2009 *International Journal of Applied Aviation Studies* Volume 9, Number 2 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 9, Number 2, 2009

REVIEW PROCESS

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

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

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

International Journal of Applied Aviation Studies
Volume 9, Number 1
ISSN Number: 1546-3214

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	Gary Condley
Deputy Superintendent	Sunny Lee-Fanning

PUBLICATION COMMITTEE

Manager, Regulatory Standards Division	David Long
Manager, Technical Operations Training Division	Jessie McMullen
Manager, Air Traffic Division	William J. Mumper
Manager, Airports & International Training Division	Sherry Reese

EDITORIAL STAFF

Managing Editor	Kay Chisholm
Associate Editor	Deann King
Research Consultant	Todd P. Hubbard

CONSULTING EDITORS

International

Ruwantissa I.R. Abeyratne	ICAO, Montreal, Canada
Pierre R. Band	Institut du Cancer de Montréal, Québec, Canada
Simon Bennett	Institute of Lifelong Learning, Leicester, UK
Robert Bor	Royal Free Hospital, London, UK
Peter Brooker	Cranfield University, Cranfield, UK
Bryan Burke	Macquaire University, Sydney, Australia
Michael B. Charles	Southern Cross University, Tweed Heads, NSW Australia
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
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
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
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

Wen-Chin Li	National Defense University, Taipei, Taiwan, R.O.C.
Jörg Leonhardt	Deutsche Flugsicherung GmbH, Langen, Germany
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
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
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
Richard H. Y. So	University of Essex, Colchester, UK
M. "Mattie" Tops	Leiden University, Leiden, The Netherlands
Steve Thatcher	University of South Australia, Mawson Lakes, Australia
Claudia van de Wal	VALK Foundation, Netherlands
Lucas van Gerwen	VALK Foundation, Netherlands
Leo Vermeulen	University of Pretoria, Pretoria, South Africa
Stephen Walsh	Interaction Training Associates, Netherlands
Damien J. Williams	University of Bristol, Bristol, UK
Robert Wolfger	Austrian Airlines, Strasshof, Austria
Rogier Woltjer	Linköping Institute of Technology, Linköping, Sweden
Beat Zimmermann	Air Navigation Institute, Thun, Switzerland

U.S.A.

Robert I. Aceves	St. Cloud State University, St. Cloud, MN
Vicki Ahlstrom	FAA, William J. Hughes Technical Center, Atlantic City, NJ
Amy Alexander	Aptima, Inc., Woburn, MA
Steve Anderson	St. Cloud State University, St. Cloud, MN
Rick Anglemyer	SCSI, Colorado Springs, CO
Francis "Frank" Ayers	Embry-Riddle Aeronautical University, Daytona Beach, FL
Larry Bailey	FAA Academy, Oklahoma City, OK
R. Kurt Barnhart	Kansas State University, Salina, KS
Robert Baron	The Aviation Consulting Group, Myrtle Beach, SC
Ellen Bass	University of Virginia, Charlottesville, VA
J. Matthew Beaubien	American Institutes for Research, Washington, DC
Wendy Beckman	Parks College, St. Louis University, St. Louis, MO
Dennis B Beringer	Civil Aerospace Medical Institute, Oklahoma City, OK
Robert O. Besco	Professional Performance Improvement, Dallas, TX

Aleta Best	Department of Transportation, Washington, DC
Kate Bleckley	Civil Aerospace Medical Institute, Oklahoma City, OK
James P. Bliss	Old Dominion University, Norfolk, VA
Erin E. Block	Purdue University, West Lafayette, IN
Deborah Boehm-Davis	George Mason University, Fairfax, VA
Philip Bos	University of Central Missouri, Warrensburg, MO
Brent Bowen	University of Nebraska at Omaha, Omaha, NE
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 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
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
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
John Deaton	Florida Institute of Technology, Melbourne, 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
Nancy Dorighi	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
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
A. F. Grandt	Purdue University, West Lafayette, IN
Mavis Green	Harvard University, Cambridge, MA
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
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
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
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	Nortwestern 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
Chien-tsung Lu	Central Missouri State University, Warrensburg, MO
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
Steven Marks	Oklahoma State University, Stillwater, OK
Royce Ann Martin	Bowling Green State University, Bowling Green, OH
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
Mark I. Nikolic	The Boeing Company, Seattle, WA
Michael S. Nolan	Purdue University, West Lafayette, IN
Gary J. Northam	Parks College, St. Louis University, St. Louis, MO
Dale B. Oderman	Purdue University, West Lafayette, IN
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
Allen J. Parmet	Midwest Occupational Medicine, Kansas City, MO
Manoj Patankar	Parks College, St. Louis University, St. Louis, MO
Donald A. Petrin	Purdue University, West Lafayette, IN
James J. Picano	Mind Quest, LLC, Suisun, CA
Jean Potvin	Saint Louis University, St. Louis, MO
Julia Pounds	Civil Aerospace Medical Institute, Oklahoma City, OK
Roni Prinzo	Civil Aerospace Medical Institute, Oklahoma City, OK
Edward Pugacz	FAA, William J. Hughes Technical Center, Atlantic City, NJ
Yandong Qiang	Johns Hopkins University, Baltimore, MD
Stephen M. Quilty	Bowling Green State University, Bowling Green, OH
William B. Rankin II	Florida Memorial College, Miami, FL
Esa M. Rantanen	Rochester Institute of Technology, Rochester, NY
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
Jason J. Saleem	Indiana Univ.-Purdue University, Indianapolis IN
Steven M. Samuels	US Air Force Academy, Colorado Springs, CO
Nadine Sarter	University of Michigan, Ann Arbor, MI

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	Ohio University Airport, Albany, OH
John W. Sheremeta, Jr.	American Airlines, Lake Ronkonkoma, NY
Mark A. Sherman	Ohio University, Athens, OH
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	Parks College, St. Louis University, St. Louis, MO
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
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
Richard D. Wright	Volpe National Transportation Systems Center, Weston, MA
John Young	Purdue University, West Lafayette, IN

PHILOSOPHY STATEMENT

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

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

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

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

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

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

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

Go to Table of Contents

EDITOR’S NOTES

Papers

In *An Evaluation of Collision Avoidance Technologies Using Empirical Function Allocation,* Hardman, Colombi, Jacques, Hill, and Miller present a quantitative method that uses empirical data to more objectively determine the function allocation between human operators and computers. The method was applied in evaluating traffic collision avoidance for unmanned aircraft.

The results of a study conducted at Middle Tennessee State University for the Commercial Pilot Certificate are presented by P.A. Craig in *Evaluating Pilots Using a Scenario-Based Methodology: A Guide for Instructors and Examiners.* The effectiveness of the scenario-based training methodology is analyzed and some "best practices" are presented.

In a nationwide survey on the use of the Microsoft Flight Simulator software by pilots, over 85% of responding pilots indicated that they use MSFS to preview approaches at unfamiliar airports. In *Pilot Perspective on the Microsoft Flight Simulator for Instrument Training and Proficiency,* W. S. Beckman this and other findings from the survey.

In *Motion Sickness Prevention by Stroboscopic Environment during Simulated Military Transport,* Webb, Estrada, and Athy assess the use of 4 and 8 hertz (Hz) stroboscopic environments as countermeasures for MS. The motion profiles of a Black Hawk helicopter and an amphibious vehicle were produced using a Multi Axis Ride Simulator (MARS).

The causes of deviations in aircraft maintenance procedures and their implications are discussed in R. I. Baron's *An Exploration of Deviations in Aircraft Maintenance Procedures.* The author addresses organizational pressure, individual complacency, and deficiencies in aircraft maintenance documentation itself.

In *Workplace Preferences of Millennials In the Aviation Industry,* M. Niemczyk and J. W. Ulrich examine the generational differences to determine work environment preferences in the aviation industry. Results of this study portray a complex combination of relationship, personal growth, and organizational structures that determine ideal workplace preferences.

A.J de Voogt and B. Heijnen analyze general aviation accidents with small aircraft occurring over the Pacific Ocean in *A Review of General Aviation Accidents in Pacific Ocean Operations.* Differences between helicopter and airplane operations are highlighted and the authors discuss the causes and prevention of general aviation emergencies over the open ocean.

KC

International Journal of Applied Aviation Studies

Volume 9, Number 2 2009

Papers

International Journal of Applied Aviation Studies, Volume 9, Number 2

An Evaluation of Collision Avoidance Technologies Using Empirical Function Allocation

Nicholas Hardman, John Colombi, David Jacques, Raymond R. Hill
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433-7765
USA
nhardman@afit.edu
and
Janet E. Miller
711th Human Performance Wing
Human Effectiveness Division
Wright-Patterson AFB, OH 45433-7765
USA

Abstract

The growing capability and complexity of automation in aviation necessitates an improved process for analyzing the function allocation between human operators and computers. This paper describes a quantitative method that uses empirical data for these decisions, thus making the process more objective and adaptable to changes in mission or technology. We apply this method to evaluate the issue of traffic collision avoidance for unmanned aircraft. The results demonstrate that some automation options meet the equivalent level of safety of piloted aircraft, but currently no option independently meets the level of safety mandated for operation in unrestricted airspace.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government. This work was sponsored by the 711th Human Performance Wing.

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.

Return to Table of Contents

An Evaluation of Collision Avoidance Technologies
Using Empirical Function Allocation

Though function allocation is critical in new system development, it remains insufficiently supported by quantitative methods. Function allocation involves matching system functions, actions, and decisions with hardware, software, humans, or some combination of them (NASA, 2007). New system development efforts are seldom a blank slate; they often have function allocation constraints due to budget or compatibility considerations. This often leaves the designer with a limited number of explicit allocation decisions; however, these decisions have far-reaching influence over the entire system architecture. Function allocation must be done in a manner that correctly balances cost, schedule, and performance with an acceptable level of safety. If one uses empirical data for function allocation decisions, those decisions will be more objective and will enable easier adaptation of new missions or technology.

A significant subset of function allocation concerns the distribution between humans and computers, that is, automation analysis. Automation is defined as the execution by a machine agent of a function previously carried out by a human (Scerbo, 1996). It is difficult, but possible, to quantify the effect of automation on cost and schedule requirements. The designer can measure the expected effect of automation on operational efficiency and expense and compare that with the expected penalties of higher equipment costs, more complex integration, and longer testing schedules. The alternative option, manual execution, consequently places more of the cost and schedule burden on labor, personnel, training, and human factors. We propose a method to assist designers in making function-allocation decisions between human operators and computers. The method is quantitative and considers system context. We then demonstrate this method in examining the issue of traffic collision avoidance for unmanned aircraft.

Method Overview

This method is intended to be performed in the framework of a comprehensive systems engineering effort. This method involves the following six steps:

1. The first step is to classify the lowest level functions of the functional decomposition as machine-only, human-only, or either. With computers becoming increasingly ubiquitous, a machine-only function has become a computer-controlled function. Examples include demodulate transponder signal or charge battery. Examples of human-only functions are express preferences or determine intent. These allocation decisions occur during the initial system architecting effort. Function allocation decisions for machine-only functions are relatively straightforward. They are based on such issues as the choice of distributed or hierarchical design and the details of the network structure.

2. For functions that indicate human involvement, more analysis is necessary. The next step is to perform a preliminary task analysis on these functions. This gives a better understanding of the *what* before the *how* or *who* of function allocation. Task analysis is much more mature than methods of function allocation, and there are multiple accepted techniques. For a thorough review of these see Stanton, et al. (2005). For human-only functions, this data is useful for user interface design. The remaining functions, those in the

either category, require an analysis for automation. To do this, the function must be further decomposed into information processing stages. One common nomenclature for these stages is sensor, processor, decision-maker, and actuator (Parasuraman, Sheridan, & Wickens, 2000). The dynamic between human and computer for a function can be very different based on the allocation of these stages.

3. Next, designers must quantify performance criteria for each information processing stage of each function under study. This forms the threshold and objective performance requirements for automation analysis. This quantification must consider overall workload and an acceptable margin of safety. For complex scenarios, the use of established methods of cognitive task analysis will form a better picture of workload (Crandal & Klein, 2006). A test program on the composite system in realistic scenarios could be cost prohibitive. To avoid this, we quantify the critical technical parameters necessary to achieve success and use those values as performance requirements. Using these lower-level criteria reduces time, cost, and makes laboratory and ground testing possible. There are several different performance models to compare performance (Rouse & Rasmussen, 1981).

4. The next step is to quantify the performance of humans for each task in the functional decomposition. These will be specific to each function and will vary under diverse real world conditions. In the application section, we show how empirical data from basic research in physiology and psychology can be used to quantitatively characterize human capabilities and limitations. Though existent data is wide-ranging, this step may require additional, more specific, basic research.

5. Next, the performance of machines (sensors, processors, or actuators) for the same tasks is quantified. This data is obtained through trade studies and consultations with industry.

6. The final step is to compare the alternatives and make the allocations. In the past, designers approached automation as a way to eliminate the supposedly inefficient and untrustworthy presence of humans. This migration from this paradigm is discussed in Dixon & Wickens (2006). The primary fault is that this paradigm fails to appreciate that the human components add necessary resilience for unexpected situations. Anyone that has ever observed a robot stuck in a rut can appreciate the need for system resilience. A modern paradigm is that automation exists to augment human capabilities and to assist the operator in achieving system goals (Dixon & Wickens, 2006). The level of automation resides on a spectrum between the machine-only and human-only extrema. There are multiple sources for the delineation of this spectrum (Albery & Khomenko, 2003; Miller & Parasuraman, 2007; Parasuraman et al., 2000). In operator role theory, the four levels of automation are defined as direct performer, manual controller, supervisory controls, and executive controller (Folds, 1995). These are determined for each information-processing stage. Figure 1 shows this graphically. One may also need to examine the need for adaptive automation. Adaptive automation may be necessary when operator workload is widely variant (Scerbo, 1996).

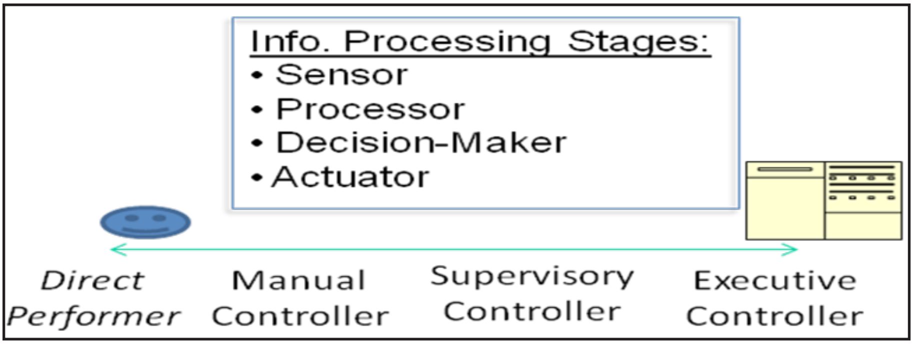

Figure 1. A graphical portrayal of operator role theory stages and levels

Application

We now use this method to guide design decisions for traffic collision avoidance in a new multi-role unmanned aerial system (UAS)[1]. While the method can apply to a broad range of systems, the unique challenges of UAS design make it a great example application to highlight the advantages of the proposed methodology over current practice. The new UAS, being considered by the US Air Force (USAF), must have the capability to transition civil airspace[2] (Jean, 2009). Industry officials state that the goal for UAS performance in the civil airspace is an equivalent level of safety, including collision avoidance, when compared to piloted aircraft (Warwick, 2004); however, the 2009 aircraft collision over the Hudson River in New York reignited the debate over the sufficiency of collision avoidance in even piloted operations (Baker & Grynbaum, 2009).

Function Classification

The first step of the method is to classify the functions regarding human and machine potential interaction. The regulations governing traffic deconfliction are contained within the "Right of Way Rules" section of the federal aviation regulations (FAA, 2005). By design, there are four layers that maintain safe operations in aviation. They are airspace regulation, air traffic service (ATS), cooperative avoidance systems, and finally, "see and avoid" as the last line of defense.

Airspace Regulation. One of the primary ways that aviation authorities attempt to avoid traffic collisions is by defining airspace. Table 1 explains the airspace classes as designated by the International Civil Aviation Organization (ICAO). Though pilots are still required to practice see and avoid, the responsibility for conflict avoidance changes with the type of airspace. In addition, there is a speed limitation of 250 knots below 10,000 feet above sea level and 200 knots in Classes C & D (ICAO, 2001).

---

[1] Unmanned aerial system (UAS) is the current official designation for what was known as a unmanned aerial vehicle (UAV) and its support components. When specifically the aircraft, or vehicle, is intended, UAV is appropriate. It must be noted that some are advocating a change to the term remotely piloted vehicle (RPV).

[2] Civil airspace is used to define that airspace under the control of a civil aviation authority, sometimes called the national airspace system (NAS), as compared to restricted or special use airspace.

As Table 1 shows, safe *unmanned aerial vehicle (*UAV) operation in Class B or C is technically possible and would only be necessary for short transitions. Due to the high traffic volume, however, there is likely to be continued opposition to granting widespread access to UAVs. Currently, UAVs are not allowed in any class of airspace without special permission. In the US, the FAA requires a certificate of authorization (COA) which requires at least a 30-day notice to local administrators, visual meteorological conditions (VMC)[3], a route clear of all populated areas, and constant ground control by a certified pilot (FAA, 2002). Therefore, for regulatory reasons, route planning and airspace situational awareness must remain a human-only function. For example, the American military's Global Hawk UAV is given frequent approval to travel in Class A airspace over most of the world and with a minimum of pre-coordination, but the flight request must verify that it will first climb to altitude within restricted airspace (AF Tech, 2009).

Table 1

ICAO Airspace Designations and UAV Operations (ICAO, 2001)

Class	Flight Ops	ATC equipment & services provided	Radio Required	Transponder Required	UAV integration problems
A	IFR only	Radar Conflict resolution & separation	Yes	Yes	None ACAS primary
B	IFR/VFR	Radar Conflict resolution & separation	Yes	Yes	No technical problem, but traffic density increases risk. ACAS primary
C	IFR/VFR After contact	Radar Separation (IFR), or traffic advisories (VFR)	Yes	Yes	No technical problem, but traffic density increases risk. ACAS primary
D	IFR/VFR After contact	Tower Separation (IFR only)	Yes	No	ACAS insufficient, DSA primary
E	IFR/VFR	Separation (IFR only)	Yes	No (<10,000´)	ACAS insufficient, DSA primary
F	IFR/VFR	Traffic advisories (IFR)	No (<10,000´)	No (<10,000´)	ACAS insufficient, DSA primary
G	VFR	None	No (<10,000´)	No (<10,000´)	Moderate problem due to lack of coverage, *DSA primary*

Notes: IFR—instrument flight rules, VFR—visual flight rules. ACAS—Automated collision avoidance system. DSA—detect, see and avoid.

---

[3] Instrument Flight Rules (IFR), Visual Flight Rules (VFR), Instrument Meteorological Conditions (IMC), and Visual Meteorological Conditions (VMC) are defined more completely in FAR 91.

Air Traffic System. The second layer of safety for all aviation is the air traffic service or air traffic system (ATS)[4]. The system involves a great deal of technology. In addition to the primary surveillance radars that detect all traffic by reflected radar energy, the secondary surveillance radars detect cooperating traffic transponder returns. Automated ATS systems analyze these radar returns to predict possible collisions. However, for the near future, this level will continue to involve human-only functions ranging from flight plan coordination and approval to traffic conflict intervention.

Many technical and regulatory changes are planned for the future ATS. One of these technical concepts involves inter-aircraft data links that would allow automated route deconfliction. These proposed future capabilities are analyzed in Hardman (2006). Unfortunately for UAVs, these changes do not appear to eliminate the need for independent deconfliction capability because (a) participation in the deconfliction data links will not be universally mandatory and (b) all aircraft will still require a backup capability in the event of network failures.

