination between crewmembers. The process of information exchange between flight attendants and pilots has been further complicated due to physical and procedural changes since September 11, 2001. Thus, the effectiveness of typical intra-crew communication and coordination has areas that need improvements and in many cases are cited as the causation factor in a review of accidents and incidents. This study examines the effectiveness of current communication and coordination between flight attendants and pilots as well as crew resource management (CRM) programs. This study used a self-reporting survey to question 112 pilots and 230 flight attendants worldwide. Results indicate the dire need for improved flight attendant/pilot communication, coordination, and several recommendations for improvements are discussed.*

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

## **The Efficacy of Flight Attendant/Pilot Communication in a Post-9/11 Environment: Viewed from Both Sides of the Fortress Door**

Recent airline operational and institutional changes have been epic, particularly as a response to 9/11. According to the Honorable Jerry F. Costello testifying before the U.S. House of Representatives, Aviation Subcommittee, regarding the US Airways 1549 Congressional Hearing (2009a), "The current economy has the entire workforce being asked to do more with less, including work longer hours, this situation also highlights the association between training, workforce development, and aviation safety" (p. 20). Changes such as strengthening cockpit doors, specific procedures to enter and exit the flight deck, substantial pay cuts, increased duty days, and reduced rest periods have placed additional strain on the aviation industry. US Airways pilot Chesley Sullenberger told the House of Aviation subcommittee that his pay has been cut "40 percent in recent years and his pension has been terminated" (p. 26). Sullenberger further stated,

Americans have been experiencing huge economic difficulties in recent months, but airline employees have been experiencing those challenges and more for eight years. We have been hit by an economic tsunami, including; September 11, airline bankruptcies, fluctuating fuel prices, mergers, loss of pensions, and revolving-door management (p. 26).

These changes have evoked a significant human and social impact on flight attendants and pilots that have affected how crews work as a team onboard commercial aircraft. Traditionally crews are trained in Crew Resource Management (CRM) that is designed to enhance teamwork skills and communication. The Federal Aviation Administration (FAA) requires pilots and flight attendants to undergo CRM training with the goal of preventing accidents and learning to cope with stressful situations by improving communication and performance through enhanced coordination (FAA, 2004). Although air carriers are required to provide CRM training, joint flight attendant and pilot CRM training is not required by regulation. Joint CRM training is useful for gaining mutual understanding of the issues that affect different groups and useful in reconciling incompatible training practices (FAA, 2004).

Helmreich and Foushee (1993) identified communication as a critical link for coordination and teamwork between cockpit and cabin crew and essential to "prevent accidents through improved communication in air carrier operations, and keep safety at the highest possible level" (p. 14). Past research cites communication as the causation factor in investigated accidents and incidents (e.g., Cushing, 1995; Faith, 1996; FAA, 2004; Foushee, 1986; Foushee & Helmreich, 1988; Kanki & Palmer, 1993; Kirvonos, 2005; Wiener & Nagel, 1988). Chute and Weiner (1995) found that flight attendant/pilot communication is also impeded by physical and psychological barriers that have affected safety of the flight environment.

While studies have recommended communication training for pilots and flight attendants, they have also shown the need for joint CRM training (Baker & Frost, 1994; Butler, 1993; Chidester & Vaughn, 1994; Chute & Wiener, 1996; Edwards, 1992; FAA, 1988; Kirvonos 2007; Sexton & Helmreich, 2000; Young, 1994). In spite of this research, gaps still exist that impede communications between the two groups (Chute & Weiner, 1995; Brown & Niehaus, 2009) and indicate few airlines have acted upon the recommendations from such research.

Chute and Weiner completed a study in 1995 to investigate the status of crew interactions. They conducted a survey at the NASA Ames Research Center of “302 crew members: 177 pilots and 125 flight attendants at two U.S. airlines. The instrument utilized was a 30-item questionnaire composed of multiple choice, five point Likert-type scale responses, and open-ended questions” (p. 5). Chute and Weiner identified barriers that “isolate and alienate the cabin and flightdeck” (p. 16). Some of the factors identified by Chute and Weiner were historical, psychosocial, regulatory, and organizational. As a result of the Chute and Weiner study, several recommendations were suggested such as the reorganization of pilots and flight attendants under the same administrative structure or the creation of a liaison between the flight operations and in-flight departments. Chute and Weiner (1995) also stated, “If zero accidents are truly the goal of the aviation community, we must encourage professional, mature, unambiguous, and open communication between pilots and flight attendants. Anything less is a compromise with flight safety” (p. 17). The study completed by Chute and Weiner has made a significant scientific contribution to evaluate and understand intra-crew interactions; however, much has changed since 1995, particularly post- 9/11. Therefore, a current global study of intra-crew relations is needed.

Previous research (Helmreich & Foushee, 1993; Loftus, 1979; Wiener, Kanki, & Helmreich, 1993) addressed why communication breakdowns can occur between highly trained, technically skilled crew members; however, very little research addresses the level of communication between flight attendants and pilots. Human factors literature and research continues to focus on pilot-to-pilot and pilot-to-air traffic control (ATC) communication, team-oriented research, (Kanki, & Palmer, 1993; Morrow, Rodvold, & Lee, 1994; Orlady & Orlady, 1999; Salas, Bowers, & Edens, 2001; Salas, Stagl, & Burke, 2004) and air traffic controllers (Hartel, & Hartel, 1995; Smith-Jentsch, Baker, Salas, & Cannon-Bowers, 2001).

Hackman (1993), Johnston (1993), and Diehl (1991) note that team skills and the principles of CRM should be introduced early in training, refreshed, and evaluated. Likewise, Johnston (1993) stresses that pilots and flight attendants should be trained as a crew from the beginning in order to perform as a team in an airline setting. This was evident in the ditching followed by an evacuation of US Air Flight 1549, on January 15, 2009. The successful ditching of the A320 aircraft provides a poignant example that further delineates the importance that effective communication and crew coordination play in aircraft safety outcomes. “The successful landing of US Air Flight 1549 on the Hudson River without the loss of life was nothing short of a miracle and the performance of the flight and cabin crew was exemplary” (Brown & Niehaus, 2009, par. 1), which can be attributed to training and the high level of experience of the crew. Captain Sullenberger repeatedly pointed out in his testimony to the Aviation Subcommittee (2009a), “the positive outcome of Flight 1549 was a team effort” (p. 21). Clearly, training played a central role and proves the importance of effective pilot and crew training, coordination, and communication. Candace Kolander, Coordinator, Air Safety, Health, and Security Association of Flight Attendants-CWA, AFL-CIO stated, “When things start to fail in the cabin, we are left to rely solely on our training” in her testimony to the Aviation Subcommittee (2009b, p. 58).

Evidence of poor intra-crew communications has been cited as a causation factor in several fatal accidents such as the 1989 Air Ontario crash in Dryden, Canada that killed 24 people (Moshansky, 1992) and the Kegworth Air disaster

that occurred on the January 8, 1989, when a British Midland, Boeing 737 crashed in Leicestershire England (Air Accidents Investigation Branch, 1990). In addition, the failed attempt to bomb Northwest Airlines Flight 253 on December 25, 2009, revealed deficiencies, which deserve serious consideration. According to the FAA Administrator Randolph Babbitt, the pilots of Northwest Airlines Flight 253 were not immediately alerted that a passenger had tried to ignite a bomb on the flight from Amsterdam to Detroit. The pilots said they had a problem when the flight landed in Detroit. There was a communication gap between the cabin and the flight deck crew, which left the pilots unaware that there had been an alleged bombing attempt onboard. Administrator Babbitt told a House subcommittee, on February 4, 2010, that the flight deck crew reported they had someone who had attempted to set firecrackers off, so it didn't elevate to anyone — whether it was the cockpit or air traffic control — to anything of great seriousness at that point. Thinking the alleged bombing was no more than a prank, air traffic controllers took no extreme action like routing the aircraft to a remote location of the airport, nor did the cockpit get very excited about it, Babbitt said it was not until the aircraft was on the ground did the cockpit crew and airport personnel become aware of the alleged extent of the bombing attempt. These minutes and seconds could potentially be used to communicate with ground operations, warn other aircraft in the air, or land the aircraft, (U.S. House of Transportation and Infrastructure Committee, 2010).

Research conducted in the area of flight attendant/pilot communication remains limited, even though the need is clearly apparent (Clark, 2007; Edwards 1992; FAA 1988; Kirvonos, 2005). The “Pilot/Fight Attendant Communication and Joint Training” (Brown & Niehaus, 2009) study sought to identify gaps that impede effective communication in a post-9/11 environment from a global prospective. Some of the same issues (pre-9/11) were identified in the previously mentioned study completed by Chute and Weiner (1995). The Chute & Weiner study (p. 13) asked the crewmembers what they thought could be done to improve cockpit/cabin communication. The primary request that the cabin crew made was for mandatory briefings and introductions. Secondly, they asked for more respect and understanding of each other's duties, responsibilities, and workloads. They also requested joint training to teach both pilots and flight attendants communication skills. Data from the “Pilot/Fight Attendant Communication and Joint Training Survey” (Brown & Niehaus 2009) shows that many of these same issues exist globally today.

## **Method**

The “Pilot/Fight Attendant Communication and Joint Training Survey” (Brown & Niehaus 2009), funded by Western Michigan University, Faculty Research, and Creative Activities Award, was developed to investigate the status of recent crew interactions. This study evaluated a) flight attendant/pilot relations, b) the effects of lack of joint CRM training exercises, c) flight attendant reluctance to contact the flight deck, d) the impact of the mandated cockpit door strengthening requirements, and e) if traditional CRM programs adequately address communication issues between the pilots and flight attendants.

## **Survey Instrument**

The instrument was a 25-item questionnaire composed of multiple choice 5-point Likert-type scale responses and open-ended questions.



Current professional pilots and flight attendants were consulted as focus group members to aid in the development and evaluation of the survey instrument. To facilitate this study, a literature review was conducted to determine the most common perceived gaps in flight attendant/pilot communications and CRM training events. The questions were determined regarding the experiences in communicating on board an aircraft in normal and abnormal situations, as well as areas pilots and flight attendants feel are lacking in training for these situations. Additionally, questions were tailored for creating ways to produce more secure and cohesive aircraft environment. A coded matrix was used to identify the common themes of the interviews. Data collected during the study were used to develop a valid and reliable survey instrument. Results garnered from the focus groups allowed refinement to the instrument. The refined survey was distributed to a second focus group for evaluation.