Cooperative Traffic Avoidance. The third layer consists of onboard systems that cooperatively work to deconflict traffic. These systems are independent of, but compatible with, ATS systems. Such systems would enable automated deconfliction between aircraft. The current weakness of cooperative systems is the necessity for all aircraft involved to have a compatible functioning system. Multiple proposals are being explored for ways to add information on nonparticipatory traffic. This has great potential as an independent automated traffic control technology, but for the near future, it cannot function as the sole source of deconfliction.

Detect, See and Avoid (DSA). The last layer of safety is the independent ability for each aircraft to detect, see, and avoid (DSA)[5] other aircraft. In uncontrolled airspace, the inherent freedom means that all responsibility for separation lies with the pilot. In other airspace, the see and avoid principle is still required to be practiced to the maximum extent possible. This principle is not specifically mentioned in ICAO regulation, but the FAA describes it as, “When weather conditions permit, regardless of whether an operation is conducted under IFR or VFR, vigilance shall be maintained by each person operating an aircraft so as to see and avoid other aircraft” (FAA, 2005, ¶ 91.113b). This regulation is satisfied in piloted aircraft when crewmembers perform a disciplined scan out the windshield (though more advanced aircraft also have augmenting technology, such as radar). DSA on a UAS is more complicated because any manual steps must be done remotely.

Preliminary Task Analysis.

This second step is to perform a preliminary task analysis on the functions under study. For the UAS traffic avoidance scenario, the DSA function needs further

---

[4] This has traditionally been referred to as air traffic control (ATC), but its future name has been deemed air traffic management (ATM) to highlight the eventual evolution towards less controlling and more managing.

[5] Some sources use the alternate terms: sense-and-avoid (SAA), non-cooperative collision avoidance, or traffic deconfliction when referring to new systems. “See-and-avoid” is the primary term used in the regulatory sense. In this report the term DSA will be used throughout to mean any concept or system in effect for the primary purpose of preventing collisions between aircraft that have not been deconflicted by other means.

examination. Since see and avoid is already performed by manned aircraft, this can be studied directly rather than predicted. We begin by stating the high-level objective: provide traffic conflict information in sufficient time to prevent midair collisions. The task analysis performed in Hardman (2006) defined the basic tasks of DSA as the following steps: Note: The information processing stage where this occurs is shown in parenthesis:

Scan field of regard (Sensor)

Detect traffic (Sensor)

Predict conflict (Processor)

Calculate feasible action (Processor)

Choose best action (Decision-Maker)

Execute chosen action (Actuator)

These tasks must all be performed continuously and simultaneously to satisfy the function of DSA. Current technology readiness and regulatory requirements prevent the last two steps from being performed without human involvement. Therefore, the automation analysis is done regarding the sensor and processor stages.

Quantify Performance Criteria.

Aviation regulations do not define the required level of DSA performance, but they have established see and avoid areas of responsibility (ICAO, 1962) as shown in Table 2. The DSA system must complete its detection, tracking, and predicting of an incursion with adequate time for an avoidance maneuver to be performed. For UAS operations, the time requirement must also include the relay time for any processing stages that are not performed onboard.

Table 2

Required Performance (Multiple Sources, as listed)

Parameter	Required Performance	Source of Requirement
Time to Collision Warning	No value given Sufficient for a safe miss distance (>500’). Speed dependent	FAA’s AIM, Ch 6, Sec 6 FAA-Order 8700.1
Detection Range	No value given Sufficient to achieve warning time requirement	N/A

Revisit Rate	Sufficient to achieve tracking within warning time requirement	N/A
Resolution	Sufficient to achieve tracking at required range.	N/A
Field of Regard (FOR)	+/-110° Azimuth +/-30° Elevation	(ICAO Annex 2) “Rules of the Air”
Traffic Volume	Sufficient for most crowded airspace (up to 12)	Derived from EUROCONTROL website statistics link

The required detection range (R) is a function of closure velocity (v_c) and necessary warning time (t_{c-}) as follows:

$$R = v_c \bullet t_{c\,-} \qquad (1)$$

Closure velocity is a function of the angles in azimuth (Ψ) and elevation (θ) between the two aircraft’s velocity vectors (v_1, v_2). This is defined by:

$$v_C = dR/_{dt} = ((v_1 \bullet \cos(\Psi_1) \cos(\theta_1)) + ((v_2 \bullet \cos(\Psi_2) \cos(\theta_2)) \qquad (2)$$

Above 10,000 feet there is no speed restriction except for those prohibiting supersonic flight. With only the Mach limitation, closure velocities can theoretically be over 1200 knots ground speed. However, at that altitude working transponders are required which means cooperative avoidance systems are capable of deconflicting traffic. As mentioned, aircraft are limited to 200 knots in airport areas and 250 knots elsewhere below 10,000 feet. This means aircraft below 10,000 feet in Class E, F, or G airspace must be able to prevent collisions with closure velocities up to 500 knots (two aircraft, traveling head on, both going 250 knots). Most UAVs and general aviation aircraft are speed limited but still could experience closure velocity of up to 400 knots (150 knots cruise speed, plus up to 250 knots for other traffic).

Necessary warning time includes both the time for the operator/processor to react and the aircraft to complete the avoidance maneuver. Engineers at the Spanish Aerospace Test Center (INTA) studied avoidance maneuvers for UAVs with various performance characteristics. In the vertical plane, most UAVs operate with too little excess thrust to perform a satisfactory abrupt zoom maneuver (rapid climb). This is true for most all small aircraft. A dive would yield the most rapid change in trajectory, but it is an undesirable option due to the effects of rapidly changing forces on the fuel system and payloads. Furthermore, unapproved changes in altitude while on an IFR flight plan may unsafely complicate the scenario for both controllers and operators. For these reasons, avoidance through a change in the horizontal plane is either necessary or at least preferred (Hardman, 2006).

Analysis was performed to calculate the time necessary to complete a horizontal plane maneuver. Figure 2 shows the geometry for the generalized worst-case scenario. In this scenario, a warning was given at the minimum alert time based on a predicted collision at time (t_f) if the aircraft continued on flight path (b). The time (t_f) is measured from the conclusion of the warning and reaction time. We defined the minimum allowable miss distance (r_f) to be 500 feet based on the official definition of a near mid-air collision (FAA, 2009). The aircraft's actual trajectory (s) is based on a maximum rate coordinated roll to the desired bank angle (ϕ). For illustration a right turn is used, but it is done so without loss of generality. The solution is achieved using the Law of Cosines, the derivation of which is well established.

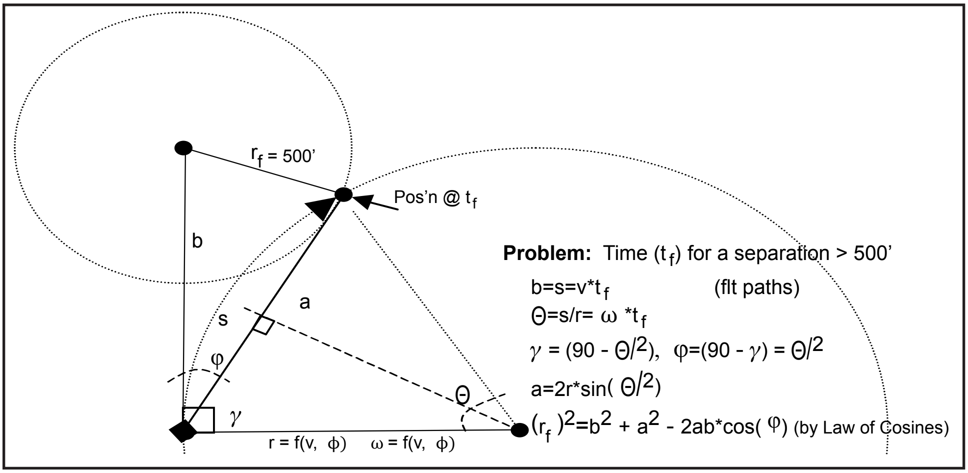

Figure 2. Avoidance Maneuver, General Solution.

The turn radius and the turn rate of an aircraft in level flight can be solved using the following established equations (USAF TPS, 2000a):

$$r = \frac{v^2}{g\sqrt{n^2-1}} \text{; and } \quad w = \frac{g\sqrt{n^2-1}}{v} \qquad (3, 4)$$

where:

v – velocity

g – force of gravity

n – load factor, which is equal to 1/cos(ϕ) in coordinated flight.

r – turn radius

ω – turn rate

Combining the equations of the generalized solution with equations 3 and 4 yields the following:

$$r_f^2 = \left(vt_f\right)^2 + \left(2r\sin\left(\frac{\omega t_f}{2}\right)\right)^2 - 2\left(2r\sin\left(\frac{\omega t_f}{2}\right)\right)\left(vt_f\right)\cos\left(\frac{\omega t_f}{2}\right) \qquad (5)$$

with all variables defined as in Figure 2 and equations 3 and 4.

The INTA study sought to deduce the proper values for the variables in equation 5. It concluded that a maximum bank angle (ϕ) of 45° should be used for several reasons. First, many UAVs and small aircraft are limited in bank angle to 60°. A maximum rate turn should not be performed to the maximum allowable bank angle due to the consequences of overshooting the bank angle limit. Secondly, the short timeline limits the amount of time available to execute the maneuver. The study found that the UAVs were capable of roll rates between 20 and 30° per second; these are common values of normal small aircraft. At these roll rates, higher bank angles would require more time than allotted for the execution of the maneuver. Thirdly, as the bank angle increases beyond this value, the viewing geometry (for pilot or sensor) becomes a factor. At some angle, dependent on aircraft type and viewing position, it will not be possible to keep the traffic in sight throughout the turn. Finally, g-force and accelerated stall speed increase inversely to the cosine of ϕ which means the rate of increase becomes very high at high angles. Based on the above study, and using equation 5 with a $\phi_{max} = 45°$, the *necessary time to complete an avoidance maneuver is* $t_f \geq 5.7$ *seconds*. If a different value for the maximum bank angle is desired, this value can be substituted in for ϕ in the equation 5.

One example of additional limitations is that satellite links frequently limit UAVs to $\phi = 15°$ during beyond LOS operations. The INTA study found that this should be programmed as a “soft stop,” but, if necessary, the DSA system should be allowed to use the maximum ϕ. The maneuver duration will be very short, so the link may not be lost. If it is lost, automated procedures for re-establishing the satellite connection are possible once the aircraft has returned to level flight.

A fact that is not intuitively obvious, but is implied by the preceding mathematical derivation, is that the warning time requirement is not speed dependent. As was shown, the required detection distance is proportional to the velocity; however, the turn rate is *inversely* proportional to velocity. In the time to collision calculations these two factors cancel out the speed dependence.

In addition to necessary warning time, aviation regulations define other performance requirements. It is necessary for a DSA system to provide coverage in the entire area of responsibility. Ideally, the system would provide +/-180° in azimuth (AZ) and +/-90 ° in elevation (EL); however, total coverage is not required. All aircraft are responsible for +/-110° in azimuth and +/-30° in elevation. Additional coverage is desired from a “defensive driving” perspective, but the DSA capability must have a field of regard (FOR) at least this big. Also, any new DSA system must be able to cope with the highest traffic densities likely to be encountered. Modern computers have no problem exceeding this number of simultaneous predictions, but the system must be able to discern the highest threat at all times.

System integrity is a measure of how well the data can be trusted. For a DSA system, it is primarily the probability of missed traffic, false alarms, or incorrect prioritization of intruders. The technical reasons for these problems include target location ambiguities, improper noise rejection, tracking ambiguity, and errors by predictive algorithms. Aviation systems have established threshold levels of safety. If a DSA system is functioning as the primary method of separation, and traffic on a conflicting flight path is non-cooperative, then a missed or incorrect detection could lead to a midair collision. The FAA requires such catastrophic events to have a probability of occurrence of less than 10^{-9} events/flight hour (FAA, 2000).

Quantification of Human Performance.

Governing agencies require UAVs to demonstrate an "equivalent level of safety" to that of manned aircraft (FAA, 2002). If an official quantitative definition of equivalent level of safety existed, then airworthiness requirements could be directly derived from that definition. Unfortunately, no such definition has been endorsed by any regulating agency. The human's capability and limitations, when acting as the direct performer for DSA, have been determined by a meta-analysis of multiple human subject studies. The most significant limitation is the required collision warning time. Research for this was originally performed by the Australian Traffic Safety Board, Civil Aviation Authority (CAA) formerly called BASI (BASI, 1991). Their results are generally accepted by the aviation community and have been cited in multiple accident investigations and subsequent research (Andre & Kukura, 2009). We summarize the results graphically in Figure 3. All parameters from the meta-analysis are listed in Table 3.

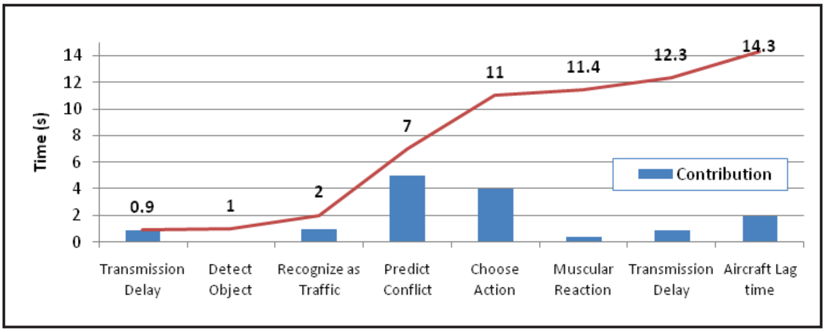

Figure 3. Time to React to a Collision Threat, Onboard Pilot

Table 3

DSA Human Ability (Multiple Sources, as listed)

Parameter	Human Performance	Source
Time to Collision Warning	Needs greater than 18.2 sec.[1]	(BASI, 1991) & calculations contained herein
Detection Range	1.14 to 1.84 NM for 90% confidence	(Andrews, 1991; Hardman, 2006)

Revisit Rate	16 sec	FAA-P-8740-51
Resolution	0.3 mrad	(Smith, 2000)
Field of Regard (FOR)	+/-180° AZ +/-30° EL	N/A
Traffic Volume	Up to 5	FAA-P-8740-51

Note: 1: Derived from 12.5 s for pilot reaction and 5.7 s for avoidance maneuver (non-fighter/aerobatic).

It is alarming to note that these results indicate that for DSA the "equivalent level of safety" standard is actually insufficient to meet the FAA's necessary level of safety against catastrophic events. Using the optimal human performance values listed in Table 3 (R=1.84 NM and necessary t_{c-} = 18.2 sec), the maximum closure velocity (v_{c_max}) safely protected by human see and avoid is v_{c_max} = 364 knots. This does not take into account the human scan rate. The FAA recommends that pilots re-scan every 16 seconds. More frequent complete area scans are a worthy goal but difficult to achieve during high workload. At this recommended rate, the necessary detection time to prevent a collision is up to t_{c-} = 34.2 sec which equates to a maximum safe closure velocity of v_{c_max} = 194 knots, well below the speeds that aircraft can legally travel!

Quantification of Machine Performance

The next step is to quantify the potential performance of automated replacement. For automated DSA systems, detection performance is a function of revisit rate, detection range, and/or resolution. This requires the use of the governing equations for optical and radar-based technologies. For radar systems, the maximum Range (R_{max}) is given by (Stimson, 1998) as:

$$R_{\max} = \sqrt[4]{\frac{P_{avg} G \sigma A_e t_{ot}}{(4\pi)^2 S_{\min}}} \quad \text{Radar Systems} \tag{6}$$

where:

P_{avg} – Average power

G – Antenna gain

σ – Target radar cross section (RCS)

A_e – Effective antenna area. (Product of the physical area and an efficiency factor)

t_{ot} – Time on target, dwell time, or integration time

S_{min} – Minimum detectable signal energy

Except for RCS and dwell time, these parameters are all limitations of the physical system. Increases in range through increases in the power, gain, or area invariably come with consequence in weight, size, power, and money.

The size and weight of the system are dependent on its necessary gain, and gain is a function of the system's design wavelength. It is given by Stimson (1998) as:

$$G = \frac{(4\pi)A_e}{\lambda^2} \qquad \text{Radar Systems} \tag{7}$$

where:

λ – Wavelength

The wavelength is equal to the speed of light (c) divided by frequency (f) ($\lambda = c/f$). Thus, the aperture size decreases or the gain increases proportionally to the square of the frequency. This makes radar systems operating at higher frequencies attractive options for UAV installation where payload is limited.

To effectively use the dwell time parameter, it is important to design a good scanning technique. The optimal dwell time is a tradeoff with revisit rate and field of regard (FOR). The longer one dwells in any one part of the sky, the longer it takes to view the total area of observation.

For infrared imaging systems, the range is related to angular resolution by the following equation (USAF TPS, 2000c):

$$R = W_R/\Theta_R \qquad \text{Infrared Systems} \tag{8}$$

where:

W_R – Linear resolution (minimum resolvable distance or diameter of target)

Θ_R – Angular resolution [rad]. The inverse of the spatial cut-off frequency ($f_{s,\,co}$)

For all imaging systems, the resolution is a function of the number of picture elements (pixels). However, high-resolution imaging creates challenges for the processing system because of the large quantities of data that must be normalized and analyzed in real time.

For systems using laser technology the maximum detection range is determined by the required power (P_R) which can be determined by the laser range equation given by the following equation (USAF TPS, 2000b):

$$P_R = \frac{P_{XMTR} D^2 \rho_T}{4R^2} \eta_{ATM}^{2} \cdot \eta_{XMTR} \cdot \eta_{RCVR} \qquad \text{Laser Systems} \tag{9}$$

where:

R – Range to target

P_{XMTR} – Power in the transmission path of the laser

D – Detector aperture diameter

ρ_T – Target reflectivity

η_{ATM} – Transmissivity, atmospheric

η_{XMTR} – Transmissivity, transmission path of the laser

η_{RCVR} – Transmissivity, receiver path of the detector

Except for systems capable of simultaneous omni-directional monitoring, knowing the maximum range is not sufficient. After detection, the DSA system must track the traffic to determine if a potential for collision exists. This requires a minimum of three scans for accurate calculation (real world trajectories are arcs). The necessary range is also a function of revisit rate. It must be assumed that the traffic is just outside of maximum detection range in the previous scan. Thus, the distance the traffic can close before being detected and tracked is equal to the closure rate multiplied by the time the system takes to perform three complete scans ($3 \cdot t_r$). Substituting this into equation 1 yields an actual time to collision.

$$t_{c-} = \frac{R-(3\, t_r\, v_c)}{v_c} = \frac{R}{v_c} - 3t_r \tag{10}$$

Predicted performance can be obtained from the given equations. The information necessary to use the equations is available from manufacturer data or applicable regulations.

To meet the derived requirements, numerous possibilities exist. There are potential tradeoffs between FOR, revisit rate, dwell time, and range (or resolution in the case of electro-optics). Thus, technologies that exceed the necessary performance in one area can make tradeoffs for improvement in others. For instance, longer dwell times can increases the detection range for radar systems. For electro-optical systems, an increase in sensitivity and a longer dwell time can increase the detection range. However, these longer dwell times increase the total area scan time. Therefore, revisit rate and maximum detection range are conflicting parameters of the time to collision requirement. These necessary tradeoffs are seldom discussed in product literature, but they are essential in helping the designer find feasible alternatives. Next, we examine the current state of th art in sensor technology and automation.

Visual Imaging. Visual imaging technology consists of using some form of camera arrangement to help establish situational awareness. It is the most analogous process to that of piloted see and avoid. Visual imaging is a semi-passive system in that no onboard illumination is needed for detection. Modern camera technology holds great possibility for small, low power cameras with very good zoom capabilities. In the future, a virtual reality system could theoretically give the remote operator the same visual scan as an onboard pilot, but the bandwidth and equipment requirements would not justify its use simply for DSA. A more realistic option is using image processing and a target recognition algorithm to analyze the input for the operator. The operator remains in the loop by cueing cameras to focus on places of interest.