## Validation

Once the initial survey was developed from the focus group forums, a face validity methodology was used. Utilizing professional pilot and flight attendants in a focus group forum ensured the survey questions were not confusing, misleading, or unintentionally measuring a different objective. The focus group data and question information was then used to modify all items mentioned.

After the initial focus group sessions were completed, the survey instrument was placed on [PulsewareSurvey.com.au](http://PulsewareSurvey.com.au) where a focus group consisting of 36 professional pilots and flight attendants were asked to read it and provide feedback on a user-friendly interface and on the instrument validity. Once common themes were identified, a ranking system was used to identify which themes are most frequently discussed and to adjust the survey instrument accordingly to ensure a valid, reliable, user-friendly instrument. Descriptive statistics were used to analyze the data collected. A comparison of the data gathered from the focus group forum and the data collected online via Pulseware were completed to identify the common themes of the data. The final survey instrument was then distributed to participants worldwide. The necessary subjects included industry flight attendants and pilots currently employed with major or regional airlines. A total of 342 crew members consisting of 112 pilots and 230 flight attendants participated in the survey. Out of 342 total survey respondents, 322 were labeled as "complete" within the system. Out of those 322 survey respondents, 291 successfully responded to the majority of the survey questions. The 291 respondents included in the survey represent 29 countries throughout the world. The countries of origin for the respondents airlines include; Australia, Austria, Belgium, Brussels, Canada, China, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, Mauritius, Mexico, New Zealand, Poland, Portugal, Romania, Slovakia, Spain, Switzerland, The Netherlands, Turkey, United States, United Arab Emirates, United Kingdom, and Venezuela.

The survey was translated to Spanish and Chinese to encourage global participation. No names or identification numbers were collected to ensure subject anonymity and confidentiality. The participants were asked if they were employed as a pilot or flight attendant and how many years they have served in their respective positions. The mean number of years as a pilot was 13 ( $SD = 10$  years) with a range of 38 years. The mean number of years as a flight attendant was 16 ( $SD = 10$  years) with a range of 38 years. No other personal data was collected.

The electronic survey was distributed via union and non-union U.S. and international airline websites, via links to the online survey, as well as links sent via email directly to individual participants. Unions and airlines were sent an informational letter via either email or U.S. Postal Service to ask for their consent in allowing the use of their websites and members. Individuals were informed via email, including a link to the survey and an explanation of the project including its importance in the industry.

Participants were asked to read and understand the contents in the consent form and the survey ground rules to prove that they understand the terms of the study and agree with the rules of participation. This form states that their identity will be anonymous. The form also informs participants of their right to decline to answer any question or to remove themselves from the study at any point without penalty or negative consequences. No research related questions would be conducted before this form is read and the "I agree" button is clicked at the bottom of the form (this simulates an electronic signature of agreement).

The <http://www.pulseware.com.au/> survey instrument developer was used to distribute the survey information to the union websites, valid internet host sites, or via email. This program collected and analyzed the data. The researchers also enlisted the help of the company Aptima, Inc., (Human Centered Engineering) to help in the analysis of the data.

## **Data Analysis**

No names of individuals, airlines, unions, or identification numbers were collected to ensure subject anonymity and confidentiality. The data was collapsed across all the airlines. Analysis of the open-ended comments required reading each response in order to establish response categories. Categories were taken from the actual responses with as little inference as possible in order to preserve the integrity of the data. There may be some overlap in categories. However, since open-ending replies do not always fall within clear categories, some judgment of the data was needed in the analysis. The categorization was iterative until the raters agreed.

## **Limitations**

There appears to be no published research of this kind in a post-9/11 environment. Therefore specific factors have not been identified through literature review as key variables involving flight attendant/pilot communications. Additional research questions that arose during this study went beyond the scope of this project and may be more appropriate for a future research study with smaller sample sizes.

## **Results**

Question 1. How would you rate your current airline's level of effective communication between pilots and flight attendants?

Flight attendant and pilots rated their airline's level of communication effectiveness. Out of 291 responses, three percent (3%) of the sample indicated poor communication with no pre-departure briefing and little or no communications in-flight;

although 3% is a relatively low percentage, it is important to note that all of these responses came from U.S. Airlines.

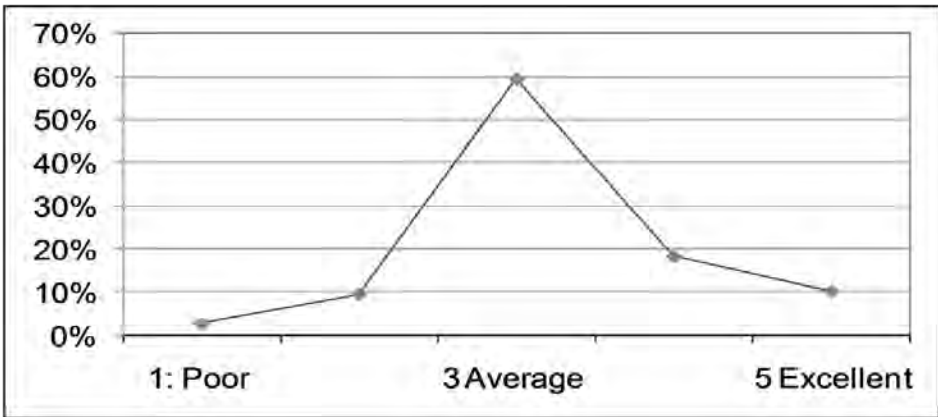


Figure 1. Flight Attendant/Pilot Communication Effectiveness

Ten percent (10%) indicated the typical communication between pilots and flight attendants at their airline consisted of a pre-departure briefing only; again, the data showed that the majority (all, but two responses) of the 10% were from U.S. Airlines.

Fifty Nine percent (59%) indicated average communication in-flight with a pre-departure briefing; (*Including, respondents from the US, Poland, Turkey, Canada, Switzerland, Mexico, Australia, Venezuela, New Zealand, Japan, UK, Hong Kong, Mauritius, Mexico, Greece, Portugal, Spain, Paraguay, Argentina, and Romania*).

Eighteen percent (18%) indicated above-average communications with both pre-departure and post-flight briefings. (*Belgium, France, Brussels, Slovakia, Romania, Australia, Poland, Germany Greece, Portugal, Switzerland, Spain, Finland, Netherlands, Turkey, U.S., and Canada*).

Ten percent (10%) indicated *excellent communications in-flight with both pre departure and post-flight briefings and excellent communications in-flight*. (*Belgium, France, Australia, Germany, Greece, Portugal, Switzerland, Finland, Netherlands, Canada, and the UK*).

*\*Note: 98% of the respondents from France indicated that their level of communication was above-average to excellent (4-5, on the likert scale). 90% of the respondents from France indicated that their communication was at level five on the Likert scale. Out of the sampling, 100% of the pilot and flight attendants from Belgium indicated that their airlines level of effective communication was above-average -to excellent.*

Question 2. How does a thorough pre-flight briefing affect communications in-flight?

Out of 291 responses, six percent (6%) of the sample indicated that a thorough pre-flight briefing might not affect communications in-flight (1-2 on Likert scale). Sixteen percent (16%) indicated the pre-flight briefing might affect communication between pilots and flight attendants in-flight (3 on Likert scale). Seventy Eight percent (78%) indicated that a thorough pre-flight briefing would affect communications in-flight (4-5 on Likert scale).

Comments: I like it when the pilots...

- *Come to introduce themselves to the entire crew (pre-departure briefing is only for FA1). Invite us to jumpseat for takeoff/landing. Make us feel that concerns are taken seriously.*
- *Begin the flight with open and respectful dialogue. I feel valued and respected as part of a team when my input is sought and we are treated as equals rather than subordinates. I feel more like a team when we are treated with respect for the safety aspect of our job.*

Conduct the required pre-departure briefings with all Flight Attendants, not just the number one or purser. The briefings that include weather/turbulence reports, time of flight, review of security when cockpit door is opened, etc.

Question 3. Which areas have created barriers that influence effective communication between pilots and flight attendants? (Select all that may apply)

- Time and operational constraints (n= 170)
- Fortress door (n=150)
- Lack of scenario based CRM training (n=92)
- Crew scheduling (n=88)
- Job understanding (n=76)
- Organizational structure (n=75)
- Procedures (n=68)
- Gender or *assumed* sexual orientation (n=48)
- Misunderstanding of the sterile cockpit rule (n=32)
- Aircraft interphone (n=29)

Comments:

- *I do not like it when pilots ignore me (I am a male FA) or do not provide me with the courtesy of an introduction or briefing. Sexual orientation can affect communication, such as gay male flight attendant with straight male pilots. Pilot's may keep contact with a flight attendant to a minimum due to suggested "morality issues."*
- *Any kind of expression of racism based on gender, color, & sexual orientation is a major obstacle for mutual trust on a professional level. Discomfort with different sexual orientation causes communication barriers.*

Question 4. What is the most visible organizational obstacle that affects flight attendant/pilot communications?

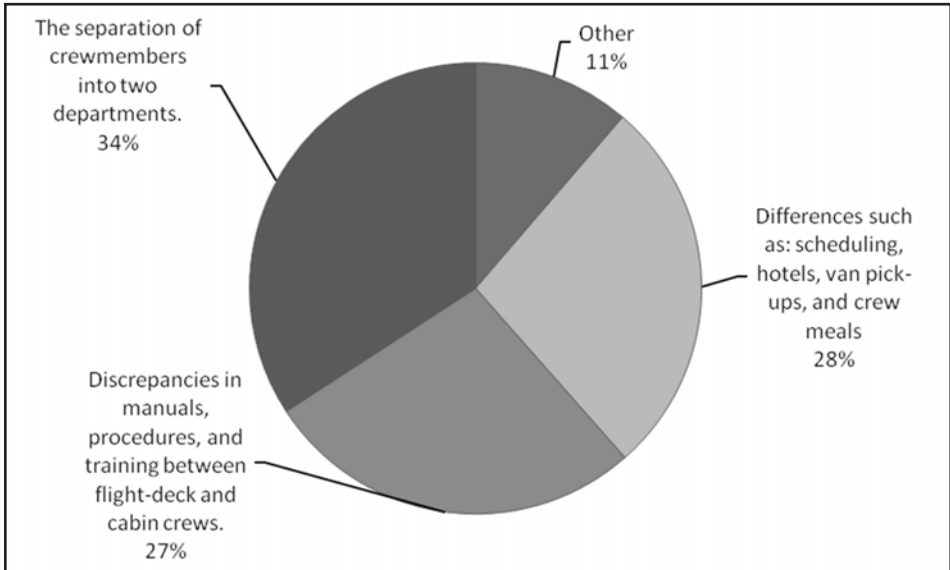


Figure 2. Organizational Obstacles

Comments:

- *FA's feel betrayed by pilots because they got raises and FA's did not. Because of this, FA's do not want to interact more than necessary with pilots.*
- *Issues with management have created distrust between the two groups.*
- *I believe our 4-year ongoing contract negotiations have caused some of our crews to lose interest in doing their job well. This has included not communicating with flight attendants simply for lack of motivation.*
- *Poor moral, rapid turnover, and level of professionalism. In part due to employees leaving as a result of poor pay and working conditions.*

Question 5. Do you notice any positive work-related differences when you are paired with the same flightdeck (or cabin) crew for several legs, as opposed to one or two legs of a trip?

Out of the 265 participants, 95% noticed positive work related differences when paired with the same crew for more than 1-2 legs.

Comments:

- *Pilot and Flight Attendants often switch crews after one or two legs. This can cause misunderstanding and poor communication. Multiple crew changes during quick turns, prevent introductions and briefings since boarding is already under way.*

- *I fear that the post- 9/11 security measures & economic constraints have set CRM back twenty years. I would have to say that 8 out of 10 flights are done where the pilots do their thing and we do ours in the cabin. A briefing is always done with the Purser (per FARs) but anything above that is rare. It is not uncommon, when working in the back, to have never met our flight deck crew when they fly just one segment that is dangerous.*

Question 6. Have you as a Flight Attendant been hesitant to report a problem to a pilot due to the sterile cockpit rule, fear of being reprimanded, or lack of understanding about a problem or system?

Out of the 224 flight attendants sampled, 55% reported that they have been hesitant to report a problem and 16% indicated that they have had a situation where they did not report a problem and did not inform the flight deck because they thought they already knew. Out of the 51 pilots sampled, 41% indicated that they have had a situation where a flight attendant reported a problem and did not inform the flight deck because they thought they already knew. Additionally, 57% of the pilots indicated they have noticed that flight attendants may be hesitant to report a problem due to misunderstanding of the sterile cockpit rule or other reason.

Comment:

- *Crew members should not be afraid to speak up if they have concerns about a passenger or situation regardless of the outcome. Some are afraid due to time constraints and retribution from the airline in case of a mistake. Better training in this area will help to prevent mistakes and give assurance to individuals that fear retribution for making a poor call concerning a given situation.*

Question 7. Do you feel that allowing flight attendants to ride on the jump seat would improve their understanding of pilots' workload and improve CRM?

Out of the 224 flight attendants sampled, 68% indicated that allowing flight attendants to jumpseat would be very helpful, to improve their understanding of the pilots' workload.

This is what one pilot had to say in the comments:

- *The aforementioned familiarization flight/access to jumpseats would be a huge improvement in allowing the F.A.'s to understand more of what goes on in their airplane on a daily basis. I cannot understand why F.A. are considered safe in-flight but never allowed to view an entire flight on the jumpseat. Understandably pilot distraction is a concern. In addition, pilots being more exposed to F.A. duties, responsibilities, and emergency/medical situations would also be a great improvement. Scenario based training, perhaps in an A/C at the hangar, with the cockpit door open would allow crew members to visualize the situations and stresses the other experiences during an emergency/medical and builds better understanding/concern than I see in many of the pilots/F.A.'s I work with today.*

Question 8. Please select the areas of training that you feel are lacking for pilots and flight attendants (Select all that apply).

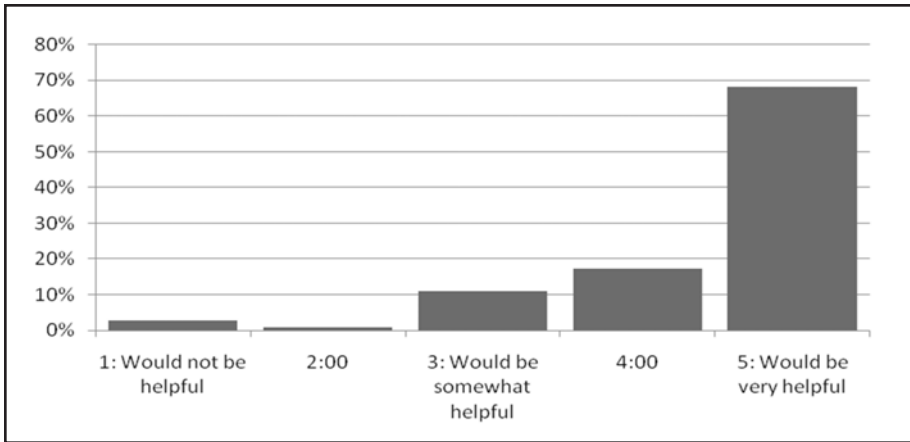


Figure 3. Jump Seating

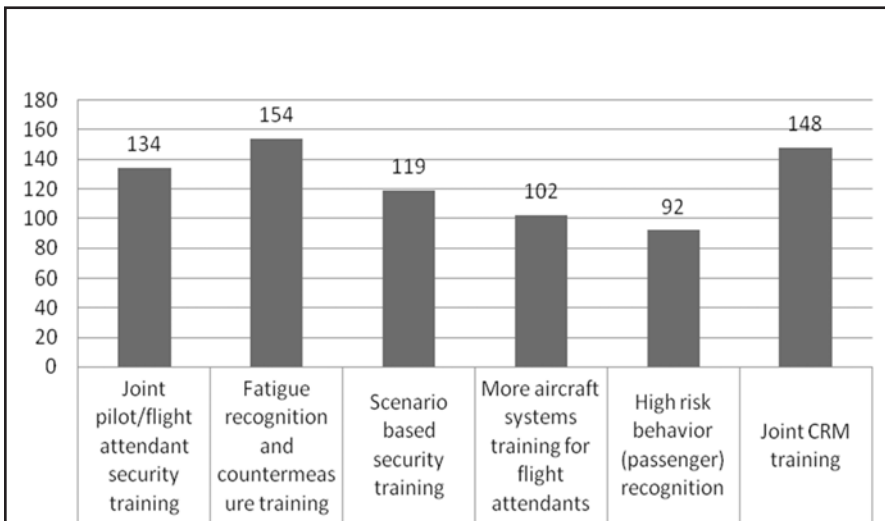


Figure 4. Areas of Training that may be Inadequate

In addition to the perceived inadequacy of joint CRM training, 73% of the pilots and flight attendants indicated that fatigue awareness training is lacking. Fatigue has been cited in recent accidents, incidents, and runway incursions.

This is what one pilot had to say in the comments:

- *I recommend that the airlines spend more time providing realistic and relevant scenario-based joint training for pilots and flight attendants that would stress and develop crew resource management skills. If there must be a silo, then the pilots and flight attendants of a crew must be in that it silo.*

Out of 228 respondents, 51% of the pilots and flight attendants indicated that their airline does provide joint communication training for pilots and flight attendants and 48% indicated that scenarios would be a very valuable addition to their



current training. Only 12% of respondents thought adding scenarios would not be valuable.

Question 9. Do you think it would be beneficial for the flight attendants and Federal Air Marshals (FAMS) to do any joint training together?

When asked, 79% of the survey respondents reported that they think it would be beneficial for flight attendants and FAMS to training together.

Comments:

- *At my airline of employment, during recurrent training one or two FAMS would come in to speak to the classes. They would discuss their role on-board the aircraft and what would be expected of the pilot/flight attendants if they were called to action. This discussion was beneficial but watching the FAMS perform actual drills would further pilot/flight attendant understanding of all roles involved.*
- *FAMS on board the aircraft are a crew resource. As with any resource, in order to use it effectively, one must know how to use it and that knowledge is best obtained through joint training. As a pilot, I have gone through some joint training with the FAMS and, as a result, I feel I have a better understanding of their use and function thereby.*
- *I believe that if a FAM and an FA were really involved in a security emergency, neither really has any idea what the other is going to do. If they are going to be expected to perform jointly in any fashion, then each should be trained what to expect and how to help each other perform their duties.*
- *This training is done in my airline and shows positive effects in cooperation.*

Question 10. Do you feel that a discreet wireless communication device would enhance safety/security onboard the aircraft?

We asked 271 pilots and flight attendants if they felt a discreet wireless communication device would enhance security or safety. In the survey, 13% indicated a discreet wireless communication device would not enhance safety, 18% indicate a slight effect, 33% indicated a device may somewhat enhance safety, and 21% indicated the device would greatly enhance safety.

Comments:

- *We need to stress better communications; the wireless panic button might be good if adequately protected from accidental activation. There has been talk of cell phone availability in-flight, which, in some form, could allow discreet communications/panic button functions inexpensively. We experienced a fire/emergency landing just last week and have been very concerned that my communications during the incident were inadequate. We were unable to post-brief as the FAA met us at the gate and whisked our FA's off for their purposes. I am sure I will get the opportunity to talk to that crew again, but wish it had been altogether as a crew and more timely.*

Question 11. Would you say that the inter-phone system provides “discreet communications” between the cabin and flightdeck?

The total number of responses were 222 of which 58% did not find the inter-phone system to be discreet, and 72% said they would be willing to wear the device to achieve wireless communication in-flight.

Comments:

- *I do feel the inter-phone is very obvious, but think wearing a blue tooth would make the F/A a target for an attack.*
- *Bluetooth, how would it be more discreet when it's visible, and so is someone speaking into it? Within galley areas or jumpseat areas a wireless system might make more sense. More defensive and/or profile training might be helpful.*
- *I think that wireless headsets would increase communication and save time during an incident or emergency.*

Question 12. Have you had an experience where an unfamiliar accent made communications difficult?

According to the study, unfamiliar accents also create barriers in communication, with 28% reported accents between flight attendants and pilots, 4% reported accents between pilots, 18% reported accents between pilots and air traffic control, 7% reported accents between flight attendants, 14% reported accents between “other” and 29% reported no accent challenges.