The primary weakness of these systems is that they are limited in the same way as the human vision. Some technologies can augment the picture for night and in haze, but performance still suffers. Another inherent difficulty with electro-optical imaging systems is that, unlike radar, they do not have the capability to directly measure range. This is a big drawback for DSA systems as this is the primary parameter for calculating traffic avoidance. One possibility is stereoptic vision with sensors on the wing tips. This would use the same principle as the human brain to discern the distance of an object by simultaneously viewing it from two different angles. Unfortunately, this method is only effective at short distances. Beyond those distances the human being uses assessments of the apparent size of an object to determine distance (Physiological Training Office, 2001). A computer could do this as well, but it must know the actual size of the object. Range rate can be determined simply by measuring the rate of change in apparent size, but requires very high resolution systems as the apparent size of an object does not change rapidly until very close. A more probable solution to the range problem is to combine the electro-optical system with one of the technologies discussed later.

Infrared. The infrared (IR) region is lower in frequency (higher in wavelength) than the visible spectrum. IR technology uses the fact all objects radiate energy at a quantity proportional to their temperature, and aircraft have a temperature contrast with the surrounding sky. Unless used with an IR illuminator, these systems are passive; they use received energy only. IR systems operate in the electro-optical area of the spectrum, and so share many of the same properties and limitations of systems in the visible range. A typical IR system requires a signal to clutter ratio greater than 19 in order to achieve a 99% probability of detection (USAF TPS, 2000d). Though IR search and track (IRST) systems have been used by the military, there are currently no IR DSA systems in use. The only known proposal is a NASA and US Navy effort to develop a supplementary IR-based DSA system with a proposed range of several miles and a FOR of +/-105° in azimuth and +/- 35° in elevation (Adams, 2001).

Laser Radar. Laser Radar technology operates in the visible and near-IR spectrum. A LASER (light amplification through stimulated emission of radiation) is a system capable of generating an intense coherent beam of light. This beam is less susceptible to the atmospheric attenuation of other electro-optical systems. When reflected this beam can be sensed by a detector which can determine the distance of the reflected object. A laser detection and ranging (LADAR) system uses this feature to make accurate range measurements at long distances. Since the range is a part of each sensed beam, the system can provide a three dimensional perspective of the reflection. Some sources also used the term light detection and ranging (LIDAR) and included the use of ultraviolet lasers as well. The advantages and limitations of LADAR systems are both related to their very precise focused beam. LADAR systems provide high resolution in range and angle, but to cover a sufficient FOR they require very fast scanning, and real-time signal processing. There is currently no DSA system in development that makes exclusive use of LADAR or LIDAR; however, its ability to augment other systems is being explored (UAVM, 2009).

Radar. Radar (radio detection and ranging) systems have been used to detect aircraft since the 1940's. Like LADAR, which was derived from radar principles, Radar is an active system that sends strong pulses of energy and analyzes the returned signal. Two locations of interest in the radar area of the spectrum are at 35 and 94 GHz. The need for large external apertures makes radar systems difficult to implement on small vehicles, and propeller-driven aircraft have a difficult time dealing with the interference issues that the propeller and engine can cause. Pusher propeller configurations allow for the installation of radar in the nose, but the size, weight, and power requirements make them currently unfeasible. Unlike the previously discussed technologies, some radar system has been developed and evaluated for DSA on UASs. Flight Safety Technologies is preparing to field a UAV version of its UNIversal Collision Obviation and reduced Near-miss (UNICORNTM) system (Flight Safety Technologies, 2009), and there are reports that both Northrop Grumman and General Atomics are working on radar-based collision avoidance systems built specifically for installation on UAVs (UAVM, 2009). According to a Sandia National Laboratories press release, they believe that they are near the creation of a synthetic aperture radar that will have an effective range of over 7 NM and weigh less than 20 pounds (Sandia National Laboratories, 2004). Currently, the most mature system is the OASys (Obstacle Awareness System) radar installed on UAVs built by Scaled Composites, LLC (Wolfe, 2004). Designers set a range objective requirement of 6 NM. Initial NASA tests found the system was capable of detection ranges between 2.5 to 6.5 nautical miles, but there were some complete misses. Based on NASA´s flight test results, listed in Table 4, this system comes close to meeting all necessary performance requirements as stated. Regarding physical characteristics, the total weight is about 55 pounds and the externally mounted antenna is 16"x16"x22" (Wolfe, 2004).

Table 4

Evaluation of OASys Radar for DSA

Parameter	System Performance	Notes
Time to Collision based on: -- Detection Range & -- Revisit Rate	2.5 –6.5 NM 150 °/sec	Generally sufficient for UAVs

Resolution	1.7 mrad (0.097°)	Sufficient to achieve tracking within warning time requirement
Tracking Accuracy	Range: <5 m	Sufficient to achieve tracking at required range including worst-case scenario ambiguities.
Field of Regard	Typical: AZ: +/- 30°, EL: +/- 11° Max: AZ: +/- 90°, EL: +25° & - 85°	Less than the ICAO requiremnet.
Other	Altitude limitation of 20k´	Problem for most UAVs

The sensor is just the first decision, but it influences the allocation of the other steps of automation. Tasks allocated for manual execution must be done remotely, but tasks allocated for automation must then be studied for the allocation of being performed onboard vs. on the ground. At this time, not fully automated onboard DSA has been fielded, even in prototype. Though such a capability would have many advantages over less independent alternatives, the numerous technical and regulatory challenges of such a system mean that it is not likely to be an option for many years.

We re-examined the decision timeline in light of a UAS scenario with a human operator acting as the manual controller. This is shown in Figure 4 and is analogous to the piloted aircraft but with added delay due to transmission. This transmission delay must be added to the total time twice, once for the alert and once for the control message. This transmission time delay consists of a propagation component and a relay processing component. Propagation time is the result of transmission range divided by the speed of light. This results in a time delay of 6.18 μsec per nautical mile. For this UAS analysis, one-way propagation time was estimated at 0.9 s, and relay processing of the message is assumed to be negligible. Attempting to satisfy the DSA function in this way results in a necessary detection range that is significantly farther than any proposed system expects to ever achieve. Thus, it is deemed infeasible.

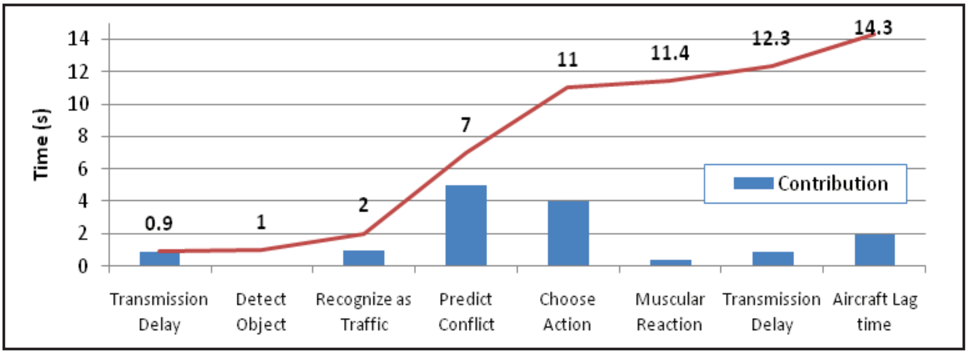

Figure 4. Time to React to a Collision Threat, Remotely Piloted

The previous conclusions guide the search for a feasible solution toward a supervisory controller option. We re-examined the decision timeline in light of a UAS scenario that uses both onboard collision prediction software and remote human interaction. Such software has been developed and tested (Chamlou, Love, & Moody, 2008), and the human-in-the-loop fulfills regulatory requirements. As shown in Figure 5, this yields a necessary warning time that is more achievable. Based on these values, the UAS requirements for DSA are listed in Table 5.

Using equation 5, we see that an automated system with a revisit rate of 1 Hz could provide the necessary alert time to an operator if it had a detection range of 2.0 NM. This is within the predicted performance of at least one sensor under development.

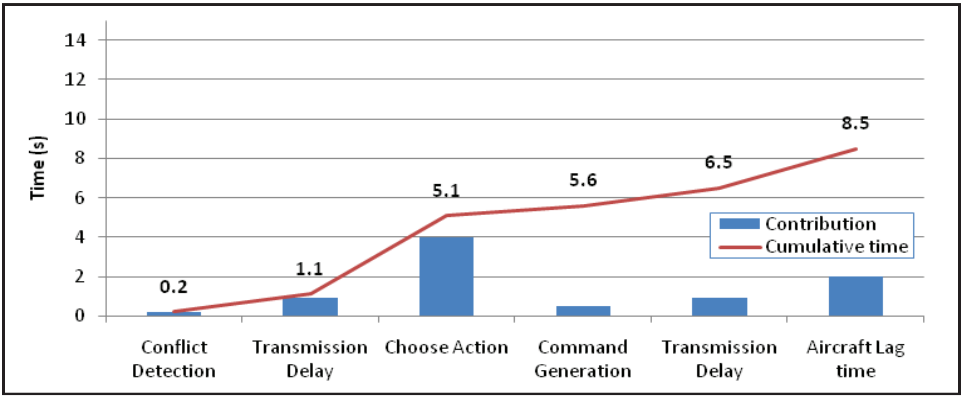

Figure 5. Time to React to a Collision Threat, Remotely Piloted

Table 5

DSA Necessary Performance of UAS Automation

Parameter	Required Performance		Source/Notes
Time to Collision Warning	Unlimited Speed limited	14.2 sec 12.2 sec	Sufficient for a safe miss distance (>500’) below 10,000 ‘. Value is for head-on traffic. Less for off angle traffic.
Detection Range & Revisit Rate	Minimum to achieve time requirements above: Unlimited Speed limited	 2.0 NM 1.4 NM	Sufficient to achieve tracking within warning time requirement.
Resolution	--		Sufficient to achieve tracking at required range including worst case scenario ambiguities.

Field of Regard	+/-110° Azimuth +/-30° Elevation	Required performance: Those dictated in the ICAO "Right of Way rules". Desired performance: Total spherical area.
Traffic Volume	> 10	Busiest airspace densities for UAV operation

Note: the speed-limited category provides for a more easily attainable requirement. Time includes transmission, reaction, maneuver, and propagation time.

Automation Selection

A comparison of the human and automatic requirements is made in Figure 6. The necessary warning times are plotted on the ordinate axis and the corresponding necessary detection ranges are plotted on the abscissa axis. The lines radiating from the origin are reference speed lines (Vc=200, 300, 400, and 500 kts respectively). Because the piloted aircraft equivalent level of safety is insufficient, UASs will ultimately have to gain certification by ICAO's alternative method, evaluation of system risks against a threshold. This method requires the advocating party to quantify the system performance and compare against an approved risk level (ICAO, 2001).

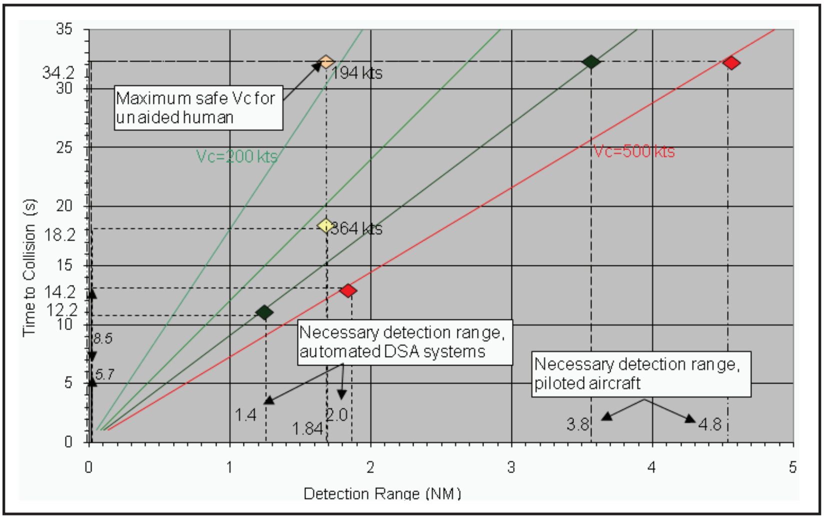

Figure 6. Human and Automated Detection Range and Time to Collision

Based on the automation analysis above, and after an examination of the available sensor technology, there are currently no solutions for small UAVs that are fully satisfactory; however, the use of a combined visual and LADAR system holds great promise for the future. Those UAVs will require some off board tracking and active deconfliction system, which means they will be limited in flexibility and dependent on another system. For larger UAVs that can handle the large expense and weight, onboard radar and conflict detection software is the best available option.

Conclusion

The method presented in this report is independent of the technology reviewed and can be used to perform quantitative function allocation between humans and computers. The results for the method, as applied to DSA in a UAS, show how automation analysis can be improved. In this particular application, the available technology makes for a difficult selection. The method made the choice clear and objective, and as these technologies mature, the analysis is easily updated to examine if the conclusions remain valid.

References

Adams, C. (2001, Feb). UAVionics in the US. *Aviation Today,* p. 20.

AF Tech. (2009). *RQ-4A/B global hawk high-altitude, long-endurance, unmanned reconnaissance aircraft*, USA. Retrieved Aug 22, 2009, from http://www.airforce-technology.com/projects/global/.

Albery, W. B., & Khomenko, M. N. (2003). Differences in pilot automation philosophies in the US and Russian air forces ground collision avoidance systems. *Technical Reports,* p. 9.

Andre, L., & Kukura, M. (2009). *See and avoid.* Retrieved 10 Sep 09, 2009, from *http://www.smartcockpit.com/flightops.*

Andrews, J. W. (1991). *Air-to-air visual acquisition handbook*. ATC-151, Lincoln Laboratory, Lexington, MA, DTIC #ADA-243807

Baker, A., & Grynbaum, M. M. (2009, Aug 10). Airspace above Hudson a highway with few signs. *The New York Times,* p. 1.

BASI. (1991). *Limitations of the see-and-avoid principle*. Sydney: Civil Aviation Authority.

Blanchard, B. S., & Fabrycky, W. J. (2006). *Systems engineering and analysis* (4th ed.). New Jersey: Pearson Prentice Hall.

Chamlou, R., Love, W. D., & Moody, C. (2008). Exploration of new algorithms for airborne collision detection and avoidance to meet NextGen capabilities. *Digital Avionics Systems Conference. DASC ‘08. IEEE/AIAA 27th,* St Paul, MN. *, 27th.* Sec. 2.d. pp. 5-1 to 5-13.

Crandal, B., & Klein, G. A. (2006). Working minds: A practitioner’s guide to cognitive task analysis. Boston: The MIT Press.

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, 48*(3). pp. 474-486.

Federal Aviation Administration (2000). *System safety handbook*. Washington, DC: Federal Aviation Administration.

Federal Aviation Administration (2002). Order 7610.4.

Federal Aviation Administration (2005). Regulations Part 91.113 and Part 91.11(a), Federal Aviation Administration (2009). Aeronautical information manual (AIM). (pp. Ch 6, Sec 6) Federal Aviation Administration.

Flight Safety Technologies. (2009). *UNICORN.* Retrieved 15 Sep, 2009, from http://www.flysafetech.com/advancing_technologies/unicorn.asp.

Folds, D. J. (1995). Using operator role theory to guide function allocation in system development. *Proceedings of the Human Factors and Ergonomics Society, 2.* pp. 1155-1159.

Hardman, N. S. (2006). *Approaches for autonomous vehicles in civil airspace: Giving sight to the blind.* (Master's Thesis, University of Tennessee).

ICAO (1962). Standard Annex 2: Rules of the Air.

ICAO (2001). Standard Annex 11: Air Traffic Services, Chapter 2; Appendix 4, .

Jean, G. V. (2009, July). Unmanned systems could fly majority of air force missions. *National Defense, 2009*, p. 34.

Miller, C. A., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors, 49*(1), pp. 57-75.

NASA. (2007). *NASA/SP-2007-6105: Systems engineering handbook* (REv. 1 ed.). Washington, D.C.: National Aeronautics and Space Administration.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *Systems, Man and Cybernetics, Part A, IEEE Transactions, 30*(3), pp. 286-297.

Physiological Training Office. (2001). *USAF physiological training handbook*. Edwards AFB, CA: Medical Group Publication.

Rouse, W. B., & Rasmussen, J. (1981). Human detection and diagnosis of system failures. *NATO Symposiumon Human Detection and Diagnosis of System Failures.*

Sandia National Laboratories. (2004). *Revolutionary radar is five times smaller than existing technology.* Retrieved 15 Sep, 2009, from http://www.sandia.gov/news-center/news-releases/2004/def-nonprolif-sec/minisar.html.

Scerbo, M., W. (1996). Theoretical perspective on adaptive automation. In R. Parasuraman, & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 37-63). New Jersey: Erlbaum Assoc., Inc.

Smith, W. J. (2000). *Modern optical engineering* (3rd ed.). Boston: McGraw-Hill Professional.

Stanton, N. A., Salmon, P. M., Walker, G. H., Baber, C., & Jenkins, D. P. (2005). *Human factors methods: A practical guide for engineering and design*. Burlington, VT: Ashgate Publishing Lmtd.

Stimson, G. W. (1998). *Introduction to airborne radar* (2nd ed., p. 136). New York: SciTech Publishing.

UAVM. (2009). *Collision avoidance: "detect sense and avoid" or "see and avoid".* Retrieved 15 Sep, 2009, from http://www.uavm.com/uavregulatory/collisionavoidance.html.

USAF TPS. (2000a). Forces of flight. *Performance phase handbook* (p. 4-1). Edwards AFB, CA: USAF Test Pilot School.

USAF TPS. (2000b). Sensors. *Systems phase handbook* (p. 4-113). Edwards AFB, CA: USAF Test Pilot School.

USAF TPS. (2000c). Sensors. *Systems phase handbook* (pp. 4-41 & 4-189). Edwards AFB, CA: USAF Test Pilot School.

USAF TPS. (2000d). Sensors. *Systems phase handbook* (p. 4-20). Edwards AFB, CA: USAF Test Pilot School.

Warwick, G. (2004, 27 Apr). Remote chance. *Flight International, 2004*, p. 1.

Wolfe, R. C. (2004). *NASA ERAST non-cooperative DSA flight test* (Final No. 2004-6). Edwards AFB, CA: National Aeronautics and Space Administration.

International Journal of Applied Aviation Studies, Volume 9, Number 2

Evaluating Pilots Using a Scenario-Based Methodology

A Guide for Instructors and Examiners

Paul A. Craig

Middle Tennessee State University
Murfreesboro, Tennessee

Abstract

From May 2008 to May 2009, over one hundred students trained at Middle Tennessee State University for the Commercial Pilot Certificate. Twenty-four of these students became participants in a study conducted during their routine Strand Checks within the syllabus. Data and observations from those checks were incorporated into an action research project. One individual acted as both Part 141 Check Instructor and research observer. The Commercial Pilot syllabus that was in use during the time of these observations utilized a scenario-based training methodology. There were four overarching research questions for the project. 1) What is the effectiveness of the scenario-based method? 2) What is the effectiveness of Learner Centered Grading? 3) Can some 'best practices' be discovered? and 4) Can recommendations be made that would guide other flight instructors and examiners as they conduct scenario-based evaluations? The project yielded several discoveries that helped answer the research questions and produced a set of "best practices" to be used by instructors and examiners when they utilize scenario-based methods in their teaching and testing.

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.

Return to Table of Contents

In early 2007, the researcher was tasked by the Federal Aviation Administration (FAA) Center of Excellence for General Aviation Research (CGAR) to write an FAA Industry Training Standards (FITS) generic commercial pilot syllabus. The syllabus was completed and accepted by FITS in the summer of 2007 and made available to the public on the FAA website. Middle Tennessee State University (MTSU) took this generic syllabus and made the minor changes necessary to submit the syllabus as an addition to the existing Training Course Outline. FAR Part 141.57 Special Curricula allows flight schools or individuals to submit alternate syllabi, and if the FAA determines that the alternate syllabus has an equal or more rigorous standard, the FAA can approve its use. In January of 2008, the Nashville FSDO approved the syllabus – making the MTSU Commercial Pilot curriculum the first in the country that was both FAA approved and FITS accepted. Students in the MTSU Professional Pilot program began using the syllabus for Commercial Pilot certification starting in the spring 2008 semester. By the fall of 2008, thirty-three students had completed the course and 88% of them had passed the Commercial Pilot Practical test on their first attempt. Students who had completed FITS training using a combined Private Certificate and Instrument rating curriculum completed the Commercial Pilot course with an average of 155.2 flight hours. Students who had no prior FITS training experience before beginning the FITS Commercial Pilot course had an average flight time when they became Commercial Pilots of 217.4 hours. (Craig, Beckman, Callender, Gossett, & Dornan, 2009).