Comments:

- *The lead Flight Attendant spoke with a heavy accent. Not only were his announcements unintelligible to passengers, I would not have known if he were giving any safety or emergency-related messages.*
- *During an emergency on board the aircraft and I was giving CPR to a passenger. I told the other flight attendant to contact the cockpit and give them the information and we must make an emergency landing. She was Asian and her accent was not able to be understood by the cockpit, it increased in difficulty as she became more nervous. The cockpit needed to talk to me in order to be clear of the emergency, but I was busy saving a life.*
- *As a deadheading crewmember on a “domestic” flight, the “lead” flight attendant had a very thick accent, and obviously, English was not her first language. During the orientation airborne welcome, she read the script directly from the announcement handbook without editing, (e.g., we will be serving breakfast, snack, lunch, dinner), all on a one hour flight. Passengers found this humorous, if I had been a working crewmember, I would wonder if her communication with the cockpit would be affected by lack of understanding of implied meaning, shared cultural mores, and nuances of speech.*

- *As a flight attendant, I have been in the flight deck when the controller was next to impossible to understand, and ground agents who were extremely difficult to understand when doing announcements on the aircraft, or in the gate area. Many FA's from China speak English only understood in a word here and there, and this has been alarming to passengers, as well as in emergency situations, such as asking for a de-fib, and I got de-caf (sic) coffee.*

## Discussion

This study examined the effectiveness of current flight attendant/pilot communication following the recommendations from the Chute and Wiener study (1995). This study focused on three primary areas: a) to what extent do traditional CRM programs address communication issues between pilots and flight attendants, b) do these CRM programs need updating to reflect the post-9/11 environment, and c) to what extent has the post-9/11 environment affected communication and coordination between flight attendants and pilots.

In this study respondents were asked to select the areas that have created barriers that affect communication between pilots and flight attendants. The fortress door along with time and operational constraints were the factors that had the highest weighted values. This suggests a post-9/11 change regarding equipment, procedures, and operational standards.

This study has identified barriers, which may leave flight attendants and pilots feeling isolated, and may impede effective communication. Although the percentages may seem relatively low, data shows that sterile cockpit is still an issue, similar to the findings in the Chute and Weiner (1995) research, and seen as a factor in previous incidents (NTSB, 1994). This gap in effective communication can lead to loss of life as in the Dryden (Moshansky, 1992) and Kegworth (Air Accidents Investigation Branch, 1990) accidents, or create a communication gap that could impede the information transfer to the pilots, as seen in the Northwest Airlines flight 253 (U. S. House Representatives, Transportation and Infrastructure Committee, 2010) attempted bombing.

Sterile cockpit issues can be identified and addressed during initial, recurrent training, and further clarified by the pre-departure briefing. This is particularly important with the less experienced workforce, which is replacing the highly experienced crews, like the US Airways flight 1549 crew. We know sterile cockpit misinterpretations have contributed to fatal accidents, and have been reported as a gap in communication in the past and seems to be a recurring issue. The flight attendants occasional reluctance to contact the flight deck with safety related information due to misunderstanding of the sterile cockpit rule, a lack of technical understanding, or fear of being reprimanded creates a gap in communication. This barrier can be partially eliminated by briefing and training, which stresses the need to keep the flightdeck informed of all safety related items.

On occasion, a flight attendant may report a situation to the flightdeck during sterile cockpit, only to discover they were wrong, such as misinterpreting APU torching. This may cause the flight attendant to feel hesitant the next time they are unclear about a situation for fear of being wrong or, in some cases, reprimanded. Although additional systems training may help, it is crucial that the flight crew set a relaxed supportive tone at the beginning of flight, encouraging teamwork and

open communication, where all are encouraged to participate. If a “team oriented” tone is not set during the briefing at the beginning of the flight, this may add to the hesitancy to contact the flight deck with a perceived safety issue, particularly, during sterile cockpit.

The survey reveals that in many cases no briefing exists, or the purser is the only crew member briefed. It is understandable given the current economic constraints and constant change of crews that there is little or no time to brief. Therefore, introductions are very important to enable the “team formation” stage (Tuckman & Jensen, 1977), which may occur during the Captain’s briefing or after check in, to set the tone and establish the crew climate. Captains need to create an environment of trust and support where a Flight Attendant is confident to pass on concerns of safety related information to the flight deck without fear of intimidation or ridicule. The relaxed supportive crew climate and tone for the flight can be accomplished during the briefing. Pilots should keep the flight attendants informed and “in the loop” throughout the flight as workload and time permits to foster mutual respect. An effective briefing is a valuable tool for the cabin crew and flight crew to help manage errors and improve outcomes. The briefing should provide a clear picture of the flight ahead, and build a common understanding and expectation among the crew. Training can be utilized to reiterate the importance of a thorough brief, and aid in the development of effective briefings. Joint training would allow the opportunity for discussion and feedback.

Chute and Wiener (1995, 1996) also found that flight attendants were almost universal in their complaint that pilots, especially captains, failed to introduce themselves. Pilots also expressed a desire that flight attendants take the initiative to introduce themselves. This investigation reveals that even when a briefing exists it is often only given to the lead flight attendant and is not always thorough, as perceived by the flight attendants. In addition to all of the pertinent information for the flight and introductions, flight attendants want to feel valued.

Communication can also be affected by the Captain’s leadership style; conversely, authority gradients can be reduced through making the briefing a necessity. Furthermore, a debriefing fosters team building and addresses any issues that may have come up during the flight. A National Transportation Safety Board (1994) study has implicated crew familiarity (actually lack of familiarity) as a factor involved in accidents.

One solution may be to collaborate between flight operations, training, management, and scheduling departments to identify and plan time for crew briefings, thereby ensuring this critical element is not omitted. Furthermore, teaching pilots how to give an effective briefing may seem obvious; however, they may only receive this in their initial or upgrade training.

The survey results also indicated that gender has the potential to impede or influence flight attendant/pilot communication. This is not limited to male pilot/male flight attendant, as research needs to consider female pilot/male flight attendant, and female pilot/female flight attendant communications and authority gradients. Wilson (2005) speaks to this in her research of attitudes toward female pilots. She states, “Because perceptions on gender differences have a pervasive and powerful effect on behavior, it is important to manage gender diversity” (p. ii). She maintains that these perceptions have important implications on crew effectiveness.

This validates the need to consider multi-gender flight crews in future research agendas. Research completed by Tannen (1990, 1995, 1996), also notes that gender is an area that requires more in-depth future research, along with the influence of flight attendant and pilot sexual orientation on crew communication.

To address some of the communication issues resulting from a post-9/11 environment – wireless communication devices and discreet “Panic Buttons” need to be closely evaluated. These technologies may provide seamless communications throughout all phases of flight, when the flight attendants may not be near or unable to get to an inter-phone. This would also allow a flight attendant or Federal Air Marshal a means to notify the flight deck (discreetly) at the first sign of a security breach. Providing flight attendants with adequate security training and tools to enable communication if the inter-phone is disabled would send a positive message enabling them to feel less isolated, and feel more valued by the airline, thereby, increasing crew performance and security. A wireless communication device may have prevented the communication gap between the pilots and flight attendants on the Northwest Airlines Flight 253, December 25, 2009. Clearly, these devices cannot prevent such attacks, however, they may increase effective communication and add another layer of security onboard the aircraft. The flight attendants are the eyes and ears for the pilots in the cabin and have the ability to provide the pilots with critical time sensitive information. These valuable minutes could be used to land the plane, divert, or communication with ground operations.

According to the survey results, these barriers may not be adequately addressed at most carriers. While it is doubtful that fortress doors or secondary barriers will be removed from U.S. Airlines, communication gaps surrounding these post-9/11 changes can be lessened with the application of these possible remedies:

- Joint scenario based training would allow flight attendants and pilots to train together to stress the need to keep the flight deck informed of all safety related items, reduce sterile cockpit communication gaps, and improve coordination.
- Videos and case studies such as the failed bombing of the Northwest Airlines Flight 253 on Christmas Day 2009, Kegworth and Dryden accidents would allow dynamic crew discussion. These accidents and incidents can be used as a stimulus for discussion of various aspects of the accident scenario, including corporate pressures, particularly those resulting from operating in a deregulated environment, crew communication, teamwork, and role definition (Hayward, 1994).
- Joint CRM (preferably scenario based) training and familiarization jump-seat rides for flight attendants may help crews to understand the information each requires to perform their job effectively; and to be aware of the possible consequences if that information is delayed, or not provided, or is incorrect.

Incorporating scenarios that practice how flight attendants give concise, accurate information to the pilots and allow pilots to practice how and what questions to ask to get the most accurate picture of what is occurring in the cabin can be most beneficial. Active listening is also a very important part of the communication process and should be included into this training. According to Kirvonos (2005),

this aspect of the training could cover the importance of listening, the cost of poor listening, the process of listening, barriers to effective listening, and steps to better listening. However, mutual trust and respect must be established first and barriers removed before effective information transfer can take place.

Reinforcing the eight NASA team skills would be very helpful. These leadership/followership skills should be introduced long before the pilot reaches a commercial airline setting. Team skills should be used to enhance current academic curricula at early stages. Flight programs and academic institutions focus on pilot training- often forgetting about the flight attendants. It is evident that we need to start from the beginning and foster pilots that are good leaders and know how to interact with the cabin crew. Helmreich, Butler, Taggart, & Wilhelm, (1994) developed the line/LOS version IV checklist (currently line/LOS checklist VI) to assess team skills for pilots. These same skills can be used to create behavioral markers to assess Joint CRM/LOFT training events with pilots and flight attendants. Kirvornos (2005) suggests the best methods for this type of training would include experiential exercises, practicing communication skills, role-playing, small group and team building exercises, discussion, case study scenarios, and self-assessment tools. Incorporating joint scenario based CRM and security training (considering possible Federal Air Marshal Coordination, when appropriate) would enable all crew members to practice emergency scenarios as a team.

Practice of high-stress situations that require clear explicit communication between all of the crewmembers, such as the loss of an engine on takeoff, ditching, or fire, would be highly beneficial. This requires actions by the pilots along with rapid information transfer between the flight crew and ATC and between the cockpit and cabin crew. These types of scenarios are often embedded into Line Orientated Flight Training (LOFT) scenarios, involving only pilots. LOFT scenarios can include flight attendants to make the LOFT session more realistic and help flight crews practice these skills. Highly practiced skills, such as effective communication, inevitably leads to better functioning crews in an emergency. This type of training can be accomplished in a Crew Orientated Flight Training (COFT) simulator session, which integrates flight attendants into the scenarios.