The syllabus utilized the three tenets of FITS: scenario-based instruction, learner centered grading, and single pilot resource management. These three tenets together form a teaching strategy that increases a pilot's ability to manage risk (Summers, Ayers, Connolly, Robertson, 2007). The syllabus is also competency-based, which means it has no minimum flight times. The syllabus has three "strand" checks incorporated into the training. The strand check is an evaluation flight of the pilot-in-training by a flight instructor that meets the FAA requirements as a check instructor and who is not already the pilot's primary instructor. The three training strands that culminate in a strand check are: the Commercial VFR strand, the Commercial IFR strand, and the Commercial Complex strand. Starting with the first semester that the syllabus was in use, the researcher began administering some of the commercial pilot strand checks to the students. The researcher became an active and first-hand observer of the effectiveness of the syllabus. All the students in the study were selected for a strand check with the researcher at random from the general population of students in the Commercial Pilot syllabus between May 2008 and May 2009.

Methodology

The project was an action research process. The project's methodology was approved by the university's Institutional Review Board. Action research is a specific branch of research that involves practitioners observing the environment that they are in, gathering data, analyzing that data, and then drawing conclusions and making recommendations to improve practice. Action research is typically conducted by teachers for teachers and focuses on problems, issues, or concerns present in the practicing environment (D.V. Craig, 2009). One of the research goals is to pass on information from one teacher/flight instructor to another that would improve overall flight training. The problem or issue at hand was how to incorporate scenario-based methods into traditional flight training to improve pilot decision-making while retaining excellent ‘stick and rudder’ skills. Since improving the practice of flight instruction and simultaneously dealing with the scenario training was the goal, the action research approach was the most appropriate method.

The researcher played two roles during the strand checks, that of pilot evaluator and action researcher. The researcher became the researcher-as-instrument in a field-intensive process. A *field-intensive* process is one that requires the researcher to take an active part in the environment being studied. The researcher is expected to be a participant observer as well as the *researcher-as-instrument* involved in the research process. A participant observer is a person who takes part in all activities in the environment being studied and interacts naturally with subjects in the environment. A researcher-as-instrument is able to rely on expertise, draw on experience, and use research skills in an unbiased manner in tasks such as conducting interviews and recording notes during observations (D.V. Craig, 2009). As a check instructor, the researcher both administered strand checks and evaluated pilot performance.

The twenty-four strand checks, with twenty-four different students, were treated as individual case studies. Case study research is a qualitative research approach in which researchers focus on a unit of study known as a bounded system. The ‘bounded’ system in this project included the individual flights and post-flight briefings conducted with each student. A case study researcher collects descriptive narrative and visual data to answer “how,” “what” and “why” questions (Gay, Mills, Airasian, 2009). The case study method fit for this project because the research questions were open-ended what and why-type questions.

Research Questions

The data collection and analysis attempted to answer four overarching research questions. 1) What is the effectiveness of the scenario-based method with regard to pilot decision making? 2) What is the effectiveness of Learner Centered Grading? 3) Can some best practices with regard to administering scenario-based evaluations be discovered? and 4) Can recommendations be made that would guide other flight instructors and examiners as they conduct scenario-based evaluations?

During and immediately after each strand check, notes on the events of the flight and quotations from the students were recorded. The student’s flight instructors were interviewed and photographs were taken where appropriate of the actions of students within the scenarios that were presented with each flight. After each flight, a post-flight briefing ranging from 45 minutes to an hour and a half was conducted. During the post-flight briefing the student and the researcher each filled out separate Learner Centered Grading (LCG) sheets that were unique for that lesson. These sheets became artifacts of the study. The comments from students and instructors recorded in a field journal, the written instructor LCG sheets, the written student LCG sheets, and in-flight observations of pilot performance provided the qualitative data for the study.

Using multiple forms of data, known as triangulation, is essential to a reliable qualitative study. Triangulation is the process of using multiple methods, data collection strategies, and data sources to obtain a more complete picture of what is being studied and to cross-check information. The strength of qualitative research lies in collecting information in many ways, rather than relying solely on one (Gay, et al, 2009).

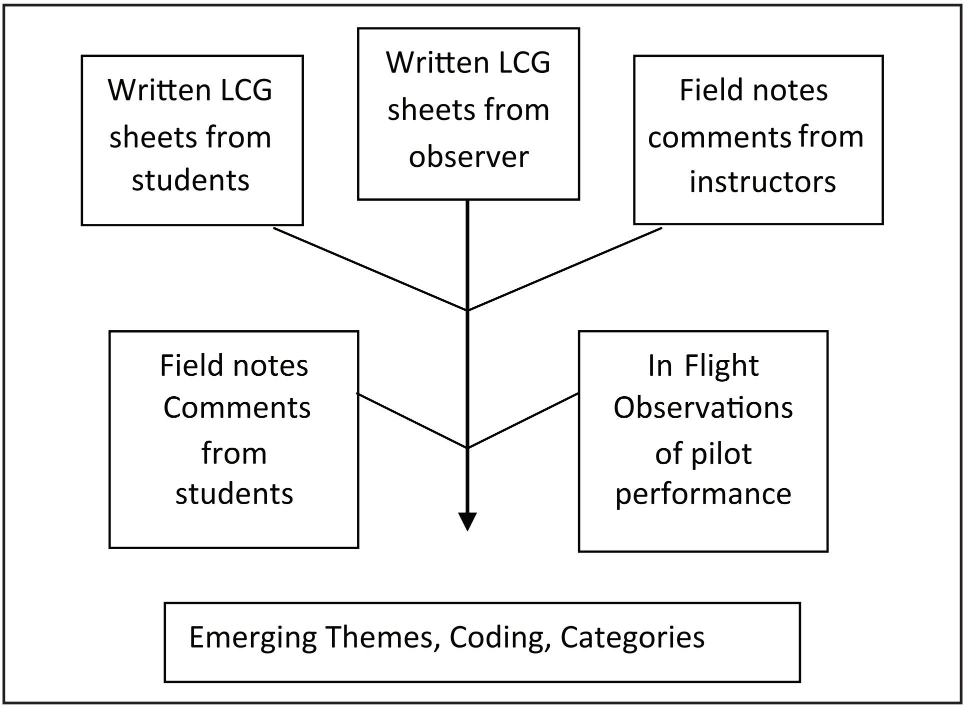

Figure 1. Multiple data sources led to the discovery of emerging themes

Coding

One of the most frequent data analysis activities undertaken by qualitative researchers is coding. This is the process of categorically marking or referencing units of text (data) with codes and labels as a way to indicate patterns and meanings. (Gay, et al, 2009). It took over a year to collect the data, but evaluation of the data was ongoing. Unlike a quantitative research method that might use parametric data, this qualitative method used narratives, field notes, quotations, training records, and LCG sheets to understand what was going on and to answer the overarching research questions. Once the data has been collected it must be organized so that some sense of it can be made. The following matrix was used to organize and categorize the data (Table 1).

The first step in organizing the data is to re-read the data and determine if any patterns or themes start to emerge. The goal of the initial step is to identify a large array of potentially important experiences, ideas, concepts, themes, etc., in the data, (Maykut & Morehouse, 1994). The first step in qualitative analysis is discovery. Within the data, the researcher looks for recurring words, phrases, and topics? Are there patterns in common between each case study? If patterns or themes are established, they are set as *coding categories*. In the categorizing and coding process the researcher seeks to develop a set of categories that provide a ‘reasonable’ reconstruction of the data that has been collected (Maykut & Morehouse, 1994).

Table 1

Overarching Questions / Data Matrix

Overarching Question	Data Set	Data Set	Data Set
1. What is the effectiveness of the scenario-based method 2. What is the effectiveness of Learner Centered Grading? 3. Can some ‘best practices’ be discovered? 4. Can recommendations be made that would guide other flight instructors and examiners as they conduct scenario-based evaluations?	Written LCG sheets from Students & Instructors	Field notes from Students and Instructors	In-Flight Observations

Discoveries and Coding Categories

Through an exhaustive review of the data in all its forms several themes did emerge in this research. Several features were commonly seen among many of the individual case study flights, interviews and LCG sheets. Following the procedure promoted by Gay, Mills and Airasian (2009) the emerging categories were labeled or named. The categories identified were: 1) The need for inside and outside the lesson scenarios, 2) Tentative Decision-Makers, 3) the skeptics, and 4) The need to use LCG sheets as a teaching tool.

Inside and Outside the Lesson Scenarios

Each case study flight employed an inside and an outside the lesson scenario strategy. A debate has been active among flight instructors in the past several years as to what is really meant by scenario-based training and why is it different than flight training that has gone on for decades. One of the findings of this study is that the difference lies in the definition of what is an inside scenario and an outside scenario. The researcher operationally defined an inside scenario as a situation that the flight instructor presents to the student while inside or during the flight lesson. An example would be a flight where a student is navigating to airport A. Along the way the flight instructor says that the weather at airport A has deteriorated and asks the student to react to this new situation. Typically the student will divert and navigate to airport B. The instructor evaluates the student’s ability to plot a new course, calculate available fuel, and execute the revised course of action to airport B. This type of scenario training has been employed in flight training for years, and good flight instructors use this method routinely. But there are elements of this situation that are not realistic. An inside scenario never states the reason why the pilot wanted to go to airport A in the first place other than it is for a training flight. An inside scenario does not consider the consequences of the diversion. If the airplane never arrives at airport A, what problems would that cause? If the student believes that the flight's purpose is for training only, then there are no negative consequences, and in fact diverting to airport B is considered a positive outcome. If an instructor uses an outside scenario then these issues are resolved and the pilot will face real-world decisions as opposed to training-world decisions. The researcher defined an outside scenario as a mission or purpose for the flight that is stated before the flight ever begins and continues after the flight is completed – the scenario takes place outside the actual time of the flight. In all 24 cases in this study, the students received an email at least 24 hours before the scheduled flight. The email described the purpose of the flight, the time constraints involved, the role that the researcher would play in the scenario and the consequences for failing to complete the mission.

> Sample email:
>
> *Hello XXXX,*
>
> *You might remember that several weeks ago a holding pond dam broke at a coal burning plant in east Tennessee. It was all in the news because when the dam broke it released some ash and sludge into the Tennessee River. This caused contamination and a health risk of those who live nearby (see attachment). Now another holding pond, located near McMinnville, Tennessee may have problems. An investigative reporter for the newspaper in McMinnville has hired you to fly over the pond so he can take photos for his news story. The story is set to run in Saturday's paper, so this must take place on Friday. We will pick up the reporter at the Warren County Airport by no later than 11:15 on Friday morning and then fly him over the lake. I will play the role of the newspaper photographer at that point. The lake is located at 35 degrees 36' by 85 degrees 52'. On this flight you will be a commercial pilot receiving compensation for the flight, but you do not work for a Part 135 company – its strictly a Part 91 and 119 flight, so review those regulations carefully. I will meet you at the airport Friday a little before 10:30, so go ahead and schedule an airplane. See you then!*

The email had all the elements to create an outside scenario. It had a specific purpose or reason to make the flight: aerial photography for a local newspaper. It had time constraints: a time that the photographer must be picked up at the airport. It defined the role of the observer/instructor: the researcher would be the photographer after we landed at the McMinnville airport. And it had consequences: failure to complete the mission would mean that the newspaper would not have the photos they needed and that would probably mean the newspaper would never call that pilot for another job. The attachment to this email had a copy of a newspaper story about the actual ash spill and a photo of a house damaged when the dam broke. A screen shot from Goggle Earth of the pond that we were going to photograph was also attached. A camera was taken on the strand check and after the student found the correct pond using a combination of GPS and pilotage, the researcher, playing the role of photographer, and took many photos of the pond from the air. The student was asked if the airplane could get lower for a better shot. The student was also asked to move the wing out of the way to get a better shot. The researcher was attempting to lure the pilot into going too low and/or abruptly yawing the airplane – the student ultimately, and wisely, overruled the researcher's request. In fact, the student was actually performing a traditional turn-around-a-point maneuver, but the maneuver was placed in a realistic context and the researcher never mentioned the maneuver by name. On the return leg of the flight, an inside scenario was introduced. The student was told that the throttle cable had come loose or had broken and he would not be able to adjust the throttle for the remainder of the flight. One recommendation that came from this research project is that flight instructors should use both an inside and outside scenario for scenario-based lessons. Instructors should introduce inside scenarios within the larger context of an outside scenario.

Tentative Decision-Makers

A second overriding theme that was observed on many flights was a tentative nature toward assuming PIC responsibilities on the part of the pilot being tested. The scenarios that were used in the strand checks all placed the pilot in the role as sole decision-maker on the flight but many were uncomfortable in that role. At the start of each flight, it was explained to the pilot that the researcher was merely going to be an observer and that they were completely in charge of the flight. It became clear that in their past training experiences, many of these students had relied heavily on their flight instructor for decision-making and for confirmations of their decisions. When the person in the role of advice-giver and instructor was removed, many of the pilots were unsure of themselves in various situations. Many of the pilots were non-assertive in their decision making. They often hesitated when a decision was at hand. Eventually many made a competent decision, but they sought confirmation that their decision was correct. They would often make statements, but end with a question in their voice. Instead of saying, "We are going to divert to Shelbyville," they would instead say "I think we should divert to Shelbyville?". (participant observation) This was an unspoken request for confirmation that the decision was acceptable to the observer. In this way they were treating the check instructor/observer like their flight instructor. There was no doubt in these cases that previous instructors had either accidentally or purposefully supported or confirmed their past decisions to the point that they had been over instructed. In some cases the pilot, facing an imminent decision, would make no decision at all, for fear of making the wrong decision in front of the observer. It was very difficult for some pilots to break out of their shell and truly act as pilot in command. On one occasion the pilot was approaching an uncontrolled airport VFR. There were multiple aircraft in the pattern and others approaching the airport. In addition, there was an aircraft conducting a practice IFR approach to the same airport. As the student got closer to the airport it became clear that to avoid the traffic he could not enter the pattern directly. The pilot in training recognized the conflict but did not immediately react. Several tense moments passed as the situation became more precarious. Finally the pilot said, "If this were me, I would swing out wide and reenter the pattern" (participant observation). To the researcher this was a very interesting comment. The pilot was actually in flight and confronted with an actual conflict. A decision needed to be made, so why didn't he already think that this was about him? The pilot knew what to do, but became unsure of himself and unable to act on his own decisions with conviction. He was more worried about what the observer would think of his decision than becoming assertive with his own decision.

For some of the pilots, scenario-based training was awkward or uncomfortable at first. The reason given by students for this was that scenario-based training was just different from what they had experienced previously. Consequently a spin-off research question arose: would continued exposure to scenario-based training reduced this awkwardness? When interviewed, the students and instructors related that students do become more comfortable with the use of scenarios when they become comfortable with what to expect. Instructors report that repeatedly placing students in decision situations with the expectation that they make decisions on their own eventually increased the pilot's confidence level with their decisions and lowered the awkwardness with the method that some initially felt.

The Skeptics

A few pilots on these strand checks did not fully respect or buy-in to the scenario as a real situation. They were skeptical of the scenario and the degree to which the scenario would be used. The research strategy was always to follow through on the scenario to its logical conclusion. But many of the pilots on these strand checks found it difficult to step out of the role of trainee and into the role of a true pilot in command. For example: The set-up for one scenario flight was that we were going to fly to another airport and drop off a company executive. The day before the flight, the pilot received an email. The email told the pilot where we would be taking the executive, told him the observer's weight, and the student was asked to calculate how much the back-seat passenger could weigh if the airplane was to remain under the maximum takeoff weight. On the day of the flight the pilot told the researcher exactly how much the passenger could weigh. But there was no actual person riding along in the back seat. The company executive was make-believe. When the flight arrived at the airport where the executive was to be dropped off, the pilot asked, "Do you want me to make a full-stop (landing)?" the researcher replied, "How will the passenger get out if we don't stop?" (participant observation) The pilot had not completely bought in to the scenario, but it wasn't all the fault of the pilot. The flight only had a make-believe passenger, so it was reasonable for the pilot to do a make-believe landing. The discovery that most pilots do not have total buy-in to the scenario pushed the researcher to improve the scenarios, attempting to make them even more real. Later, non-pilot colleague was asked to ride along in the back seat of a strand check. In that situation the passenger was not a make-believe passenger but an actual person who had a real reason to travel to another airport. That scenario had much greater buy-in on the pilot's part because there was a real person sitting in the back seat. Through interviews with students and instructors, it was learned that the lack of scenario buy-in was a product of two factors. First, many of the pilot's own flight instructors themselves learned to fly prior to the era of scenario-based training. For the most part, these were young and low-time instructors who did not have a wealth of experiences to draw scenarios from. Some instructors admitted that they began a lesson with a scenario but did not carry it through the entire flight. Sometimes they would revert back to how they were trained and do some maneuvers without any scenario context just to fill in the remainder of the time allotted for the flight. Second, some of the instructors did not have a high scenario buy-in themselves and this had biased their students. The instructors would set up a scenario for a flight, but at some point during the flight, the scenario would conclude and the lesson was continued in a traditional way.

This meant that the pilot-in-training had to act as the sole decision-maker some of the time, but switched roles in mid-flight to that of a trainee at other times. This made the scenario seem false to the pilot and lowered the buy-in.

Training of instructors in the use of a scenario-based training method must be an on going effort. Most pilots and instructors gained a high scenario buy-in from the beginning and understood the strategy of using scenarios. But the quality of the scenario set-up was a factor in the level of buy-in. Greater scenario realism yielded greater buy-in. Greater buy-in yielded greater training benefits.

Learner Centered Grading as Teaching Tool

One of the tenets of the FITS method is student involvement in the learning process. Previous research has demonstrated that the feedback/debriefing portion of scenario based training is crucial, as it plays a significant role in the effect of the scenario based training episode on performance (Blickensderfer, 2007). To that end, each lesson of the syllabus has a separate LCG sheet. This sheet is a one-page list of the major elements or tasks of the lesson with a range of boxes to check. The boxes correspond to how well or how poorly the pilot did on each element. At the conclusion of each strand check, little or no feedback was given to the student immediately after the flight because the researcher did not want to influence what the student thought of their performance. Instead, the student tied down the airplane, gathered up bags, headsets, and tach cards without any discussion of the flight. Once back in the office, the student was given a copy of that lesson's LCG sheet and was asked to fill it out. The researcher made a separate evaluation using a second copy of the same LCG sheet. After completing both sheets a comparison of the evaluations was conducted. At first the LCG sheets were seen by instructors and students alike as more unwanted paperwork. The instructors were already required to fill out and sign the student's logbook, the Part 141 required forms, a daily activity report on the student and the forms for the business office. The LCG sheets were seen as a burden of paperwork and some instructors chose not to use them. A few of the pilots tested had never before seen the LCG sheet and did not know how to score it. Some instructors balked because they reported that the grading sheet took too long to fill out, but the researcher completed an LCG sheet and asked the pilot to do the same on each one of the strand checks conducted in the project. The longest time interval recorded during the study to fill out the LCG sheet was four minutes. Eventually many instructors began to realize that, used properly, the LCG sheets were not additional paperwork but a post-flight debriefing tool. All the instructors interviewed said that they always conducted post flight evaluations with each student. Once instructors began viewing the LCG sheets as a debriefing tool, resistance to its use diminished. When the student's and instructor's LCG sheet are compared, it is clear at a glance where differences of opinion are located. Differences in grading become instant discussion points. More often than not, when there was a difference between what the student thought and what the instructor thought, the student had graded themselves more strictly than the instructor. These instant discussion points help target the instructor's post-flight debriefing to where the discussion was needed most and made students more apart of the learning process. Instead of being told they did particular items incorrectly by their instructor, the LCG sheets helped the student take some responsibility for their own progress. One of the study's research questions was, what is the effectiveness of LCG sheets? It was discovered that when instructors use the LCG sheets as a method to enhance post-flight debriefings, rather than viewing it as unnecessary paperwork, students do become more involved and take more ownership of their own training. Used as intended, the LCG sheets are an effective learning tool.

Findings and Recommendations

The purpose of the action research process is to ultimately improve practice. In this study, the practice is the teaching and evaluation of pilots using a scenario-based methodology. To that end, two of the research questions targeted improvements in this area. Can some best practices with regard to administering scenario-based evaluations be discovered? And can recommendations be made that would guide other flight instructors and examiners as they conduct scenario-based evaluations? After an evaluation using data collection triangulation, and after an analysis of that data using coding of themes, the research concludes that a set of 'best practices' can be established and several recommendations can be made.

It became clear that the more realistic the scenario was the more challenged the pilot became. Real-world scenarios require higher-level thinking and problem solving skills. Consequently, the more realistic the scenario becomes the greater level of buy-in the pilot will have. If a pilot, who undertakes scenario-based flight training, possesses a high level of scenario buy-in, they will gain more confidence in their decision-making because of exposure to the scenario. Therefore, the success of this teaching strategy depends on how real the scenarios are made to seem. When an element of make-believe enters the pilot's mind, they are less likely to buy-in and therefore less likely to gain a benefit in the form of higher-level thinking and decision making skills. The primary benefit of scenario-based training is to teach the pilot to become a safe, consistent and assertive decision maker.

During the stand checks that were administered, many different outside scenarios were used that were each designed to increase the realism of the scenario. The following is a list of a few of the scenario set-ups used in this study:

- Flying to an NCAA basketball game.
- University recruiter travels to speak to area guidance counselors.
- Take a reporter from the student newspaper to an interview.
- Charter flight to drop off a company executive at a store's grand opening.
- Pick up a Compact Disk and deliver it to Music Row.
- Fly the head football coach on a recruiting trip.
- Guest speaking at a science fair.
- Fly to a rock music festival.
- Deliver a legal deposition to the courthouse in a neighboring county.
- Pilot flies to speak at the Career Day at a high school across the state.
- Pick up a student at a distant airport and fly her back for a summer camp.
- Deliver a construction company's proposal to a bid opening.

- Pick up aircraft parts at another airport and bring them back to be installed.
- Aid tornado victims by flying in parts for the electric company.
- Environmental aerial photography mission.
- Retrieve a computer that was repaired after its hard drive crashed.
- American Red Cross blood drive delivery.

These evaluation flights produced qualitative data in several forms. Observations of flight performance/decision-making by the researcher were recorded on the instructor's Learner Centered Grading sheet. Student comments were recorded in the form of field notes by the researcher. Instructor comments were recorded in the form of field notes by the researcher and LCG sheets recorded by the student. Based on the data collected in these ways the researcher can recommend the following as 'best practices' for the teaching and evaluating of pilots using the scenario-based methodology.

1. Use props.

Realism is increased when tangible objects are used with the scenario whenever practical. One example scenario set-up was a mission to pick up aircraft parts from an FBO at airport that was 35 miles away. Before meeting with the pilot for the flight, the researcher stopped at the school's maintenance facility and asked the Director of Maintenance if he had any shipments that day that he didn't need for several hours. He provided a box that had the shipping label still on it and actual aircraft parts inside. The box was hidden from the student by placing it inside a back pack. The back pack was carried in the airplane during the strand check. When the pilot arrived at the destination airport, the box was pulled from the back pack. The pilot then had an actual package to deliver. Upon the return home the pilot carried the package to the maintenance staff and made the delivery.

Other props that were used include a Compact Disk, a camera, an actual deposition in a legal case, newspaper articles, an igloo cooler marked Blood Donation, and photos. Sometimes the props are actual people, as in the case of a university recruiter who rode in the back seat of the airplane and delivered promotional material to a guidance counselor in another city. There were times when the researcher was the prop. In those cases the pilot-in-training was told that the researcher would play a particular role and to treat the researcher just as they would an actual passenger. Throughout the project the researcher played the role of surgeon, attorney, real estate developer, newspaper photographer, store owner, contractor, and football coach. On one occasion, the pilot taking the strand check was himself used as a prop. The scenario set-up called for the pilot to fly back to his hometown so he could speak at a high school science fair. After arriving at the destination airport, the researcher went into the FBO and asked everyone who happened to be in the lobby to come outside and hear the pilot give a talk regarding how airplanes fly. Soon the pilot was using the airplane as the teaching tool for about half a dozen impromptu science fair students. Flight instructors are already very creative people, so the researcher recommends that innovative props be used in scenario training.

2. Use current events

The realism is increased if the scenarios are not repeated time and time again. To assist in keeping the scenarios new and real, build the scenarios around local current events. This will provide a steady stream of new ideas. During the year of the project many local events took place that were easily adapted into scenarios for pilot training. In April, a tornado touched down just north of the airport and many homes were damaged or destroyed. Electric power was cut off for much of the county for several days. A scenario flight a few days after the tornado was designed to fly to a neighboring country to pick up electrical components to assist in restoring power to the area. A box with the city's electric company logo on it became the package to be delivered. The MTSU women's basketball team earned a berth in the NCAA tournament and that event was used to set up a scenario. In that case the pilot was asked to fly the radio broadcasters to the game. Each summer there is a major rock music festival in the state called Bonnaroo. A scenario was created to fly a reporter to a pre-concert news conference. The researcher posed as a reporter for *Rolling Stone* magazine complete with a Bonnaroo ID badge. On the actual flight the pilot had to follow a business jet to the runway. Once on the ground the student parked next to the jet and coincidently saw the rock band AC/DC get off the jet on their way to perform. Playing the role to the fullest extent, the researcher went over and asked the band members some questions as an actual reporter. This added a level of true realism to the scenario that could not have been anticipated, but sometimes scenarios work out to be extremely real.

The syllabus used by the university does have some built-in scenarios, but it was discovered that the ready-made scenarios are enhanced by using current events. An example of this is the final end-of-course complex airplane strand check. That lesson has a built-in scenario that is based around a job interview for the pilot. In the scenario the pilot who is taking the end-of-course test must fly with a company check airman in order to get a charter pilot job. The researcher posed as the check airman for the flight, but added other elements to this ready-made scenario. The additional elements added indicated that economic times are tough for the company, so the check airman decided to combine the job interview flight with a charter flight. The charter flight customer was a college student from the Journalism Department who needed to do a video interview for YouTube. Taking the customer somewhere is not only a scenario opportunity but also an opportunity for the instructor to discuss Part 135 operations.

3. Set Time Limits and Deadlines

In every scenario, real-world time limitations and deadlines should be set. Time pressures have a great impact on the quality of pilot decision making. Unfortunately, in much of flight training, pilots are not exposed to time pressures and therefore do not react well when they experience a time crunch. An advantage of scenario training is that instructors can place students under pressure and this helps the student prepare for the day when they actually face this pressure. A tight schedule was included with every scenario set-up. If the scenario called for the pilot to deliver a passenger to a certain location, then the passenger also had a strict deadline time by which they needed to arrive. Basketball games start at a certain time, contracts must be delivered before a specific bid opening time, blood donations must arrive in a timely manner, reporters have publication deadlines to meet – all these are examples of time sensitive scenarios. In these cases, if the flight cannot be completed on an exact schedule, then there is no need to take the flight at all. Putting the student in these time-crucial situations increases the realism and helps them transition from trainee pilot to professional pilot. On one strand check an actual passenger came along on the flight who, in fact, used the flight to make a delivery. He arranged for someone to come out to the destination airport on their lunch hour to pick up an envelope. This was an IFR strand check so before takeoff the pilot was told that even though the weather was actually clear he should assume a ceiling in the area of 1,000 feet. Anytime the airplane was above 1,000 AGL, he would use a view-limiting device and anytime below 1,000 AGL he would take off the device. The pilot elected not to pick up the IFR clearance on the ground and consequently had problems making contact with ATC in the air below 1,000 AGL. The pilot actually had to fly the wrong way in order to achieve radio reception and this put the flight well off schedule. The non-pilot back seat passenger (without any prompting) asked the pilot several times if the flight would be on time. This placed additional, real-world pressure on the pilot that did affect his performance. One scenario in the syllabus simulates an on-demand charter flight where a customer calls in without any prior arrangement and needs to be taken immediately to another airport. The set-up could include a situation where the customer has a relative that has been in an accident. They must race to get to the relative as soon as is humanly possible. The customer calls to say they need an immediate flight, tells the destination, and is now driving to the airport to depart. In this situation the pilot has only a short time to safely plan the flight and prepare the airplane. The instructor tells the student that the customer will be arriving in 30 minutes and they must be ready to depart immediately when the customer arrives. The instructor must be prepared to cancel the flight if the student can not be prepared in time. Canceling a flight because the pilot could not prepare quickly yet safely really drives home the realism. This type of imposed time pressures will increase the realism of the scenario and will challenge the student. Time pressures also will help the student see the difference between being a pilot-in-training and being a commercial pilot on-the-job.

4. Employ Consequences

The true difference between the traditional use of scenarios in flight training and those used in this methodology are the implied consequences. In every scenario, there must be consequences for failing to complete the mission. In the typical inside the lesson diversion scenario, an instructor tells the student that the weather at their destination has deteriorated. The student understands this to be just a training exercise and they elect to divert to another airport or turn around. When the student does this there are no consequences associated with not arriving at the original destination. But what if the reason the pilot was flying to the original destination in the first place was to deliver a human kidney for transplant? If diverted, the kidney would not make it to the patient – facing the potential of the patient dying would be a huge consequence and in the real-world the decision to divert or turn around would not be so easy. Many accidents have taken place because a pilot pressed ahead for fear of the consequences of failing the mission of the flight. This has been called get-there-itis or get-home-itis by the FAA. In every scenario there must be an implied consequence. If a newspaper photographer cannot make his deadline because the pilot didn’t get the job done on time, then that newspaper will never call that charter company again. Pilots who lose business for their company are quickly out of a job. Holding the consequences over the pilot’s head will increase the tension and this increases the realism of the scenario.

5. Play the scenario all the way through.

 To the extent possible, let the student fly through the scenario to its logical conclusion. Sometimes things do not work out as planned. When this happens the tendency is for the flight instructor to offer suggestions that will change the course of the lesson. Resist this temptation. Allow the pilot to live with the results of their own decision making – especially when the decision making has been faulty. A greater lesson is learned by the pilot when he or she can physically see the repercussions of their own actions. Of course, doing this sometimes feels like working without a net and can create airplane scheduling problems. At a busy flight school flight lessons don't normally end when the learning has ended, they end when the airplane is due back for another flight. But scenarios don't have a fixed ending. However, with practice and when using a well thought out scenario the researcher discovered that it is possible to anticipate the length of the lesson very well.

On one occasion a student stopped the researcher in the flight school building and asked if he might be available to conduct a strand check with him. The student had done poorly on his initial IFR strand check with another check instructor and needed to redo the evaluation flight. The researcher told the student that he would be available the next day and asked him to schedule an airplane for a two-hour block. The student said, "we won't need that much time, because all we need to do is enter a holding pattern" (participant observation). The researcher told the student that scenarios really do not work that way and that when they flew, an entire scenario set-up would be used. The student did not like that answer. The researcher/instructor believed that without other distractions that the student could fly into a holding pattern properly. But could the student perform up to standards amidst additional complicating circumstances (scenario)? The student's failure to pass the original strand check was not because he didn't enter the holding pattern properly – that was just a symptom of the larger problem. Scenario training goes farther than traditional training in this regard. Scenarios can expose deeper student problems and point instructors to focus their training where it is needed most.

Final Research Question

The most important of the research questions was: What is the effectiveness of the scenario-based method with regard to pilot decision making? The data collected throughout this study revealed that challenging pilots to apply textbook knowledge to real-world situations is not something pilots do with ease. Logic would suggest that anytime one practices something, they get better at it. Likewise, when pilots routinely are placed in situations that demand decisions and made to independently work through the problems, their decision making improves. The pilots on the strand checks tended to agree. When asked to respond to the statement: "After the FITS Commercial Pilot course, I am now more confident in my aeronautical decision making skills," 95% of the students recorded that they either "agree" or "strongly agree." (Craig, et al 2009). To the open ended question: What did you like best about the scenario-based method? representative responses were:

- "I enjoyed the scenario based lessons, as they helped me to realize what I could do with my commercial license, as well as the responsibilities I would assume as a commercial pilot."

- “The ‘missions’ made FITS a little more down to earth and gave a good sense of things that commercial pilots may actually do.”
- “I liked how the scenarios were based on realistic situations and simulated real world pressure.”
- ”The real-life scenarios which were presented throughout each lesson allowed the application of regulations and piloting skills which were being taught, giving me a better understanding of the material.” (Craig et al 2009).

Flight instructors also commented on various aspects of the scenario-based syllabus. They believed that the flexibility of the syllabus saved their students time and money. The instructors said they enjoyed teaching with scenarios but needed support and guidance to utilize the method. Several instructors said it was just more fun than traditional training. Many said that in their opinion scenario-based instruction was even better suited for Commercial Pilot training over primary training because the students were applying what had already been taught (VFR flying, IFR flying, ATC communications, etc)

In conclusion, the combined qualitative data generated from this research (field notes from students and instructors, written instructor LCG sheets, written student LCG sheets, and in flight observations of pilot performance) tends to support the statement that scenario-based instruction, using real-world situations, does increase a pilot’s critical thinking skills and makes them more comfortable and assertive in decision making circumstances. The evidence indicates that the more realistic and believable the scenario is the greater the training benefit.

References

Blickensderfer, E. (2007). *Examining the Learner Centered Grading Approach*. Retrieved, from the World Wide Web: http://www.faa.gov/training_testing/training/fits/research/media/lcg.pdf

Craig, D.V. (2009). *Action Research Essentials.* San Francisco, CA: Jossey-Bass.

Craig, P., Beckman, W., Callender, M., Gossett, S., Dornan, W. (2009). *Results from the first FAA Industry Training Standards (FITS) Commercial Pilot Training Course – A Student’s Perspective*. In Proceedings, 15th International Symposium on Aviation Psychology. Dayton, Ohio.

Federal Aviation Administration. (2007, July). *FITS Generic Commercial Airplane Single Engine Land Syllabus*. Retrieved, from the World Wide Web: http://www.faa.gov/education_research/training/fits/training/generic/media/commercial_asel.pdf

Gay, L., Mills, G., Airasian, P. (2009). *Educational Research, Competencies for Analysis and Applications, Ninth Edition.* Upper Saddle River, NJ: Pearson Merrill.

Maykut, P. & Morehouse, R. (1994). *Beginning qualitative research: A philosophic and practical guide*. Washington, DC: Falmer Press.

Summers, M., Ayers, F., Connolly, T., Robertson, C., (2007) *Managing Risk through Scenario Based Training, Single Pilot resource Management, and Learner Centered Grading.* Retrieved, from the World Wide Web: http://www.faa.gov/training_testing/training/fits/guidance/media/RM_through_SBT.pdf

International Journal of Applied Aviation Studies, Volume 9, Number 2

Pilot Perspective on the Microsoft Flight Simulator for Instrument Training and Proficiency

Wendy S. Beckman

Department of Aerospace
Middle Tennessee State University

Abstract

The debate over the effectiveness of computer-based software for instrument training and proficiency has been ongoing since such software first became available. While many studies on the efficacy of such devices have been and continue to be conducted, pilots are in large number utilizing such packages. A nationwide survey was conducted to determine how the Microsoft Flight Simulator (MSFS) software package is being used by pilots for both initial instrument training and for maintaining instrument proficiency. Over 1,300 survey respondents indicated that the skills of instrument approach procedures, holding patterns, basic attitude instrument flight, and enroute navigation are frequently practiced on MSFS and are found to be effective for both initial training and for maintaining proficiency. In addition, over 85% of responding pilots indicated that they use MSFS to preview approaches at unfamiliar airports, and 88% of these pilots find the software package effective for this task. These findings indicate that pilots have embraced MSFS as a useful training aid.

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.

Return to Table of Contents

Pilot Perspective on the Microsoft Flight Simulator for Instrument Training and Proficiency

The Microsoft Flight Simulator (MSFS) software series was first made available in 1980, and over the past 28 years there have been ten editions released (Grupping, 2007). In the early days of the software, both the graphics and processing capabilities of computers and the level of sophistication of the software resulted in the program not being able to portray flight in a very realistic manner. This caused certificated pilots to view the software as solely a game, an entertaining and fun diversion, but not something that could be used for training or proficiency purposes. However, in the last decade, both the software and the capabilities of relatively inexpensive computers have evolved to the point of being able to provide a fairly realistic flight experience. This improvement has led to the use of the MSFS package by pilots both for training and proficiency purposes, even though the Federal Aviation Administration (FAA) does not allow the time spent using the package to be logged to meet instrument experience requirements. In fact, the software is so popular among pilots, two books have been published detailing strategies and methods for using the MSFS package for training (Van West & Lane-Cummings, 2007; Williams, 2006).

History of Positive Transfer in PCATD/BATD

Since the inception of computer-based training devices, a number of studies have been conducted to determine their effectiveness for both instrument proficiency and initial instrument training. The FAA, in 1993, termed these devices Personal Computer Aviation Training Devices (PCATD) and commissioned studies on their efficacy. In preparing a 1999 report to Congress (Federal Aviation Administration, 1999) the FAA, "…reviewed about 700 studies and articles; analyzed and summarized the most relevant data-based literature; interviewed government, academic, and private sector flight instruction experts on the use of the devices" (p. 2). The overall finding was that the transfer of training from PCATD to aircraft is largely positive, with some areas of training having higher levels of effectiveness than others. Since this 1999 report, research has continued into the effectiveness of PCATD for various types of pilot training: for example, Dennis and Harris (1998), Taylor, et al. (1999), Johnson and Stewart (2005), and Roessingh (2005). These studies have shown that PCATD are effective for training tasks that are primarily instrument-based, while they are not as effective for visually-based maneuvers and motor skill training applications. Given the documented effectiveness of PCATD for instrument training, the use of up to 10 hours of initial instrument training in a PCATD has been allowed by the FAA since 1997 (Federal Aviation Administration, 1997).

With regards to proficiency, Taylor, et.al. (2003), found that, 1) PCATD were effective in maintaining instrument recency of experience, as required by Federal Aviation Regulations, 2) Practice in a PCATD was at least as effective for recency of experience as practicing in an actual aircraft, and 3) PCATD were as effective as flight training devices (FTD) for recency of experience purposes. In fact, this paper recommended that the FAA begin permitting the use of PCATD to fulfill recency of experience requirements for instrument pilots. In a secondary study, Taylor et.al. found that there was no difference in subsequent aircraft performance by pilots who were given an instrument proficiency check in a PCATD, a FTD, or an aircraft (Taylor, et al, 2004).

In July 2008, Advisory Circular 61-136, “FAA Approval of Basic Aviation Training Devices (BATD) and Advanced Aviation Training Devices (AATD)” replaced Advisory Circular 61-126, which had provided guidance on the use of PCATD since 1997. The new Advisory Circular classifies devices which were previously known as PCATD or FTD into either the BATD or AATD category. Most former PCATD are now classified as BATD, while most former FTD are classified as AATD. The FAA has approved the use of BATD for up to ten hours of initial instrument training and for use in meeting the recency of experience requirements of 14 CFR 61.57(c)(1) (Federal Aviation Administration, 2008). Much like the former PCATD requirements, for a training device to qualify as a BATD, physical controls for the following items are required: landing gear, wing flaps, cowl flaps, carburetor heat control, mixture, propeller, and throttle controls. In addition, the following controls must be able to be set without using a keyboard or mouse: master/battery, magnetos, alternators, fuel boost pumps, avionics master, pitot heat, and aircraft lights.