To decrease the extra time and cost of scenario-based training, the flight attendants and pilots could complete some of the redundant recurrent training items online or via a home study course, saving some of the training time for joint scenarios. One example is Australian Airlines who have developed and implemented "integrated crew training" (Based on the Australian Airlines expanded Model), with very positive results, in "Australian Airlines Pilot/Flight Attendant Integrated Crew Training" (Baker & Frost, 1994). They have reported experiencing tangible benefits in terms of improved understanding, communication, coordination, and cooperation among crews because of this endeavor. The reaction of crews has been reported as "overwhelmingly positive and supportive."

## **Conclusion**

This study has revealed that current CRM programs need updating to reflect our post-9/11 environment, (particularly in the United States). These updates are required to address the changes of crew interactions due to new procedures, fortress doors installed and kept closed, bolted and barricaded, pilots confined to the



flight deck, and passengers viewed with suspicion (Chute, 2002). The results point to operational, procedural, physical, and psychological issues that have developed as result of changes in the workplace after September 11, 2009.

Joint CRM can provide participants with new operationally relevant and an interactive training experience with a human factors orientation, which can not only meet regulatory requirements, but also provide a vehicle for the development of better teamwork, communication, and improved relationships amongst crewmembers. Joint CRM can also be used to enhance effective crew coordination, practice communication skills as part of a team, understand, and appreciate each other's role, both on a daily basis and in an emergency. The results of the study indicate that some airlines have already implemented this type of training, outside of the U.S.

Although these programs have been beneficial, it is evident that joint CRM, training has to go beyond putting both pilots and flight attendants in the same room and completing existing training programs. While the concept of joint CRM training is not new, the design of effective joint scenarios where crews are together for repetitive sessions, ideally in an aircraft environment (such as a simulator or cabin trainer), may be more effective than training in isolation. Scenario based training, with both flight attendants and pilots in an actual aircraft or cabin trainer with a flightdeck would allow crewmembers to experience emergency/medical situations to build better understanding.

It would also be beneficial for the flight attendants to know about basics in the flightdeck such as- how to adjust the seat, squawk an emergency code, and basic use of the radio (if a pilot is incapacitated). These simple introductions would allow the flight attendants and pilots to further mutual respect, job understanding, and team cohesion.

With pilots safely barricaded behind their reinforced cockpit doors, and with instructions in place to limit their exposure to threats, it is crucial that airlines consider adopting some of the remedies recommended in this study. Whether it is using all crew inclusive scenario-based training or reliable communication tools to allow the aircraft crew on both sides of the door to communicate with one another, crews must work harmoniously together.

Organizational obstacles were viewed as the number one area that affects flight attendant/pilot communication, so it is important to take a closer look at the individual obstacles. The most prevalent obstacle is the separation of crews in two departments, this was also an issue followed by recommendation from Chute and Weiner (1995) following their study. Fifteen years later, this is still a primary issue. "This can result in conflicting goals, inconsistent instructions, manuals, and lack of communication" states Chute and Weiner (1995, p. 8). They also have argued this issue in previous papers, (Chute and Wiener, 1995, 1996) stating "that such organizational segregation also de-emphasizes the safety function of the cabin crew. If safety is truly the primary responsibility of the cabin crew, then they should be in the same department as the pilots" (p. 8).

While joint scenario based training, development, and scheduling may not be realistic for all carriers in these lean economic times, improvement of sterile cockpit training, and improved crew briefings are both viable and crucial. Coordination



and synergy of the flight deck and cabin crew has never been more significant or more challenging than our current operating environment, making the crew briefing critical. Collaborating efforts between flight operations, training, management, and scheduling departments to identify and plan time for crew briefings can ensure this critical element is not omitted. Furthermore, teaching pilots how to give an effective briefing may seem obvious; however, they may only receive this in their initial or upgrade training.

It is clear operational conditions for commercial airlines have drastically changed since 9/11 and so must our response to training. While any addition to expenditure is difficult to justify in the economic climate, we encourage operators to consider these recommendations.

### **Acknowledgments**

This research was funded by Western Michigan University through a FRACAA grant award. The authors would like to thank the hundreds of aviation industry professionals from 29 countries, who contributed their time and experience to this study. Additional thanks to Dr. Maria Novis, (AEPA) for translating the survey to Spanish. Dr. Jiazhong Yang, (Civil Aviation Flight University, China) for translating the survey to Chinese, Dr. Amy Alexander (Aptima, Inc.) for focus group data analysis, John Niehaus and Beth Seiler for contributions to the survey design, and development. Additionally, we would like to thank AFA, EAAP, APA, IFALPA, ALPA, and AEPA members for their participation and assistance in the distribution of the survey.

## References

- Air Accidents Investigation Branch. (1990). *Report on the accident of Boeing 737-400 G-OBME near Kegworth, Leicestershire on 8 January 1989* (Report No. AAIB AAR 4/1990). Department of Transportation, United Kingdom. Retrieved from [http://www.aaib.gov.uk/cms\\_resources.cfm?file=/4-1990%20G-OBME.pdf](http://www.aaib.gov.uk/cms_resources.cfm?file=/4-1990%20G-OBME.pdf).
- Baker, R., & Frost, K. (1994). Australian airlines pilot/flight attendant integrated crew training. *The CRM Advocate*, 4(1).
- Brown, L., & Niehaus J. (2009). Pilot/Flight Attendant Communication to Improve Aviation Safety. *Proceedings of the Corporate Aviation Safety Symposium*. Orlando, Florida: Flight Safety Foundation.
- Butler, R.E. (1993). LOFT: Full-mission Stimulation as Crew Resource in Management Training. In E. Wiener, B. Kanki, & R. Helmreich (Eds.), *Cockpit Resource Management*, 231-259. San Diego, CA: Academic Press.
- Chidester, T. & Vaughn, L. (1994). Pilot flight attendant coordination. *The CRM Advocate*, 94(1), January-November.
- Chute, R.D. (2002) Cabin/Cockpit Communication: Post-9/11. *Proceedings of the Assault and Hijacking Workshop sponsored by the Human Factors Group of the Royal Aeronautical Society*, April 10, 2002 London Heathrow.
- Chute, R. D. & Wiener, E. L. (1995). Cockpit/cabin communication: I. A tale of two cultures, *International Journal of Aviation Psychology*, 5(3), 257-276.
- Chute, R. D., & Wiener, E. L. (1996). Cockpit/cabin communication: II. Shall we tell the pilots? *International Journal of Aviation Psychology*, 6(3), 211-231.
- Clark, B. (2007). *Plane talk: Flight attendant identity construction and inter-crew communication*. (Unpublished master's thesis). Retrieved from School of Modern Languages of Queen Mary, University of London.
- Cushing, S. (1995). Pilot-air traffic control communications. *Flight Safety Digest*, 14(7), 1-10.
- Diehl, A. (1991). The effectiveness of training programs for preventing aircrew "error". In *Proceedings of the Sixth International Aviation Psychology Symposium*. Columbus: Ohio State University, 640-655.
- Edwards, M. (1992). Crew coordination problems persist, demand new training challenges. *Flight Safety Foundation Cabin Crew Safety*, 27, 1-5.
- Faith, N. (1996). *Black box the final investigations*. London: Macmillan Publishers.
- Federal Aviation Administration. (1988). *Advisory Circular No. 120-48: Communication and coordination between flight crew members and flight attendants*. Washington, DC: Author.
- Federal Aviation Administration, (2004). *Advisory Circular No. 120-51E: Crew resource management training*. Washington, DC: Author.
- Foushee, H.C. (1986). The role of communications, socio-psychological, and personality factors in the maintenance of crew coordination. *Aviation, Space, and Environmental Medicine*, 53, 1062-1066.

- Foushee HC, & Helmreich RL. (1988). *Human Factors in Aviation*. San Diego, CA: Academic Press.
- Hackman J.R. (1993). *Teams, leaders, and organizations: New directions for crew-oriented flight training*. Harvard University: Boston.
- Hayward, B., (1994). The Dryden accident: A training video for pilots and flight attendants. *The CRM Advocate*, 94(1).
- Hartel, C., & Hartel, G. (1995). *Controller resource management: what can we learn from aircrews?* Washington, DC: Federal Aviation Administration.
- Helmreich, R. L., & Foushee, H.C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E. Wiener, B. Kanki, & R. Helmreich (Eds.), *Cockpit Resource Management*, (pp. 3-45). San Diego, CA: Academic Press.
- Helmreich, R. L., Butler, R. E., Taggart, W. R., & Wilhelm, J. A. (1994). *The NASA/University of Texas/FAA Line/LOS Checklist: A behavioral marker-based checklist for CRM skills assessment*. Austin, Texas: The University of Texas at Austin.
- Johnston, N. (1993). CRM: Cross-cultural perspectives. In E. Wiener, B. Kanki, & R. L. Helmreich (Eds.), *Crew Resource Management*, (pp. 367-398). San Diego, CA: Academic Press.
- Kanki, B.G., & Palmer, M.T. (1993). Communication and crew resource management. In E.L Wiener, B.G. Kanki, & R.L. Helmreich (Eds.), *Cockpit Resource Management*, (pp. 99-136). San Diego, CA: Academic Press.
- Kirvonos, P. D. (2005). Communication in aircraft cabin safety: lessons learned and lessons required. *Proceedings of the 22nd Annual International Aircraft Cabin Safety Symposium*. Los Angeles, CA.
- Kirvonos, P.D. (2007, June). *Lessons learned and lessons required*. Paper presented at the 2007 Regional Seminar of the Australia and New Zealand Societies of Air Safety Investigators, Wellington, New Zealand.
- Loftus, G.R. (1979). Short-term memory factors in ground controller/pilot communication. *Human Factors*, April, 21(2), 169-181.
- Moshansky, V. P. (1992). *Commission of inquiry into the Air Ontario crash at Dryden, Ontario: Final Report*. Government of Canada Publications (Catalogue No. CP32-55E). Toronto, Canada.
- Morrow, D., Rodvold, M., & Lee, A. (1994). Non-routine transactions in controller-pilot communication. *Discourse Processes*, 17(2), 235-258.
- National Transportation Safety Board (1994). *A review of flight crew-involved major accidents of U.S. air carriers, 1978 through 1990*. (NTIS Report No. PB94-917001) Washington, DC: Author.
- Orlady, H.W., & Orlady, L.M. (1999). *Human factors in multi-crew flight operations*. Aldershot, UK: Ashgate.
- Salas, E., Bowers, C.A., & Edens, E. (2001). An overview of resource management in organizations: why now? In E. Salas, C.A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations: Applications of resource management training* (pp. 1-5). London, UK: Lawrence Erlbaum Associates.

- Salas, E., Stagl, K., & Burke, S. (2004). 25 years of team effectiveness in organizations: Research themes and emerging needs. *International Review of Industrial and Organizational Psychology*, 19, 47-49.
- Sexton, J. B., & Helmreich, R. L. (2000). Analyzing cockpit communications: The links between language, performance, error, and workload. *Human Performance in Extreme Environments*, 5(1), 63-68.
- Smith-Jentsch, K., Baker, D., Salas, E., & Cannon-Bowers, J. (2001). Uncovering the differences in team competency requirements: the case of air traffic control teams. In E. Salas, C.A. Bowers, & E. Edens (Eds.), *Improving teamwork in organizations: Applications of resource management training*, (pp. 294-320). London, UK: Lawrence Erlbaum Associates.
- Tannen, D. (1990). *You just don't understand: women and men in conversation*. New York, NY: Quill/HarperCollins.
- Tannen, D. (1995). *Talking from 9 to 5: women and men at work*. New York, NY: Quill/HarperCollins.
- Tannen, D. (1996). *Gender & Discourse*. New York, NY: Oxford University Press.
- Tuckman, B. W., & Jensen, (1977). Stages of Small-group development revisited. *Group & Organization Management*, 2(4), 419.
- US Airways Flight 1549 accident: Hearing before the Subcommittee on Aviation of the Committee on Transportation and Infrastructure, U. S. House of Representatives, 111th Cong. 1 (2009). Washington, DC: Government Printing Office.
- Airline Safety and Pilot Training Improvement Act, 111 U.S.C. § HR 3371-1-18 (2009). Washington, DC: Government Printing Office.
- U. S. House of Representatives, (2010). Transportation and Infrastructure Committee Aviation Subcommittee. The Agency's Call to Action on Airline Safety and Pilot Training. Washington, DC: Government Printing Office.
- Wiener, E. L., Kanki, B. G., & Helmreich, R. L. (Eds.). (1993). *Cockpit resource management*. San Diego, CA: Academic Press.
- Wiener, E. L., and Nagel, D. C. (Eds.). (1988). *Human Factors in Aviation*. San Diego, CA: Academic Press.
- Wilson, J. (2005). *Gender Based Issues in Aviation*. (Doctoral dissertation, University of Pretoria). Retrieved from <http://upetd.up.ac.za/thesis/available/etd-03302005-094856/unrestricted/00front.pdf>.
- Young, L. B. (1994). Communications and pilot/flight attendant crew performance. *The CRM Advocate*. 94(1). Retrieved from [http://s92270093.onlinehome.us/crmdevel/resources/crmadvocate/94\\_1/94\\_1](http://s92270093.onlinehome.us/crmdevel/resources/crmadvocate/94_1/94_1).

**Previously Published Technical Report**

The article within this section of the IJAAS was originally published in whole or in part as technical report.

---

***Subtle Cognitive Effects of Moderate Hypoxia<sup>1</sup>***

Kate J. Hewett, MS

Ian P. Curry, BM BS, MSc

Steven J. Gaydos, MD, MPH

*Science Information Center*

*U.S. Army Aeromedical Research Laboratory*

*Fort Rucker, AL 36362*

*(334) 255-6883/6906*

*Fax: (334) 255-6993*

[catherine.machen\[at\]us.army.mil](mailto:catherine.machen[at]us.army.mil)

***Abstract***

*Increasing concern over the effects of moderate altitude exposure on aircrew performance in unpressurized aircraft recently prompted efforts to formulate international standards for the use of supplementary oxygen. The purpose of this study was to assess the impact of low to moderate levels of hypoxia on the cognitive performance of aircrew. Fifty participants were administered 45-minute exposures at altitudes of sea level, 8000, 10,000, 12,000 and 14,000 ft. Altitudes were simulated using the Reduced Oxygen Breathing Device. Participants completed the CogScreen®-Hypoxia Edition (CogScreen®-HE) to measure cognitive performance. Saturation of peripheral oxygen showed that although the participants did become hypoxic ( $p < .001$ ), there was no statistically significant change in reaction time ( $p = .781$ ), accuracy ( $p = .152$ ), or throughput ( $p = .967$ ) with increasing altitude. The results indicate that healthy individuals do not experience significant cognitive deficit, as measured by the CogScreen®-HE, when exposed to moderate levels of hypoxia at or below 14,000 ft.*

---

<sup>1</sup> Original publication citation: U. S. Army Aeromedical Research Laboratory. (2009). *Subtle Cognitive Effects of Moderate Hypoxia* (Defense Technical Information Center [DTIS] No. ADA511511). URL: <http://handle.dtic.mil/100.2/ADA511511>.

---

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](mailto:kay.chisholm@faa.gov).

Return to Table of Contents

## Introduction

Due to recent rotary-wing accidents in the current theatre of military operations, there is increasing worldwide concern over the effects of moderate altitude exposure in unpressurized aircraft. To address this issue, efforts have been made to formulate international standards for the use of supplementary oxygen (Air and Space Interoperability Council, 2007). Hypoxic hypoxia is a deficiency in alveolar oxygenation due to inadequate ventilation, ventilation-perfusion mismatch, or, in aviation operations, reduced partial pressure of oxygen in inspired air at altitude. At altitude, hypoxic hypoxia can result in a spectrum of symptoms including drowsiness, poor judgment, and impaired coordination to central nervous system failure, cardiovascular collapse, and death at extreme altitudes (Table 1).

Table 1

*Symptoms of hypoxia from Aeromedical Training for Flight Personnel (Department of the Army, 2000).*

Altitude (thousands of ft)	0 - 10	10 - 15	15 - 20	20 - 25
Estimated SpO <sub>2</sub>	99 – 90%	89 – 80%	79 – 70%	69 – 60%
Symptoms	Decrease in night vision	Drowsiness Poor judgment Impaired coordination efficiency	Impaired flight control handwriting speech vision intellectual function judgment Decreased coordination pain sensation memory	Circulatory failure CNS failure Convulsions

Research findings have led to a common assumption that aircrew can function perfectly well at altitudes up to 12,000 ft and even higher for limited periods (Bahrke and Shukitt-Hale, 1993; Reed, Youngs, and Kanid, 1994). Ernsting (1978) maintained that ascent to 10,000 feet (ft) produces no hypoxia symptoms in resting individuals. However, there is growing concern that hypoxia at moderate altitudes may cause cognitive deficits.

With air transport, personnel can be moved from sea level to over 10,000 ft within minutes, a far shorter time than required for acclimatization. In a survey of Australian helicopter aircrew, approximately 75% of physically active helicopter aircrew who returned surveys reported experiencing at least one hypoxic symptom during flight between 8000 and 10,000 ft (Smith, 2005). The survey also showed

non-pilot aircrew reported a significantly higher number of hypoxia symptoms than pilots. A follow-up study demonstrated that altitude hypoxia may be exacerbated greatly by physical exertion typical of the duties of aircrew personnel (Smith, 2006). These studies found hypoxia effects at altitudes previously thought to be too low for significant concern.

The Federal Air Regulations (FAR) require supplemental oxygen as follows (14 CFR § 91.211):

1. *At cabin pressure altitudes above 12,500 feet (MSL) up to and including 14,000 feet (MSL) unless the required minimum flight crew is provided with and uses supplemental oxygen for that part of the flight at those altitudes that is of more than 30 minutes duration;*
2. *At cabin pressure altitudes above 14,000 feet (MSL) unless the required minimum flight crew is provided with and uses supplemental oxygen during the entire flight time at those altitudes; and*
3. *At cabin pressure altitudes above 15,000 feet (MSL) unless each occupant of the aircraft is provided with supplemental oxygen.*

The crews of U.S. Army rotary wing aircraft in operations around the world are exposed to repeated incidences of moderate altitude (up to 18,000 ft). Current flight regulations (Department of the Army, 2008) list the following oxygen requirements at altitude in unpressurized aircraft:

*(1) Aircraft crews.*

- (a) On flights above 10,000 feet pressure altitude for more than 1 hour.*
- (b) On flights above 12,000 feet pressure altitude for more than 30 minutes.*

*(2) Aircraft crews and all other occupants.*

- (a) On flights above 14,000 feet pressure altitude for any period of time.*
- (b) For flights above 18,000 feet pressure altitude, oxygen prebreathing will be accomplished by aircrew members.*

To date, most of the literature has assessed gross cognitive change after multiple hours of exposure (e.g., Shukitt, Burse, Banderet, Knight, and Cyerman, 1988; Balldin, Tutt, and Dart, 2007), whereas current studies focus on subtle changes in cognition (e.g., Rice, et al., 2005; Balldin, Hickey, Sundstrom, Pilmanis, and Doan, 2006). A recent study demonstrated slight but statistically significant decrements in the cognitive performance of resting individuals for 20-minute (min) exposures at 12,000 ft (Balldin et al., 2006). Rice et al. (2005) sought to approximate the initial altitude a cognitive decrement would present at moderate altitudes as measured by the CogScreen®-Hypoxia Edition (CogScreen®-HE), and noted vigilance performance decrements at 15,000 ft.



A major aim of the current study was to expose participants to moderate levels of hypoxia in smaller increments than previous studies in an attempt to assess whether there is a gradual change in cognitive functions with increasing altitude. This information may more accurately inform policy and countermeasure strategies. The overall purpose of this study was to assess the impact of low to moderate levels of hypoxic hypoxia on the cognitive performance of aircrew personnel.

## Methods

The study was conducted by U.S. Army Aeromedical Research Laboratory (USAARL) personnel with logistical and technical assistance from the U.S. Army School of Aviation Medicine (USASAM). The protocol was reviewed and approved by the USAARL Human Use Committee. Fifty participants were evaluated during the study. Each participant was exposed to five normobaric simulated altitudes: sea level, 8000, 10,000, 12,000, and 14,000 ft while at rest conducting a cognitive test battery. The research intervention or independent variable that the research volunteers experienced was a condition of hypoxic hypoxia that simulated the amount of oxygen in the atmosphere at defined altitudes. These hypoxic conditions were generated with a Reduced Oxygen Breathing Device (ROBD). The ROBD (EnviroNics® Series 6202) is a portable, computerized, gas-blending instrument that produces hypoxia without changes in atmospheric pressure. It uses thermal mass flow controllers (MFC) to mix breathable air and medical nitrogen to produce the equivalent atmospheric oxygen partial pressures for altitudes up to 34,000 ft. The MFCs are calibrated on a primary flow standard traceable to the National Institutes of Standards and Technology.