MSFS is NOT a BATD

Given these restrictions, a typical MSFS computer station will not meet the FAA requirements for approval as a BATD. This means that MSFS cannot be used for fulfillment of instrument experience requirements when training for an instrument rating or for meeting instrument recency of experience requirements. However, the fact that time cannot be “logged” for FAA purposes does not mean that there is not value to utilizing the package. Anecdotal evidence suggests that pilots have seen the value in practicing instrument skills using MSFS and that a number of both instrument students and instrument rated pilots utilize MSFS for their own purposes, regardless of not being able to log the time. An FAA approved BATD typically costs several thousand dollars. MSFS, along with a set of flight controls, can be purchased for less than $200. The purpose of this study was to discover how the MSFS package is currently used by pilots both as they are training for their instrument rating and for maintaining proficiency once they have obtained their instrument rating. An assessment of their perceptions of the effectiveness of the package for various skill applications was also explored.

Methodology

Middle Tennessee State University (MTSU) Institutional Review Board approval was received to conduct this human subject research study. A 12 question survey instrument was developed in electronic format, and was reviewed for clarity and content by two members of the Graduate Faculty in the MTSU Aerospace Department. The survey began with two demographic questions: What year the respondent received their instrument rating, and under what part of the Code of Federal Regulations the respondent primarily flies. The first part of the survey was directed at determining whether or not the participant used MSFS during their initial instrument training. If a respondent indicated using MSFS for instrument training, they were asked to indicate which of the following areas the package was used for: Procedures and checklists, basic attitude instrument flight, enroute navigation, instrument approach procedures, holding patterns and holding pattern entries, instrument equipment failures, radio procedures/phraseology, avionics set up and usage, and aeronautical decision making. Participants were then asked to rate the effectiveness of MSFS for each of the skills they indicated practicing with the software package, using a standard Likert scale with the possible responses: Very Effective (5), Effective (4), Neutral (3), Not Effective (2), and Very Ineffective (1). The last question of this section asked the participant to estimate the total number of hours they spent using MSFS while working on their instrument rating.

The second section of the survey was designed to see how MSFS was utilized once pilots obtained their instrument rating. Thus, the first question asked was whether or not the pilot has used the MSFS to maintain or improve their instrument proficiency since achieving their instrument rating. If this question was answered in the affirmative, the participant was asked to indicate which of the following skills they have practiced: procedures and checklists, basic attitude instrument flight, enroute navigation, instrument approach procedures, holding patterns and holding pattern entries, instrument and equipment failures, radio procedures, maintaining avionics proficiency, aeronautical decision making, and previewing approaches at unfamiliar airports. Respondents were asked to rate the effectiveness of each of the skills they practice, using the same Likert scale as indicated above. Participants were then asked to estimate the number of hours per month they have spent using MSFS to practice their instrument flying skills and whether they typically have practiced a specific skill or fly a scenario when using MSFS. Finally, participants were provided an open response area in which they could provide any additional comments about their use of the MSFS package.

The survey was distributed to potential participants by the daily electronic newsletter AvWeb, which has a subscriber list of over 200,000 pilots and other aviation professionals. AvWeb published a paragraph describing the study, along with an internet link to the survey, in a January 2009 electronic daily newsletter. All instrument-rated pilots were invited to participate in completing the survey. Over 1,300 responses were received within one week of the survey announcement, and the survey was closed to respondents at that point.

Results and Discussion

Respondents were asked to identify whether they flew primarily under Code of Federal Regulations (CFR) Part 91, 135, or 121. Of the respondents, 87% indicated flying primarily under CFR Part 91. The mean year of obtaining an instrument rating was 1986 with a standard deviation of 117.22, and 48% of all respondents indicated using MSFS during their initial instrument training. Since the software package was initially released in 1980 and has evolved greatly since that time, the use of the software package by pilots who achieved their instrument rating during various time periods was examined. As can be seen in Table 1, the use of MSFS by pilots training for their instrument rating has steadily increased over time. While only 18% of pilots who earned their instrument rating in the 1981-1985 time period reported using the package for initial instrument training, for those pilots that earned their instrument rating since 2005, 82% reported using the software during training.

Table 1

Percentage of pilots using MSFS during initial instrument training, based on date of obtaining instrument rating

Earned instrument rating between:	Percentage that used MSFSX during training	Number (n) that earned their instrument rating in this time period
1981-1985	18%	49
1986-1990	17%	117
1991-1995	37%	125
1996-2000	61%	175
2001-2005	77%	251
2006-2009	82%	250

For those pilots using the software during initial training, a mean of 51.6 hours (standard deviation of 45.41) of time was spent using the package. Respondents indicated which skills they practiced and how effective MSFS was perceived to be for those skill areas as seen in Table 2.

Table 2

Percentage of pilots that reported practicing specific skills on MSFS while working on their instrument rating and the effectiveness of MSFS for practicing those skills

Skill	Percent that practiced this skill	Skill effectiveness rating for MSFS by those that practiced this skill (5=very effective, 1=very ineffective)		
		Mean	*Std Dev*	*n*
Procedures and Checklists	43.7%	3.45	.999	488
Basic Attitude Instrument Flight	80.9%	4.12	.831	595
Enroute Navigation	79.5%	4.36	.779	591
Instrument Approach Procedures	93.2%	4.55	.698	646
Holding Patterns and Hold Entries	81.3%	4.42	.794	607
Instrument/Equipment Failures	49.7%	3.67	.995	509
Radio Procedures/Phraseology	23.9%	2.78	1.08	459
Avionics Set Up and Usage	69.0%	3.70	1.06	563
Aeronautical Decision Making	35.6%	3.17	1.06	471

It can be seen that the areas of instrument approach procedures, holding patterns, basic attitude instrument flight, and enroute navigation were both practiced by the majority of users and received a mean effectiveness score of over 4.0. Avionics set up and usage were also ranked as both being commonly practiced and fairly effective, while instrument/equipment failures were practiced by about half of the respondents and were rated as fairly effective. However, areas such as procedures and checklists, aeronautical decision making, and radio procedures were not often practiced and were not rated as being very effective in the MSFS package. It was worth noting that neither the percentage that practiced a particular skill nor the rated effectiveness of a particular skill changed appreciably over time. For example, pilots who received their instrument ratings between 1981 and 1985 reported practicing the same skills on MSFS as those that received their instrument rating between 2006-2009.

Next, the use of MSFS by instrument-rated pilots was examined. Over all survey respondents, 69% indicated having used MSFS to maintain or improve their proficiency since obtaining their instrument rating. When the use of the package for proficiency was examined by date of achieving their instrument rating, the results were as indicated in Table 3.

Table 3

Percentage of pilots using MSFS for continued proficiency, based on date of obtaining instrument rating

Earned instrument rating between:	*Percentage that now use MSFSX for proficiency*	*Number (n) that earned their instrument rating in this time period*
Prior to 1980	64%	254
1981-1985	70%	49
1986-1990	64%	117
1991-1995	73%	125
1996-2000	74%	175
2001-2005	71%	251
2006-2009	71%	250

It is interesting to note that there is very little difference between those pilots that achieved their instrument rating prior to 1980 and those that have obtained their instrument rating since that time, with regards to using the MSFS package for proficiency purposes. This means that even pilots who were not exposed to the software package during their initial instrument training are now choosing to use the software to assist in maintaining their instrument skills. For those pilots that reported using MSFS for proficiency, the skills which they practice and how effective they regard MSFS to be for practicing those skills can be seen in Table 4.

Table 4

Percentage of pilots that reported practicing specific skills with MSFS after obtaining their instrument rating and the effectiveness of MSFS for practicing those skills

Skill	Percent that practiced this skill	Skill effectiveness rating for MSFS by those that practiced this skill (5=very effective, 1=very ineffective)		
		Mean	*Std Dev*	*n*
Procedures and Checklists	42.9%	3.57	.973	589
Basic Attitude Instrument Flight	80.2%	4.13	.847	761
Enroute Navigation	76.9%	4.29	.768	741
Instrument Approach Procedures	97.2%	4.54	.649	871
Holding Patterns and Hold Entries	83.3%	4.45	.741	770
Instrument/Equipment Failures	50.3%	3.75	.933	613
Radio Procedures/Phraseology	20.8%	2.91	1.04	522
Avionics Set Up and Usage	51.3%	3.66	1.04	613
Aeronautical Decision Making	39.7%	3.42	1.03	568
Previewing Approaches at Unfamiliar Airports	86.2%	4.40	.789	790

For pilots using MSFS to maintain instrument skills, the areas of instrument approach procedures, previewing approaches at unfamiliar airports, holding patterns, basic attitude instrument flight, and enroute navigation are the skills that were indicated as most commonly practiced. Each of these areas also received a mean effectiveness rating of over 4.0 by the respondents that reported practicing the particular skill. The areas of maintaining avionics proficiency, instrument/equipment failures, and procedures and checklists were indicated as being practiced by around half of the respondents and as being somewhat effective. Radio procedures and aeronautical decision making were not ranked as being often practiced or as being particularly effective. The respondents who use MSFS for proficiency indicated practicing a mean of 5.51 hours per month (standard deviation 6.98) using MSFS, and 61% reported practicing using scenarios versus practicing a specific skill in isolation.

Conclusion

Not surprisingly, there has been a sharp increase in the use of MSFS by pilots engaged in training for their instrument rating over the past 30 years. While only 18% of instrument trainees utilized the package in the early 1980's, 82% of those that have trained for an instrument rating in the past three years used the package during training. In addition, approximately 70% of the instrument rated pilots who responded to this survey indicated that they have used MSFS to help maintain their instrument skills, practicing a mean of 5.51 hours per month (standard deviation 6.98). The areas that are often practiced and that are seen as highly effective were the same for both those working on their instrument rating and for those that use the package to maintain proficiency. These areas include instrument approach procedures, holding patterns, basic attitude instrument flight, and enroute navigation. Those pilots holding instrument ratings also indicated that MSFS was frequently used to preview approaches at unfamiliar airports before conducting a flight.

While no statistical analysis was performed, it appears that the maneuvers that are indicated as most commonly practiced on MSFS are those that are rated as most effective. This is expected, as pilots would not continue to practice tasks in MSFS that were perceived as ineffective at either developing or maintaining instrument skills. Since there is agreement between the lists of effective areas for using MSFS in both the training and maintaining proficiency arenas, it may help pilots at the very beginning of their instrument training if their flight instructor indicated to them which tasks can be practiced most effectively in MSFS.

There are two major limitations of this study. First, as is true with many survey methodologies, the study employed a self-report mechanism which was subject to the limitations of the retrospective perceptions of the participants. This methodology is not as strong as an experimental design, where actual achievement on various tasks is able to be measured. However, the ratings by the pilot participants did tend to show agreement with the experimental research that has been published to date regarding the effectiveness of BATD's for instrument training and proficiency. The second limitation is that the participants in this study all received the link to the survey via an electronic newsletter. This being the case, it is likely that the recipients of the survey link tended to be more computer-oriented than those instrument-rated pilots that do not receive such electronic newsletters. This predisposal toward computer use may cause the MSFS usage numbers indicated in this study to be higher for the responding group of pilots than for the remainder of the pilot population, so the survey results are not necessarily generalizable to the entire instrument-rated pilot population.

This study reveals that pilots understand there is more to achieving and maintaining instrument proficiency than simply meeting the requirements of the FAA. Since MSFS is not a BATD, the FAA does not allow the "logging" of time spent using MSFS to meet the experience requirements to obtain an instrument rating, nor to meet the instrument recency of experience requirements of CFR 61.57. However, pilots continue to use MSFS as a training aid and have identified a number of skills that can be practiced effectively in MSFS, and are proceeding to practice these skills in large numbers. Whether or not the time can be logged, pilots have realized that MSFS is a relatively inexpensive and convenient way to both improve upon and maintain their instrument skills. Given the expense of aircraft flight time, the development of any skills which can be achieved on the ground should be encouraged. This will allow more time in flight to develop aeronautical decision making skills by experiencing complex tasks that will be required when operating under IFR, such as cross country flights and flights in actual IMC. Allowing for more training which is focused on operating in the IFR environment should ultimately lead to both more competent pilots and to pilots who are more comfortable and confident in their instrument skills. Therefore, the use of any training aid that assists in the development of these skills should be encouraged by both flight schools and individual instructors.

References

Code of Federal Regulations. (2009). 14 CFR Chapter I Part 61 Certification: Pilots, Flight Instructors, and Ground Instructors. Retrieved March 10, 2009, from: http://ecfr.gpoaccess.gov/

Dennis, K. and Harris, D. (1998). Computer-based Simulation as an Adjunct to Ab Initio Flight Training. *International Journal of Aviation Psychology, 8(3),* 261-276.

Federal Aviation Administration. (2008). *FAA Approval of Basic Aviation Training Devices (BATD) and Advanced Aviation Training Devices (AATD)* (AC-61-136). Washington, DC. Government Printing Office. Retrieved March 25, 2009 from: http://rgl.faa.gov/Regulatory_and_Guidance_Library/rgAdvisoryCircular.nsf/0/37E40A5E69B18FEF8625748E005B327A?OpenDocument

Federal Aviation Administration. (1999). Aviation Safety: Research Supports Limited Use of Personal Computer Aviation Training Devices for Pilots. Government Accounting Office. Retrieved March 11, 2009 from:http://www.gao.gov/archive/1999/rc99143.pdf

Federal Aviation Administration. (1997). *Qualification and Approval of Personal Computer-Based Aviation Training Devices* (AC 61-126). Washington, DC. Government Printing Office. Retrieved March 25, 2009 from: http://rgl.faa.gov/Regulatory_and_Guidance_Library/rgAdvisoryCircular.nsf/0/D 0D8CDA61E05EB90862569BA0052B8F6?OpenDocument

Grupping, J. (2007). Flight Simulator History. Retrieved February 23, 2009 from Flight Simulator History Web Site: http://fshistory.simflight.com/fsh/timeline.htm

Johnson, D. and Stewart, J. (2005). Utility of a Personal Computer-Based Aviation Training Device for Helicopter Flight Training. *International Journal of Applied Aviation Studies, 5(2),* 287-305.

Roessingh, J. (2005). Transfer of Manual Flying Skills From PC-Based Simulation to Actual Flight-Comparison of In-Flight Measured Data and Instructor Ratings. *International Journal of Aviation Psychology, 15(1),* 67-90.

Taylor, H., Lintern, G., Hulin, C., Talleur, D., Emanuel, T., and Phillips, S. (1999). Transfer of Training Effectiveness of a Personal Computer Aviation Training Device. *International Journal of Aviation Psychology, 9(4),* 319-335.

Taylor, H., Talleur, D., Bradshaw, G., Emanuel, Jr., T., Rantanen, E., Hulin, C., Lendrum, L. (2003). *Effectiveness of personal computers to meet recency of experience requirements.* (DOT/FAA/AM-03/3). Retrieved March 10, 2009 from: http://www.faa.gov/library/reports/medical/oamtechreports/2000s/media/0303.pdf

Taylor, H., Talleur, D., Rantanen, E., Emanuel, Jr., T. (2004*). The Effectiveness of a Personal Computer Aviation Training Device, a Flight Training Device, and an Airplane in Conducting Instrument Proficiency Checks Volume 1: Subjective Pilot Performance Evaluation.* (AHFD-04-12/FAA-04-5). Savoy, IL: Aviation Human Factors Division. Retrieved March 10, 2009 from: http://www.tc.faa.gov/logistics/grants/pdf/2001/01-G-037%20Volume%201.pdf

Van West, J. & Lane-Cummings, K. (2007). *Microsoft Flight Simulator X for Pilots: Real-World Training.* Indianapolis, IN: Wiley Publishing, Inc.

Williams, B. (2006). *Microsoft Flight Simulator as a Training Aid*. Newcastle, WA: Aviation Supplies and Academics.

International Journal of Applied Aviation Studies, Volume 9, Number 2
Copyright © 2009, FAA Academy, Oklahoma City, OK

Motion Sickness Prevention by Stroboscopic Environment during Simulated Military Transport

Catherine M. Webb, Arthur Estrada III,

and

Jeremy R. Athy

United States Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362
catherine.webb@us.army.mil;

Abstract

Previous studies have shown stroboscopic illumination to reduce the severity of motion sickness (MS) symptoms when retinal slip is a significant factor. The present study assessed the use of 4 and 8 hertz (Hz) stroboscopic environments as countermeasures for MS. The motion profiles of a Black Hawk helicopter and an amphibious vehicle were produced using a Multi Axis Ride Simulator (MARS). Each participant attended three experimental sessions over a five-day period in which they read passages while experiencing 20 minute sessions on the MARS. To assess MS, participants completed the Motion Sickness Questionnaire, a balance assessment, the Psychomotor Vigilance Test, as well as a subjective questionnaire evaluating the effectiveness of the stroboscopic environment. Although there was no conclusive evidence of stroboscopic illumination as a MS countermeasure in the objective performance measures, there was evidence of its effectiveness in subjective reports. Future research with a more MS-susceptible population is warranted.

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.

Return to Table of Contents

Motion Sickness Prevention by Stroboscopic Environment during Simulated Military Transport

Motion sickness (MS) has been well known for thousands of years, as ancient seafaring nations were very familiar with this malady. It has become increasingly prevalent in the modern world with the development of many forms of vehicular travel. Symptoms include dizziness, nausea, vomiting, drowsiness, pallor, sweating, and overall malaise (Benson, 2002). The most accepted theory concerning the cause of MS is the sensory conflict theory proposed by Reason and Brand (1975). This theory suggests that sickness results when the vestibular, visual, and proprioceptive senses perceive motion information that conflicts with expectations based on past experience.

Studies have shown that retinal image velocity (retinal slip) contributes to space and terrestrial MS (Han, Kumar, Somers, Reschke, & Leigh, 2005; Reschke, Somers & Ford, 2006). Retinal slip results when our eyes fail to hold an image stationary on the retina. This problem has implications for Soldiers who are visually engaged (e.g., reading and/or navigating) while being transported in the back of various military vehicles. For example, Cowings, Toscano, DeRoshia, & Tauson (1999) reported a negative impact on crew performance and health when the participants attended to computer screens while in a moving command and control vehicle. Motion sickness was reported by 100% of the subjects with 55% indicating moderate to severe symptoms. The authors also report that 15% of the participants experienced vomiting and that drowsiness was the most frequently reported symptom.

In operational environments, MS should be treated with the most effective countermeasures which yield the fewest negative side effects. Many of the currently available pharmaceutical countermeasures must be given in high doses to be effective. Unfortunately, high doses of antiemetics often produce sedation, which is unacceptable in terms of mission effectiveness. Hence, the development of non-traditional, non-pharmacologic MS and nausea remedies would be of great benefit to the operational military community.

A possible solution to the problem of MS related to retinal slip is the use of stroboscopic vision. Stroboscopic illumination is believed to prevent retinal slip by presenting snapshots of the visual surroundings that are brief enough so each image is stationary on the retina (Stroboscopic, 2005). It is believed that prevention of retinal slip will reduce symptoms of MS. This field of research began serendipitously as a result of a research project exploring adaptation of the vestibulo-ocular reflex employing optically reversing prisms which induced MS symptoms (Melvill-Jones & Mandl, 1981). Melvill-Jones and Mandl discovered what they term a “particularly interesting” finding: None of the subjects ever experienced nausea or associated symptoms in 4 hertz (Hz), or cycles per second, stroboscopic light (strobe-light conditions).

The results of a study by Reschke, Somers, and Ford (2006), comparing the efficacy of strobe lighting and shutter glasses (both at 4 Hz) as a treatment for MS, were very similar to those of Melvill-Jones and Mandl (1981). Reschke et al. reported that stroboscopic illumination, both by ambient illumination or by shutter glasses, reduced the severity of MS symptoms and “appears to be an effective countermeasure where retinal slip is a significant factor in eliciting motion sickness due to either self- or surround-motion" (p. 2). A review of these studies provides compelling evidence that stroboscopic technology may provide a method of preventing MS in the mounted Warfighter. Estrada (2007), in a preliminary, but suggestive airborne test of the stroboscopic shutter glasses in the U.S. Army Aeromedical Research Laboratory's (USAARL) research helicopter, found the results to be consistent with the reports by Reschke et al. and Han et al. (2005). Although efficacy of the shutter glasses as a countermeasure for MS was not implied by Estrada's test, the results did indicate that stroboscopic technologies, such as the shutter glasses, demonstrated promise, especially at 8 Hz, and should be explored as a non-pharmacological MS prevention strategy.

Despite the research reporting the benefits of stroboscopic vision as a countermeasure for MS, it should be noted that a small percentage of the population is adversely affected by flickering or flashing light. According to the National Society for Epilepsy (NSE) (n.d.) and the Epilepsy Foundation (n.d.), photosensitive epilepsy (sometimes called flicker-induced epilepsy) has been reported in about 3 to 5% of the people who have epilepsy (1 in 200) and is more common in children and adolescents between the ages of 5 and 19 years. The NSE lists the most common triggers of photosensitive epilepsy as visual fire alarm strobe lights, television screens, video games, computer monitors, and exposure to string environmental lights. The rarity of this condition is documented in a study by Doose and Waltz (1993) where only 2 to 10% of individuals possessing electroencephalogram (EEG) markers of seizure liability (photoparoxysmal response) developed seizures due to photic stimulation. The frequency range at which seizures are induced varies according to the information source. According to the NSE and Epilepsy Foundation, seizures are generally triggered by flashes between 5 and 30 Hz while DeHart and Davis (2002) suggest the triggering frequencies are between 8 to 14 Hz. As expected, the critical frequency varies from person to person although it is uncommon to have photosensitivity to flashes below 5 Hz (NSE).

The objective of this study was to determine the effectiveness of 4 and 8 Hz stroboscopic environments for alleviating MS symptoms and ameliorating performance declines elicited by aircraft and amphibious vehicle motion profiles. The current study explored the potential of turning the cabin area (passenger section) of military vehicles into an ambient stroboscopic environment as a MS countermeasure. It was hypothesized that symptoms of MS would be reduced under the two stroboscopic conditions (i.e., 4 and 8 Hz) compared to the no-strobe condition (i.e., normal room illumination).

Methods

Study population

Data were collected on 18 participants. Participants were active duty military and Department of Defense civilian employees. In addition, participants were screened for a history of epilepsy to reduce the risk of photosensitive epilepsy induced by the stroboscopic environment. Participants were divided into two groups based on the motion profile they were to experience. Data from one participant from the terrestrial/aquatic group were excluded from all analysis due to failure to complete all testing sessions. In addition, data from the Effectiveness Questionnaire from one participant in the airborne group were not included in the analysis as the participant chose to abstain from completing the survey. Of the nine participants that experienced the aircraft motion, five were males and four were females. The average age of these participants was 29.00 years (*SD* = 5.98 years). Of the eight participants that experienced the amphibious vehicle motion, seven were males and one was female. The average age of these participants was 29.63 years (*SD* = 4.27 years).

Equipment

The motion profiles of an Army UH-60 Black Hawk helicopter and a Marine AAVC7A1 Amphibious Assault Vehicle were produced using the USAARL's Multi Axis Ride Simulator (MARS). In general, the MARS is a Stewart style motion platform (Figure 1), in which a seat is attached. The MARS duplicates the actual vibrations, movements, thrusts, and jolts of various military vehicles. The usable frequency range of the table is 0 to 40 Hz, with acceleration peaks of 3g. All motion profiles used were within exposure limits defined by the International Organization for Standardization (ISO) standard 2631-1 (ISO, 1997).

Figure 1. The Multi Axis Ride Simulator

Motion data for the MARS were collected from an actual UH-60 Black Hawk helicopter and a Marine AAVC7A1 Amphibious Assault Vehicle using a tri-axial accelerometer. The motion profile for the Airborne group consisted of two signals repeated for a total of 20 minutes. The motion profile for the Terrestrial/Aquatic group consisted of a 10 minute simulation of movement on land (comprised of three signals) followed by a 10 minute simulation of movement on water (also comprised of three signals). Detailed frequency information is presented in Table 1. Note that the frequency data is for the vertical (Z) axis.

Table 1

Details of airborne and terrestrial/aquatic MARS motion profiles

	Duration	RMS	Peak	Major frequency components
	(sec)	(m/s2)	(m/s2)	(Hz)
AIR1	80	1.02	3.77	17.2, 21.5, 30.2
AIR2	60	1.29	4.17	5.9, 17.2, 21.5, 30.3
LAND1	45	6.47	17.68	35-40
LAND2	45	1.61	7.71	1.6, 4, 35-40
LAND3	45	2.08	10.41	1.6, 4, 28.5, 35-40
WATER1	165	1.4	8.17	1 - 7
WATER2	48	1.14	6.11	1 -7
WATER3	53	1.32	17.66	1 - 7

For the purposes of this study, the MARS was surrounded by a black curtain to prevent participants from seeing stabilizing outside visual references. A 750 watt strobe light provided the ambient stroboscopic effect (4 or 8 Hz). The strobe light measured 220 equivalent candelas at the intensity setting used for the experiment. A 90 watt bulb provided the ambient reading light for the no-strobe condition. Both light sources were mounted overhead.

In order to induce retinal slippage, all participants were asked to read a passage from a military novel and answer questions regarding the material. The passage was presented on 8.5 x 11 inch paper with 20 point Times New Roman font. The participants' heads were not supported or restrained and were generally in a forward and downward facing position while attending to the hand-held reading task. Observations and video recordings of the sessions confirmed minimal variation in head movements.

Data collection instruments

Motion History Questionnaire. Developed by Kennedy and Graybiel (1965), the Motion History Questionnaire (MHQ) was used to ask participants about their experiences in environments that may engender MS-like symptoms; judged susceptibility to motion sickness, nausea and dizziness; and likes and dislikes for activities which produce such symptoms in some persons. Participants' responses on the MHQ were used to compute a "Perceived Susceptibility" score ranging from 0 to 15, where a higher score indicates a greater susceptibility to MS (Kennedy et al., 2001).

Psychomotor Vigilance Task. In order to test for changes in alertness, basic reaction time was tested through the Psychomotor Vigilance Task (PVT). Participants were required to monitor a screen on which an LED stimulus was presented randomly every 1 to 10 seconds. The participant responded by pressing a microswitch. Reaction time (RT) and lapses (responses over 500 milliseconds [msec]) were recorded for each stimulus.

Motion Sickness Questionnaire. Subjective sickness symptoms were measured using the Motion Sickness Questionnaire (MSQ) (Kellogg, Kennedy & Graybiel, 1965). The MSQ is a self-report form consisting of 28 items that are rated by the participant in terms of severity on a 4-point scale or with yes-no answers. The questionnaire was automatically scored by computer.

Postural Balance Assessment. One symptom of motion sickness is dizziness, which can affect balance (Benson, 2002). To test for this affect, a Postural Balance Assessment (PBA) was employed. It is a 5-minute postural equilibrium test consisting of three parts (Gower & Fowkles, 1989). The first part is referred to as "walk on floor with eyes closed" (WOFEC) and requires the participant to take 12 heel-to-toe steps with her/his eyes closed and arms folded across her/his chest. The participant is scored on a scale of 0 to 12 based on how many steps she/he is able to make without side-stepping or losing balance. The second part of the PBA is the "standing on preferred leg with eyes closed" (SOPLEC) test which requires the participant to stand on her/his preferred leg for 30 seconds with her/his eyes closed and arms folded across her/his chest. The participant is scored on the number of seconds she/he is able to remain upright (to within 5 degrees) without losing balance. The third part is the "standing on non-preferred leg with eyes closed" (SONLEC) test which is the same as SOPLEC except that the participant stands on the opposite leg. The three parts of the PBA were completed three times, and the scores from all three trials were averaged for each part of the PBA.

*Effectiveness Questionnaire.*Participants' opinions regarding the effectiveness of the stroboscopic lighting conditions in reducing MS were captured with an Effectiveness Questionnaire. Due to the paucity of objective measures for motion sickness, the Effectiveness Questionnaire was developed by the research staff in order to solicit subjective, opinion-based feedback in order to more fully assess perceived efficacy of the test conditions. The responses to the first two questions were on a visual analog scale to allow analysis with parametric statistics. The questionnaire contained the following questions:

1. How effective do you feel the stroboscopic environment was in controlling motion sickness and allowing you to perform the reading task?
2. Was the stroboscopic environment distracting in any way?
3. Based on your experience in the stroboscopic environment, do you feel the strobe effect has a practical application for military helicopter passengers?
4. Provide any additional comments.

Procedure

The study protocol was approved by the USAARL's Human Use Committee. Written informed consent was obtained from all participants. In addition, participants viewed a video outlining the safety features of the MARS. Participants also completed the MHQ during the in-processing procedures.

The present study used a repeated measures design. Each participant attended three experimental sessions one day apart over a 5-day period (i.e., Monday, Wednesday, and Friday). The independent variable was frequency of the stroboscopic light (no-strobe, 4, and 8 Hz). The participants experienced one of the three lighting conditions per session, and the order of presentation was pseudo-randomized among 3 possible orders.

When participants first arrived at the testing facility, they completed the PVT, MSQ, and PBA. Next, participants experienced the 20 minute session on the MARS. Half of the participants (*n* = 9) were exposed to the Black Hawk motion (the airborne group) and the remaining participants (*n* = 8) were exposed to the Amphibious Vehicle motion (the terrestrial/aquatic group). During the MARS session, participants read selected passages from a military novel and answered questions to induce retinal slippage. After completion of the MARS session, participants again completed the PVT, MSQ, PBA, and Effectiveness Questionnaire. Before being released from the study each day, participants met with the study physician to ensure there were no lingering effects of the stroboscopic and/or motion environments.

Results

All statistical analyses were conducted using SPSS® 12.0 with significance set at $p = 0.05$. For all dependent measures, the airborne group was analyzed independently from the terrestrial/aquatic group.

Motion History Questionnaire

With regard to MS susceptibility, both groups of participants scored low on the MHQ, thus indicating a low susceptibility to MS. The mean Perceived Susceptibility score for the participants in the airborne group was 4.22 (*SD* = 2.54). For those participants in the terrestrial/aquatic group, the average score was 2.62 (*SD* = 2.20). The difference in perceived susceptibility between the two groups was not significant as revealed by an independent samples *t* test ($t(15) = 1.38$, $p = 0.19$).

Psychomotor Vigilance Task

Participants were asked to complete the PVT before and after each exposure to the MARS, and data were recorded regarding mean reaction time and number of lapses. Tables 2 and 3 present the respective data before and after the MARS exposure for the airborne and terrestrial/aquatic group by lighting condition. Difference scores were calculated by subtracting the scores from the post-administration from scores of the pre-administration. The data were analyzed using a one-way repeated measures ANOVA across the three lighting conditions (0 Hz, 4 Hz and 8 Hz). No significant *performance* differences were found among the lighting conditions for mean reaction time or lapses for the airborne group or the terrestrial/aquatic group.

Table 2

Psychomotor Vigilance Task Mean Reaction Time (msec) Data and standard error (SE)

Motion	Admin-istration	Lighting Condition					
		0 Hz		4 Hz		8 Hz	
		Mean RT	SE	Mean RT	SE	Mean RT	SE
Airborne Group	Pre	280.25	8.84	263.09	12.91	276.53	10.08
	Post	294.28	23.97	262.73	11.35	277.05	8.20
Terrestrial/ Aquatic Group	Pre	251.40	13.37	237.39	10.09	252.02	16.96
	Post	246.43	13.79	239.14	11.30	249.98	13.77

Table 3

Psychomotor Vigilance Task Mean Lapses Data and standard error (SE)

Motion	Admin-istration	Lighting Condition					
		0 Hz		4 Hz		8 Hz	
		Mean lapses	SE	Mean lapses	SE	Mean lapses	SE
Airborne Group	Pre	1.11	0.20	1.11	0.51	1.11	0.54
	Post	3.33	2.11	1.00	0.53	1.11	0.31
Terrestrial/ Aquatic Group	Pre	0.50	0.27	0.25	0.16	0.63	0.26
	Post	0.50	0.27	0.38	0.18	0.75	0.37

Motion Sickness Questionnaire

Participants were asked to complete the MSQ before and after each exposure to the MARS. Figure 2 presents the four MSQ scores before and after the MARS exposure for the airborne and terrestrial/aquatic group by lighting condition. Difference scores were calculated by subtracting the scores from the pre-administration from scores of the post-administration. The data were analyzed using a one-way repeated measures ANOVA over the three lighting conditions (0 Hz, 4 Hz or 8 Hz).

No significant differences were found among the three lighting conditions for any of the four scores for the airborne group or the terrestrial/aquatic group. Although not significant and only a observation, exposure to the 4 Hz stroboscopic environment resulted in greater changes in all four MSQ scores (pre- to post-MARS exposure) compared to the 8 Hz environment (i.e., the MSQ differences scores were larger after exposure to the 4 Hz stroboscopic environment than those after exposure to the 8 Hz stroboscopic lighting).

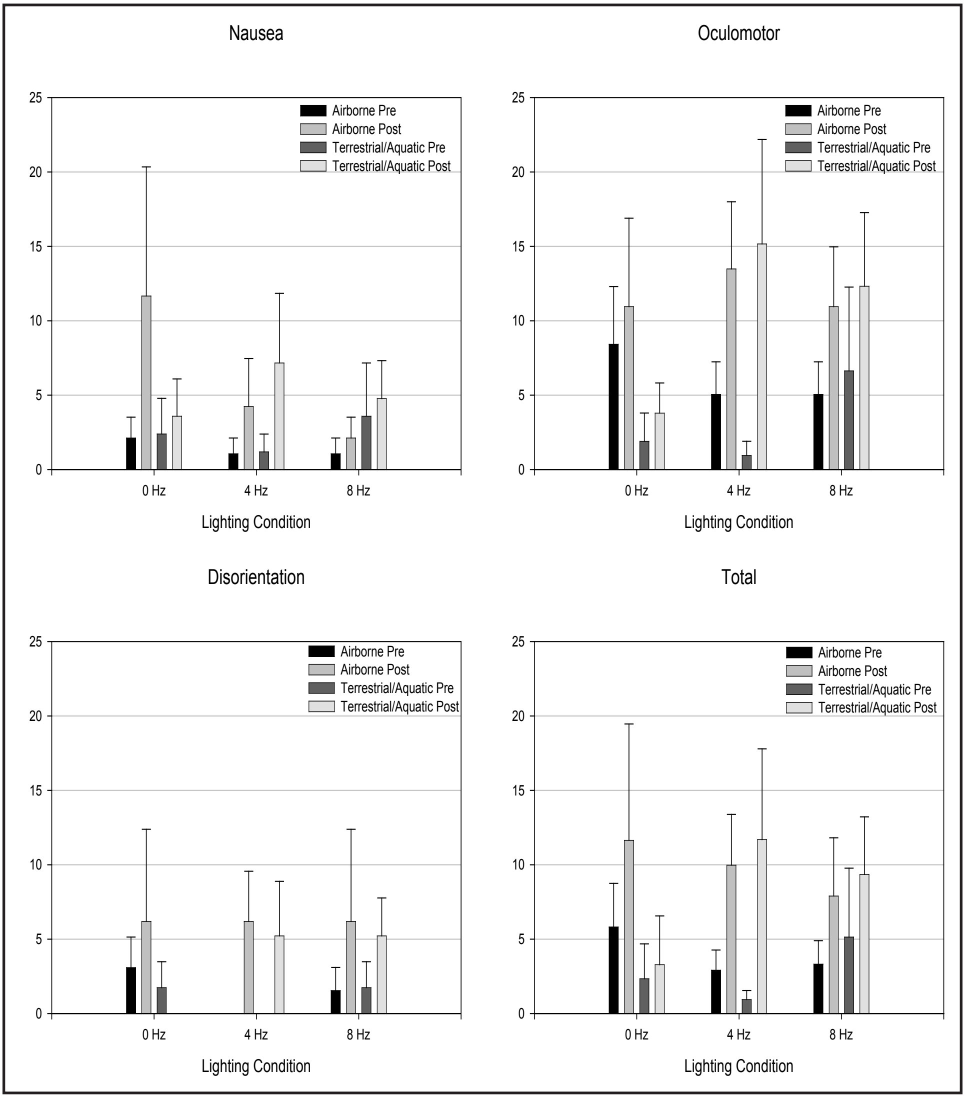

Figure 2. Mean ± SE MSQ scores

In order to determine if the exposure to the MARS resulted in significant increases in MS, paired samples *t* tests were conducted comparing participants pre- and post- MARS MSQ scores in the no strobe condition. Analysis of the no strobe lighting condition data showed that the MARS session produced increases in MSQ scores, but not to a statistically significant degree (Table 4).

Table 4

Significance Test of Pre/Post Changes in MSQ scores for the No-Strobe condition

Motion	MSQ Score		p value
Airborne Group	Nausea		.264
	Oculomotor		.545
	Disorientation		.559
		Total	.385
Terrestrial/Aquatic Group	Nausea		.685
	Oculomotor		.170
	Disorientation		.351
		Total	.563

Interestingly, when the number of symptoms reported (on the MSQ) after MARS exposure were analyzed (regardless of severity), the participants in the airborne group reported a greater mean number of symptoms after the 0 Hz condition (M = 2.89, SD = 4.57) than the 8 Hz condition (M = 2.22, SD = 2.49). However, this difference was not significant, as revealed by a one-way repeated measures ANOVA over the three lighting conditions (F(1.23, 9.87) = 0.18, p = 0.73). The most commonly reported symptoms after exposure to the MARS are presented in Figure 3.

Participants were asked to complete the PBA before and after each exposure to the MARS. Difference scores were calculated by subtracting the scores from the pre-administration from scores of the post-administration. The data were analyzed using a one-way repeated measures ANOVA over the three lighting conditions. No significant differences were found among the three lighting conditions for any of the three PBA tests for the airborne group or the terrestrial/aquatic group.

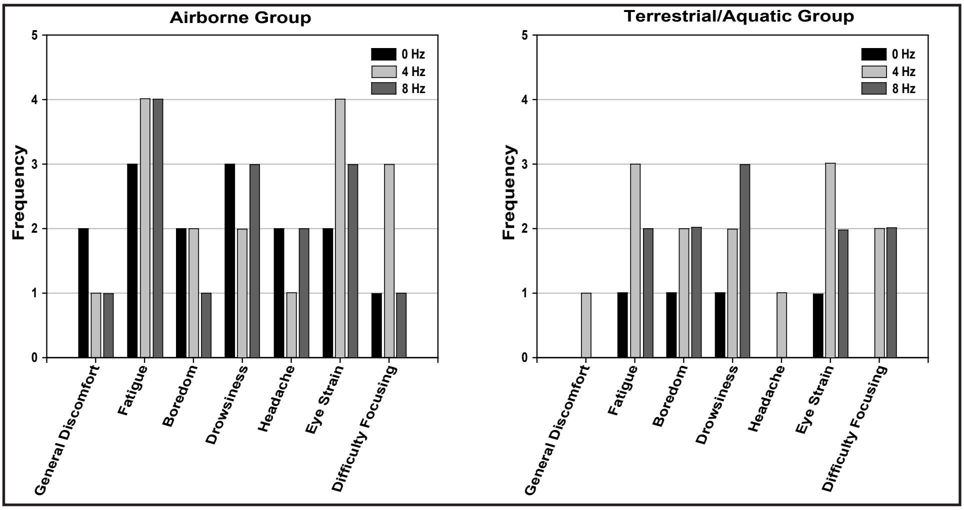

Figure 3. Frequencies of MSQ symptoms reported

Postural Balance Assessment

Upon further analysis of the PBA data, it was discovered that there were practice effects, regardless of the motion profiles. A 3 x 2 repeated measures ANOVA revealed a significant practice effect for the WOFEC test (*F*(2, 32) = 26.88, *p* < .001). Figure 4 illustrates that over the three days, participants WOFEC performance increased both before and after the MARS exposure. Participants also improved on the other two tests in the PBA over the course of the study, but the differences were not significant.