The ROBD was developed by the Naval Aerospace Medical Research Laboratory (NAMRL) and is now marketed commercially by EnviroNics® for aviation training and research purposes. The ROBD enables individuals to be safely made hypoxic, without risk of barotrauma and decompression illness under controlled conditions in such a way that these individuals can engage in the performance-based testing procedures described below that are the dependent measures for this study. The ROBD is now routinely used by the Army and the Navy for refresher hypoxia training for aircrew personnel. The ROBD provides simulation of 0 to 34,000 ft elevation, 21% to 4.4% oxygen, an integrated pulse oximeter, an integrated oxygen analyzer, and an emergency oxygen dump switch for essentially instantaneous delivery of 100% oxygen.

### Study population

Data were collected on 50 participants. Volunteers were restricted to active duty Army Soldiers. Most participants were Army aviators, student aviators, or individuals waiting to begin Army flight training. The participants were aged 19 to 45 years. Pregnant individuals were excluded due to the risk of adverse effects of hypoxia on the fetus. To limit the effect of any confounding variables, participants were disqualified if they had a history of drug abuse, addiction, or consumed more than ten beers, eight glasses wine, or eight mixed drinks, per week.

## Experimental design

The study was a within-subjects repeated measures design in which 50 Soldiers were exposed to each of five normobaric simulated altitudes (sea level, 8000, 10,000, 12,000, and 14,000 ft) while wearing a pulse oximeter to measure saturation of peripheral oxygen (SpO<sub>2</sub>). Any possible order effect was prevented by blinding the participant to the simulated altitude to which they were exposed. These altitudes were assigned to the participants in a pseudo-random fashion ensuring the final totals of each altitude in each order were equal.

The purpose of recording SpO<sub>2</sub> was to ensure acclimation occurred at each altitude before cognitive testing started. The independent variable was altitude, as generated by the ROBD, and the dependent variable was performance on the CogScreen®-HE.

## Cognitive tests

CogScreen®-Aeromedical Edition (CogScreen®-AE) was designed for the Federal Aviation Administration (FAA) to detect subtle changes in cognitive functioning by, “rapidly assessing deficits or changes in attention, immediate- and short-term memory, visual-perceptual functions, sequencing functions, logical problem solving, calculation skills, reaction time, simultaneous information processing abilities, and executive functions” (Kay, 1995).

The CogScreen®-HE is a shortened version of the CogScreen®-AE specifically designed for detecting changes in cognitive functioning due to hypoxia. The CogScreen®-HE's touch-pen technology delivers rapid, non-invasive, validated, and sensitive cognitive tests that are appropriate for repeated measures testing. The CogScreen®-HE presents four subtests, visual sequence comparison, divided attention test, pathfinder combined, and matching to sample. The program administers the subtests three times resulting in a 30 min testing session. Following test completion, the CogScreen®-HE provides several performance scores derived from the four subtests. For the purpose of this study, only reaction time, accuracy, and throughput (number of correct responses per minutes) were used.

## Procedure

Data collection for a volunteer lasted one day. Table 2 describes scheduling during the experiment. Upon arrival to USASAM, participants read and signed informed consent forms and were given the opportunity to ask the researchers questions. Participants then spent approximately 1 hour training and practicing the CogScreen®-HE. The practice session did not involve exposures to any altitude other than ambient room air (roughly 350 ft above sea level). Practice sessions ensured that test performance was asymptotic and that the measurements were made with maximum efficiency.

Following training, individuals were exposed to the five hypoxic conditions: sea level, 8000, 10,000, 12,000, and 14,000 ft. Each exposure simulated flight at specified altitude and lasted for 45 min total. This was broken down into 15 min at rest to equilibrate to the altitude and 30 min for completion of the CogScreen®-HE.

After each exposure to altitude, participants rested for 15 min, breathing ambient room air before starting the next altitude condition.

Table 2

*Testing itinerary.*

<b>Time</b>	<b>Activity</b>
08:00	In-Processing and Informed Consent
08:30	CogScreen®-HE Practice
09:00	Hypoxia Condition 1* and Cognitive Testing
10:00	Hypoxia Condition 2* and Cognitive Testing
11:00	Lunch
12:00	Hypoxia Condition 3* and Cognitive Testing
01:00	Hypoxia Condition 4* and Cognitive Testing
02:00	Hypoxia Condition 5* and Cognitive Testing
03:00	Out-Processing and Release

*\*Note: The hypoxia conditions (sea level, 8000, 10,000, 12,000, and 14,000 ft) were randomly presented to each participant.*

In the event that a participant's peripheral arterial oxygen saturation fell below 70%, oxygen would be increased to the equivalent of 13,000 ft to safeguard the health of the volunteer. This procedure was not required as all participants maintained their oxygen saturation above 70%.

## **Results**

All statistical analyses were performed using SPSS® 13.0 with significance set at an alpha level of .05 for all statistical tests. A repeated measures analysis of variance (ANOVA) was conducted to evaluate the impact of altitude on SpO<sub>2</sub>. Saturation of peripheral oxygen declined with increasing simulated altitude,  $F(2.220, 108.77) = 155.675, p < .001$ , with the Greenhouse-Geisser correction. These findings confirm the efficacy of the ROBD system and that participants were, indeed, hypoxic. Follow-up results that emerged using paired samples *t*-tests confirmed significance ( $p < .001$ ) at all levels of altitude.

Repeated measures ANOVA were conducted to evaluate the impact of altitude on cognitive performance, as measured by the CogScreen®-HE. The dependent variables were reaction time, accuracy, and throughput as measured by the CogScreen®-HE. No significant effect was found between altitude and reaction time,  $F(4,192) = .437, p = .781$  (Figure 1). Likewise, a non-significant effect emerged for altitude and accuracy, with a Greenhouse-Geisser correction,  $F(2.193, 105.245) = 1.889, p = .152$  (Figure 2). No significant effect was found between altitude and throughput,  $F(4,192) = .140, p = .967$  (Figure 3).

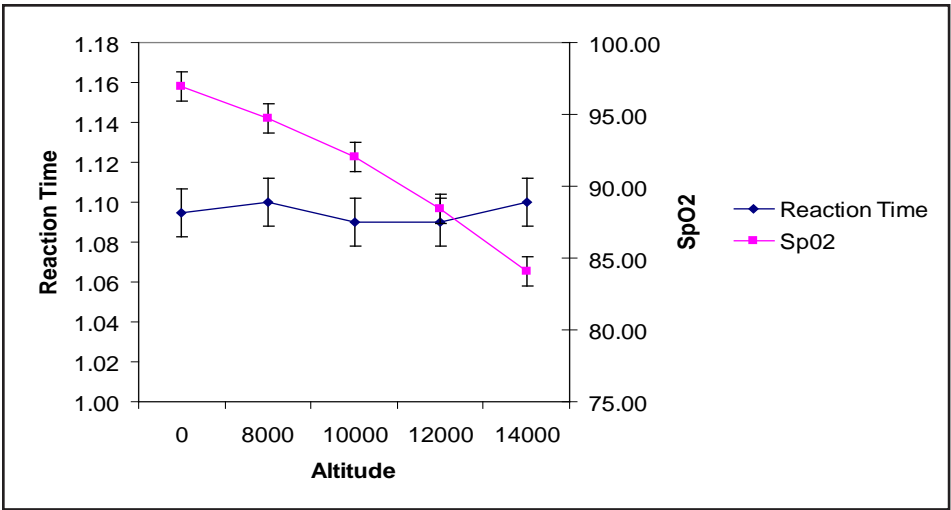


Figure 1. Graphical representation of the relationship between reaction time, SpO2 and altitude. Standard error bars for each mean are also shown.

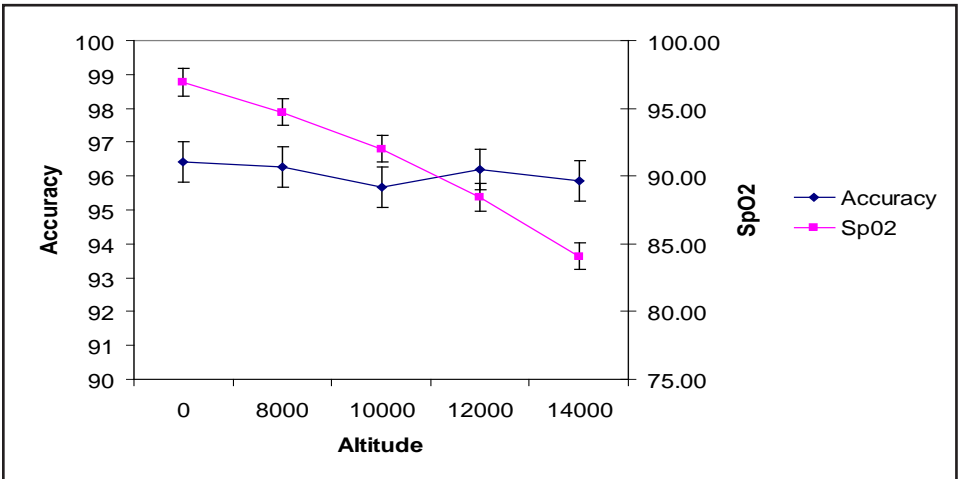


Figure 2. Graphical representation of the relationship between accuracy, SpO2 and altitude. Standard error bars for each mean are also shown.

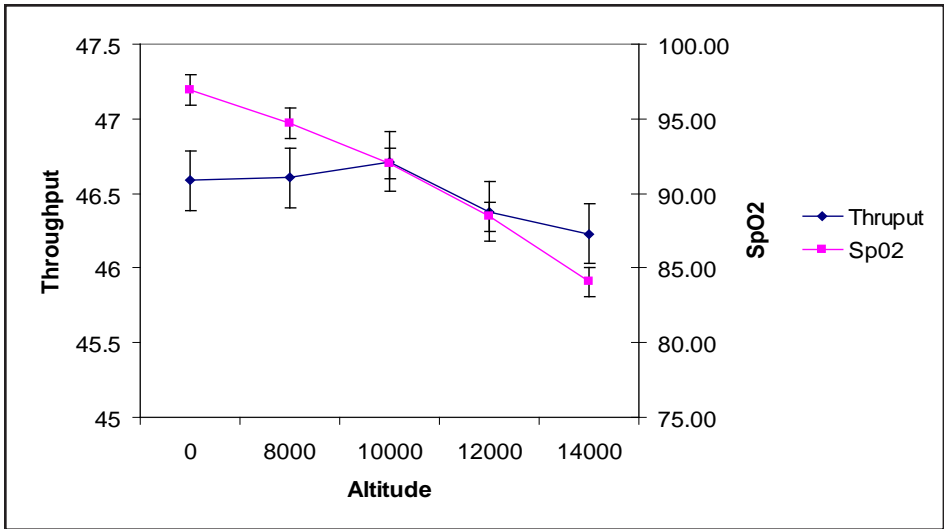


Figure 3. Graphical representation of the relationship between throughput, SpO2 and altitude. Standard error bars for each mean are also shown.

### Discussion

Results from this study demonstrated that moderate hypoxia at altitudes of 8000 to 14,000 ft does not significantly decrease cognitive performance as measured by the CogScreen®-HE. These findings suggest that current standards regulating supplemental oxygen use sufficiently protect aircrew from cognitive performance decline in the unpressurized cockpit. However, these results may be due to this specific cognitive test battery not being sensitive enough to detect subtle changes in performance due to low altitude hypoxia. Rice et al. (2005) used the CogScreen®-HE to estimate the altitude at which cognition degradation occurs. Sixty resting aviators' scores at 10,000 ft, 12,000 ft, and 15,000 ft were compared to their baseline scores. The only significant finding was in accuracy during the Vigilance subtest for 15,000 ft and the baseline scores ( $p = 0.012$ ). Analysis of reaction time and accuracy indicated no significant differences. Further research is needed to determine at what specific altitudes the CogScreen®-HE is able to detect cognitive degradation.

Another possible explanation for the lack of significance is that participants were given a practice session. Hypoxia is known to affect the learning process, and because participants were given a practice session before testing, the CogScreen®-HE tasks were no longer novel. Denison, Ledwith, and Poulton (1966) attributed increased reaction time in exercising subjects at 8000 ft to task novelty. Denison et al. found subjects who had a practice session at sea level on the Manikin test performed better at 8000 ft than those subjects who did not have a practice session. Similarly, Kelman and Crow (1969) found that impairment of mental performance, as measured by a vigilance task, occurred at 8000 ft. However, subsequent studies by Fowler et al. (1985), using the same study design as Denison, failed to demonstrate learning difficulties up to 12,000 ft. Figarola and Billings (1966) found no impairment on practiced tracking and vigilance tasks at 8000 ft; however, they did find performance decrements at 17,000 ft. In a study on both resting and exercis-

ing subjects, Paul and Fraser (1994) found that the ability to learn new tasks is not impaired by mild hypoxia at altitudes up to 12,000 ft.

In addition, perhaps the critical altitude, which causes marked performance decrements was not reached in this study. According to Nelson (1982), the decisive altitude for changes in higher cognitive functioning lies between 4000 and 5000 meters (13,123 ft and 16,404 ft, respectively). Even at 4500 meters (14,764 ft), Pavlicek et al. (2005) found no significant difference in word fluency, word association, or lateralized lexical decision performances. In addition, Schlaepfer, Bartsch, and Fisch (1992), found that mild hypoxia improved visual perception in healthy individuals. If a testing session at 15,000 ft had been incorporated in the experiment and significance was found at that altitude, the experimenters would not only know that the CogScreen®-HE was sensitive enough to detect changes in performance due to hypoxia, but it would show that, on certain tasks, performance was not negatively impacted by moderate hypoxia (8000 ft to 14,000 ft).

Many studies on moderate altitude hypoxia merely record cognitive performance, and not subjective symptoms experienced at altitude. Some of the participants reported experiencing symptoms of hypoxia, particularly at 14,000 ft (e.g., slight light-headedness and minor headaches). Although participants experienced hypoxic symptoms, their cognitive performance on the CogScreen®-HE was not significantly compromised. These reported symptoms came up in conversation between test sessions and were not recorded for later analysis. It is conceivable that, similar to Smith's subjective survey study (2005), levels of hypoxia assessed in the present study may solely impact the psychological perception of hypoxia and not the measurable, objective cognitive consequences. Further research is needed to compare both the perception of hypoxic symptoms to objective decrements in cognitive performance.

Crewmembers in the cabin of the aircraft are rarely stationary, and the cognitive effects of moderate altitude may be exacerbated by increased heart rate and metabolic demands due to physical movement often required of flight medics and crew chiefs. Physical exertion accelerates the onset of hypoxia and lowers the altitude at which symptoms occur. Paul and Fraser (1994) found exercising subjects' reaction time to be slower on the Manikin task than resting subjects. Smith (2005) surveyed Australian Army helicopter pilots and found 60% of non-pilot aircrew reported experiencing four or more hypoxic symptoms, compared to only 17% of pilots. The most common symptoms or areas of deficit experienced were light-headedness (37.7%), calculation (45.3%), and reaction time (37.7%). Further research is needed to examine how physical exertion affects cognitive performance at moderate altitudes.

## **Conclusion**

Healthy individuals aged 19 to 45 years did not experience significant cognitive deficits as measured by the CogScreen®-HE when exposed to moderate levels of hypoxia at rest for exposure times of 45 min at or below 14,000 ft.

## **Acknowledgements**

The authors wish to thank Colonel Robert Noback, Dean of the United States School of Aviation Medicine (USASAM), Major Mike Mendenhall, Chief of USASAM Physiology Section, and Staff Sergeant Thomas Rath, Atmospheric Noncommissioned Officer, USASAM.

## **DISCLAIMER**

The views, opinions, and/or findings of this study are those of the author(s) and should not be construed as an official Department of the Army (DA) or Department of Defense (DOD) position, policy, or decision, unless so designed by other official documentation. Citation of trade names does not constitute an official DA or DOD endorsement or approval of the use of such commercial items.



## References

- Air and Space Interoperability Council. (2007). Mitigation of Hypoxia in Unpressurized Aircraft at Altitudes below 18,000 feet. Draft Advisory Publication WG ASMG 07.
- Bahrke, M.S., & Shukitt-Hale, B. (1993). Effects of Altitude on Mood, Behavior and Cognitive Functioning. *Sports Medicine*, 16(2): 97-125.
- Balldin, U.I., Hickey, P., Sundstrom, J., Pilmanis, A., & Doan, B. (2006). Cognitive effects during night flying at lower altitudes without cabin pressurization. *Aviation Space and Environmental Medicine*, 77: 274.
- Balldin, U.L, Tutt, R.C, & Dart, T.S. (2007). Effects of 12 hours of Low-grade Hypoxia at 10,000 ft at Night in Special operations Forces Aircraft Operations on Cognition, Night Vision Goggle Vision and Subjective Symptoms. AFRL Technical Report No: 28-06-2007.
- Denison, D.M., Ledwith, F., & Poulton, E.C. (1966). Complex reaction times simulated cabin altitudes of 5,000 feet & 8,000 feet. *Aerospace Medicine*, 37(10): 1010-1013.
- Department of the Army. (2008). Flight Regulations (Army Regulation 95-1).
- Department of the Army. (2000). Aeromedical Training for Flight Personnel (Field Manual 3-04.301).
- Ernsting, J. (1978). Prevention of hypoxia-acceptable compromises. *Aviation, Space and Environmental Medicine*, 55: 407-410.
- Figarola, R. R., & Billings, C. E. (1966). Effects of meprobamate and hypoxia on psychomotor performance. *Aerospace Medicine*, 37: 951-954.
- Fowler, B., Paul, M. A., Porlier, B., Elcombe, D., & Taylor, M. (1985). A reevaluation on the minimum altitude at which hypoxic performance decrement can be detected. *Ergonomics*, 29(1): 61-66.
- Kay, G.G. (1995). Cogscreen Aeromedical Edition Professional Manual.
- Kay, G., Horst, R., Pakull, B., & Hordinsky, J. (1991). Automated Cognitive Function Assessment of Civilian Pilots. Paper presented to the 62nd Annual Scientific Meeting of the Aerospace Medical Association, Cleveland, Ohio.
- Kay, G., Strongin, G., Hordinsky, J., & Pakull, B. (1993). Georgetown/Russian Collaborative Project: Development of Aviator Norms for CogScreen, Paper presented at the 64th Annual Scientific Meeting of the Aerospace Medical Association, Toronto, Canada.
- Kelman, G. R., & Crow, J. (1969). Impairment of mental performance at simulated altitude of 8,000 ft. *Aerospace Medicine*, 40(9): 981-982.
- Nelson, M. (1982). Psychological testing at high altitudes. *Aviation, Space, and Environmental Medicine*, 53: 122-126.
- Paul, M.A, & Fraser W.D. (1994). Performance during mild acute hypoxia. *Aviation, Space, and Environmental Medicine*, 65(10): 891-898.
- Pavlicek, V., Schirlo, C., Nebel, A., Regard, M., Koller, E. A., & Brugger, P. (2005). Cognitive and emotional processing at high altitude. *Aviation, Space, and Environmental Medicine*, 76 (1): 28-33.

- Reed, E., Youngs, W.J., & Kanid, B.G. (1994). Cognitive Defects at Altitude. *Nature* 372, 325.
- Rice, G.M., Moore, J.L., Jernigan, C., Moore, J.L., Clemons, E., Rife, C., & Kay, G.C. (2005). Cognitive performance at simulated altitudes of 10,000, 12,000, & 15,000 ft. utilizing the CogScreen- Hypoxia Edition (CogScreen-HE). *Aviation, Space, and Environmental Medicine*, 76(3): 231.
- Schlaepfer, T. E., Bartsch, P., & Fisch, H. U. (1992). Paradoxical effects of mild hypoxia and moderate altitude on human visual perception. *Clinical Science*, 83: 633-636.
- Shukitt, B.L., Burse, R.L, Banderet, L.E, Knight, & D.R, Cymerman, A. Cognitive Performance, Mood States and Altitude Symptomatology in 13-21% Oxygen Environments. USARIEM Report SGRD-UE-HP, June 1988.
- Smith, A.M. (2005). Hypoxia symptoms reported during helicopter operations below 10,000 feet: A retrospective survey. *Aviation Space and Environmental Medicine*, 76: 794-798.
- Smith, A.M. (2006). The impact of short-duration light and moderate physical activity at 7,000 ft and 9,000 ft on the development of hypoxia. *Aviation Space and Environmental Medicine*, 77: 274.