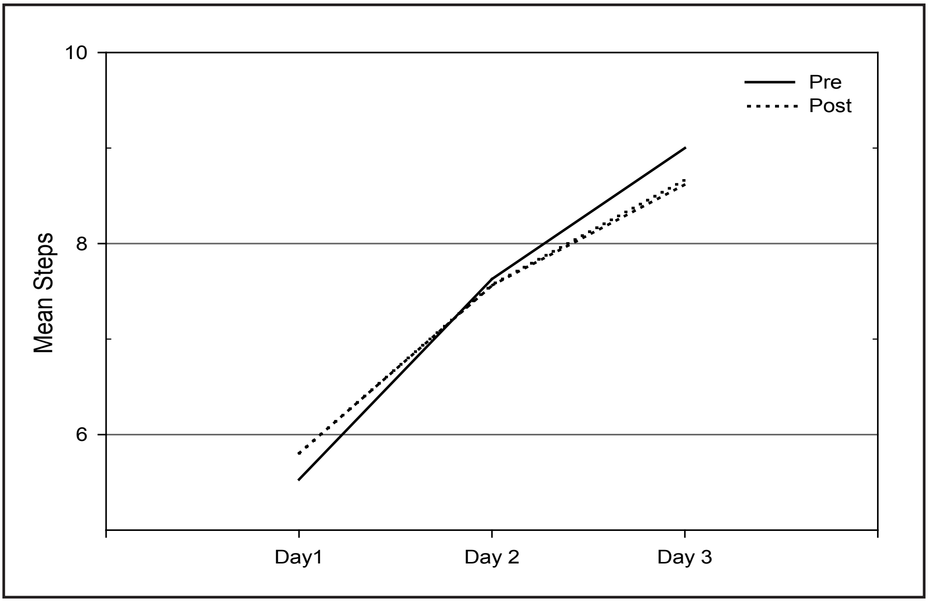

Figure 4. Practice effects during the WOFEC test

Effectiveness Questionnaire

The Effectiveness Questionnaire measured how effective and distracting the 4 and 8 Hz stroboscopic environments were to each participant. Figure 5 presents the mean ratings for the airborne and terrestrial/aquatic groups by lighting condition. A paired samples t test revealed participants in the airborne group reported the 8 Hz condition as significantly more effective in controlling MS symptoms than the 4 Hz condition (t(7) = -2.15, p = 0.03). Participants in the terrestrial/aquatic group also reported the 8 Hz stroboscopic environment as more effective with the difference approaching significance (t(7) = -1.85, p = 0.054). With regard to how distracting the stroboscopic conditions were to the participants, those in the airborne group felt the 4 Hz condition was more distracting, whereas the terrestrial/aquatic group felt the 8 Hz environment was more distracting. These differences were not significant.

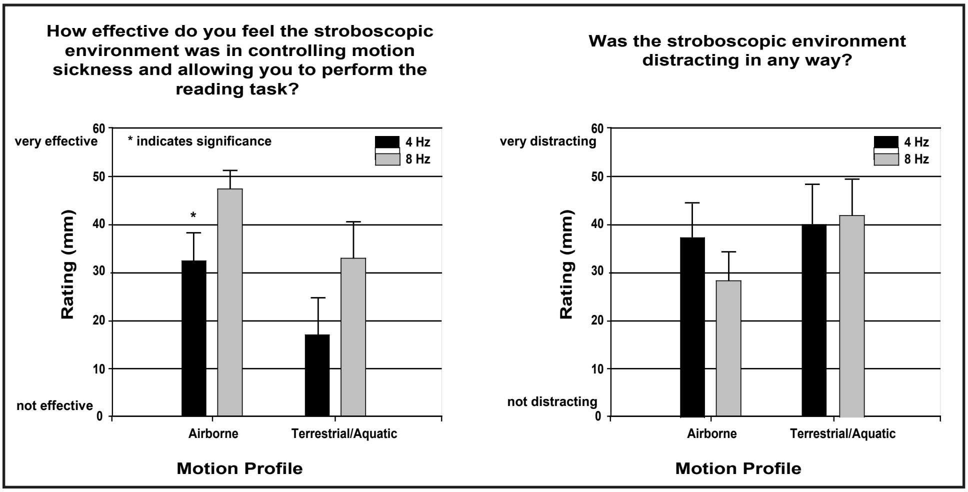

Figure 5. Effectiveness Questionnaire results (Mean ratings ± SE). Note: the higher the rating the more effective and the more distracting, respectively

Approximately 44% of the participants felt the strobe effect has a practical application for military helicopter passengers, compared to 12% who felt it did not. The remaining 44% of the participants were undecided. Study volunteers were also encouraged to provide comments regarding the effectiveness of the stroboscopic environment in reducing MS. Several participants indicated that the 8 Hz condition produced less eye strain and allowed for easier reading in comparison to the 4 Hz condition.

Discussion

Data from the Effectiveness Questionnaire showed that participants in the airborne group judged the 8 Hz stroboscopic environment as significantly more effective than the 4 Hz condition in controlling MS. This finding is similar to that found in Estrada (2007). In addition, participant comments indicated the 8 Hz environment produced less eye strain and allowed for easier reading. These results are interesting given that the participants reported the stroboscopic environment effectively controlled their MS symptoms; however there was no evidence in the other dependent measures. The present study was the first to include objective tests, such as the PVT, in the assessment of stroboscopic lighting as a countermeasure for MS. Other studies examining stroboscopic lighting have mainly relied on subjective reports of participants symptoms (Estrada, 2007; Reschke et al., 2006). There is literature that suggests that MS does not have any measureable effect on performance because motivation to perform the task is such an important factor. Motion sickness has been said to affect one's proclivity, not ability, to perform a task (Johnson, 2005). This claim may account for the results in the present study.

Another explanation for the inability to find a significant effect of the stroboscopic environment for the other dependent measures may be a lack of MS susceptibility in the study population. The two participant groups were predominantly male (i.e., 12 of 17). Some research suggests that females are more susceptible to MS (Turner, Griffin, & Holland, 2000), and this may have impacted study results. With regard to MS histories, both the airborne and terrestrial/aquatic groups tended to score low on the MHQ. A subsequent examination of the PVT, MSQ and PBA data using only those participants with the highest 25% of MHQ scores from our sample (i.e., those most susceptible to MS) failed to find a significant effect of the stroboscopic illumination; however, the statistical power was very low due to the small sample size (*n* = 4). It should be mentioned that two of the four participants most susceptible to MS reported less nausea after exposure to the MARS under the 8 Hz condition compared to the no-strobe condition. A similar observation was found in the PBA data; all four of the MS-susceptible participants were able to take more steps in the WOFEC after MARS exposure under the 8 Hz condition compared to the no-strobe condition. Future studies examining the effectiveness of stroboscopic illumination in reducing MS must focus on a MS susceptible population.

It should be noted that Reschke et al. (2006) used a modified version of the Miller and Graybiel Questionnaire which is administered and scored differently from the MS questionnaire used in the present study. For example, participants in the Reschke et al. study were asked to report any MS symptoms they were experiencing every 5 minutes during testing, while participants in the present study were only asked to complete the MSQ before and after MARS exposure. Perhaps differences in procedure and scoring contributed to the inability of the present study to find significant evidence of stroboscopic illumination as a MS countermeasure in the MSQ data.

The participants in the present study were not given time to practice any of the tests in the PBA. More practice time could have prevented the practice effects. Other studies using the PBA have also found learning effects for tests similar to the WOFEC and SOPLEC for up to as many as ten practice sessions (Hamilton, Kantor, & Magee, 1989). This practice effect prevents any conclusions from being drawn about the effectiveness of the stroboscopic environment in reducing disequilibrium related to MS.

An additional potential confound includes possible adaptation to MS due to the testing schedule. In the present study, participants were tested on a Monday- Wednesday-Friday schedule. Research has shown that individuals can adapt to MS-inducing stimuli over time (DeHart & Davis, 2002). The necessities of the testing schedule may have confounded the results with the participants exhibiting less MS over time. While there were no significant order effects for the 4 MSQ scores, mean MSQ scores after MARS exposure were lower on the last day of testing. Although the participants' head movements were not restricted during the MARS sessions, head movement variations were minimal. However, it is possible, although unlikely, that unrestrained head movements may have contributed to individual differences in MS levels.

As with many efforts examining the application of novel technologies, the examination took place in a laboratory setting and not in actual air and ground vehicles. While the MARS was a cost effective alternative, it was unable to reproduce low frequency oscillations found in actual air and ground motion profiles (i.e., below 0.5 Hz), which are most nauseogenic (Cheung & Nakashima, 2006). Analysis of the no strobe lighting condition data showed that the MARS ride produced increases in MS scores, but not to a statistically significant degree. Therefore, the authors cannot state with certainty that the MARS produced sufficient MS symptoms in most participants for the strobe conditions to counter. However, it is interesting to note that one participant in the airborne group discontinued participation due to unpleasant MS symptoms while experiencing the no-strobe condition. The strength of MS stimuli is an important variable in MS research, as there are great individual differences in susceptibility (Benson, 2002).

Conclusion

The present study was an examination of the potential for stroboscopic illumination to serve as a countermeasure for MS by using both objective and subjective measures. There was evidence of user acceptance from the Effectiveness Questionnaire, particularly for the 8 Hz condition. Future studies should examine 8 Hz stroboscopic environments, as this is the second study to reveal a preference compared to 4 Hz environments, which have been more commonly examined (Melvill-Jones & Mandl, 1981; Reschke et al., 2006). The results of this research also demonstrated the limitations associated with research involving such great individual differences in susceptibility. Clearly, examination of a susceptible population will be required for any future research examining stroboscopic illumination as a MS countermeasure.

Acknowledgements

This study is funded by the Army Medical Department Advanced Medical Technology Initiative (AAMTI) from the Telemedicine and Advanced Technology Research Center (TATRC). The authors wish to thank Ms. Edna Rath, Ms. Melody King, Mr. Brad Bumgardner, and Dr. John Crowley for their help with data collection. The opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the U.S. Army and/or the Department of Defense.

References

Benson, A. J. (2002). Motion Sickness. In K.B. Pandolf & R.E. Burr (Eds.), *Medical Aspects of Harsh Environments: Vol. 2* (pp.1048-1083). Washington DC: Borden Institute.

Cheung, B., & Nakashima, A. (2006). *A review on the effects of frequency of oscillation on motion sickness (*DRDC TR 2006-229*).* Toronto, Canada: Defense Research and Development Canada.

Cowings, P. S., Toscano, W. B., DeRoshia, C., & Tauson, R. A. (1999). *Effects of the Command and Control Vehicle (C2V) operational environment on Soldier health and performance* (NASA Technical Memorandum 1999-208786)*.* Ames Research Center, Moffett Field, CA: *National Aeronautics and Space Administration.*

DeHart, R. L. & Davis, J. R. (2002). *Fundamentals of Aerospace Medicine* (3rd ed.). Baltimore: Williams and Wilkins.

Doose, H. & Waltz, S. (1993). Photosensitivity – Genetics and clinical significance. *Neuropediatrics, 24,* 249-255.

Epilepsy Foundation. (n.d). *Photosensitivity and seizures.* Retrieved May 4, 2007, from http://www.epilepsyfoundation.org/about/types/triggers/photosensitivity.cfm

Estrada, A. (2007). *Preliminary assessment of stroboscopic shutter glasses on motion sickness in helicopter passengers (*USAARL Report No. 2007-11)*.* Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Gower, D. W., & Fowkles, J. (1989). *Simulator sickness in the UH-60 (Black Hawk) flight simulator* (USAARL Report No. 89-25). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Hamilton, K. M., Kantor, L., & Magee, L. E. (1989). Limitations of postural equilibrium test for examining simulator sickness. *Aviation, Space, and Environmental Medicine, 60,* 246-251.

Han, Y. H., Kumar, A. N., Somers, J. T., Reschke, M. F., & Leigh, R. J. (2005). Effects of retinal image slip on modulation of visual vestibule-ocular reflex during near viewing. *Annals of the New York Academy of, 1039*, 463-465.

International Organization for Standardization (ISO). (1997). *Mechanical vibration and shock – Evaluation of human exposure to whole-body vibration, Part 1: General requirements (*ISO 2631-1:1997). Geneva, Switzerland.

Johnson, D. M. (2005). *Introduction to and review of simulator sickness research (*ARI Research Report 1832). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kellogg, R. S, Kennedy, R. S, & Graybiel, A. (1965). Motion sickness symptomatology of labyrinthine defective and normal subjects during zero gravity maneuvers. *Aerospace Medicine, 36,* 315-318.

Kennedy, R. S. & Graybiel, A. (1965). *The dial test: A standardized procedure for the experimental production of canal sickness symptomology in a rotating environment* (NAMRL Report No, 113, NSAM 930)*.* Pensacola, Fl: Naval School of Aerospace Medicine.

Kennedy, R. S., Lane, N. E., Grizzard, M. C., Stanney, K. M., Kingdon, K., & Lanham, S. (2001). Use of a motion history questionnaire to predict simulator sickness. *Proceedings of the Sixth Driving Simulation Conference, Niece, France,* 79-89.

Melvill-Jones, G. & Mandl, G. (1981). Motion sickness due to vision reversal: Its absence in stroboscopic light. *Annals New York Academy of Sciences, 374,* 303-311.

National Society for Epilepsy. (n.d). *Information on epilepsy: Photosensitive epilepsy.* Retrieved May 4, 2007, from http://www.epilepsynse.org.uk/pages/info/leaflets/photo.cfm and http://www.epilepsynse.org.uk/pages/info/glossary/index.cfm#F

Reason, J. T., & Brand, J. J. (Eds.). (1975). *Motion sickness.* New York: Academic Press.

Reschke, M. F., Somers, J. T., & Ford, G. (2006). Stroboscopic vision as a treatment for motion sickness: Strobe lighting vs. shutter glasses. *Aviation, Space, and Environmental Medicine*, *77,* 2-7.

Stroboscopic goggles for reduction of motion sickness. (2005). Retrieved February 8, 2008, from http://www.techbriefs.com/component/content/article/236

Turner, M., Griffin, M. J., & Holland, I. (2000). Airsickness and aircraft motion during short-haul flights. *Aviation, Space and Environmental Medicine, 71*, 1181-89.

International Journal of Applied Aviation Studies, Volume 9, Number 2

An Exploration of Deviations in Aircraft Maintenance Procedures

Robert I. Baron, Ph.D

The Aviation Consulting Group
April 2009

Abstract

An informal study on deviations in aircraft maintenance procedures was conducted during a human factors training course in March of 2009. The purpose of the study was to pilot test the Maintenance Events Checklist (MEC) with a relatively small sample of aircraft maintenance technicians. The MEC is intended to capture participants' responses to statements related to maintenance deviations. Participants consisted of aircraft maintenance technicians who worked in non-airline operations (i.e., corporate and business aviation, helicopter operators, and FBOs). Results showed that nearly 50% of the participants' indicated they had "very rarely" deviated from the MEC content items. However, 22% indicated they had deviated "occasionally" and 5% indicated they had deviated "often." Causes and implications are discussed, which focus on organizational pressure, individual complacency, and deficiencies in aircraft maintenance documentation itself.

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.

Return to Table of Contents

An Exploration of Deviations in Aircraft Maintenance Procedures

Deviations from approved procedures have been implicated in a number of maintenance-related aircraft accidents. These deviations may stem from factors such as time pressure, stress, fatigue, lack of resources, or ambiguous or confusing documentation. These factors typically do not occur in isolation but are linked together and may increase the likelihood of skipped steps, signoffs without verification, or continuing a job without the correct tools or equipment. This was recently the case where American Airlines Flight 1400 experienced an in-flight engine fire requiring a turnback and emergency landing in St. Louis (STL). The investigation revealed that a component in the manual start mechanism of the engine was damaged when a mechanic used an unapproved tool to initiate the start of the #1 (left) engine while the aircraft was parked at the gate at STL. The deformed mechanism led to a sequence of events that resulted in the engine fire, to which the flight crew was alerted shortly after take-off (National Transportation Safety Board, 2009).

Hobbs (2002) developed the Maintenance Events Checklist (MEC) that is intended to capture aircraft maintenance technicians' (AMTs) responses to statements related to deviations. While not a scientific instrument, the MEC is an effective tool to aid researchers in identifying problematic areas in aviation maintenance activities. Once these areas are identified, further research can be conducted to address specifically those deviations which appear to be most problematic.

Previous studies have identified a number of recurring problems in aircraft maintenance tasks. Some of the more recent studies have utilized the Aviation Safety Reporting System (ASRS) database to provide a rich, and relatively new, source of information relating to aviation maintenance errors. One such study used ASRS reports to identify associations between existing ASRS codes and the area of the aircraft involved in the incident (using ATA coding) (Hobbs & Kanki, 2008). The authors point out that these associations have not been adequately examined in the past but could be a valuable source of data for human factors training, the design of procedures, and the identification of improvements in aircraft design (p. 5). A recent study in the United Kingdom also segregated maintenance errors by ATA code. In that study the researchers found that the top three most frequent maintenance errors by aircraft area were (1) Equipment and Furnishings (ATA 25: 18.27%), (2) Landing Gear (ATA 32: 10.6%), and (3) Flight Controls (ATA 27: 8.59%) (Civil Aviation Authority, 2009).

Other studies have investigated specific types of maintenance errors that occur at the AMT level. A study in the United Kingdom found that the majority of errors were incorrect installation of components, the installation of wrong parts, electrical wiring discrepancies, and tools left in the aircraft (Civil Aviation Authority, 1992). Another study found that incomplete installations, incorrect assembly or location, vehicles or equipment contacting the aircraft, material left in the aircraft, wrong part, and part not installed were the most common incidents reported by AMTs in Australia (Hobbs & Williamson, 2002). According to Hobbs and Williamson (2003) these types of errors can be broadly categorized into three groups: memory failures, rule violations, or knowledge-based errors.

Memory failures address prospective, rather than retrospective memory. In other words, an AMT forgets to do something in the future such as removing a tool from an interior aircraft work area.

Rule violations may include deviating from documented procedures or developing a completely different way to conduct a task. A study conducted in Europe by McDonald, Corrigan, Daly, and Cromie (2000) found that 34% of the aircraft mechanics in their study acknowledged that their most recent task was performed in a way that contravened formal procedures. The previously mentioned Australian study by Hobbs and Williamson (2002) found that 30% of AMTs acknowledged that in the previous 12 months they had decided not to perform a functional check or an engine run. Over 30% reported that they had signed off a task before it was completed, and over 90% reported having done a task without the correct tools or equipment.

Rule violations may be propagated at the individual or organizational levels. Patankar (2002) conducted a comprehensive study that investigated rule violations by aircraft mechanics. On the individual level he found that the most problematic areas, ordered by percentage, were (1) lack of awareness, (2) complacency, (3) time constraints, (4) lack of knowledge or experience, and (5) workplace distractions. An individual factor was defined as, “A maintenance error that seemed to be within the individual mechanic’s span of control” (p. 15). From the organizational perspective the results showed that the most problematic areas, ordered by percentage, were (1) procedures or information quality, (2) aircraft design/configuration of system or quality of parts, (3) maintenance management or leadership, (4) workplace norms/peer pressure, and (5) lack of training.

Knowledge-based errors are a result of failed problem-solving or a lack of system knowledge (Rasmussen, 1983). An AMT might be unfamiliar with a particular system and uses past experiences or trial and error to dictate the steps to be taken for a particular task. An example of this might be an AMT assuming that a red rocker switch is an avionics master when in fact the switch is an engine starter. This assumption is made based on previous experience where the red rocker switch on a particular model aircraft was used as an avionics master. While working on a different type of aircraft the AMT engages the red rocker switch and inadvertently begins to rotate the propeller. If someone is in the proximity of the engine, serious harm or even death may occur.

While much of the extant literature focuses on error types as a result of organizational or individual deficiencies, only recently have researchers begun to investigate potential deficiencies in aircraft maintenance documentation. Researchers postulate that many of the aforementioned shortcuts and bootleg procedures may be the result of poor quality documentation. For instance, procedural errors, which are defined as any information-related error involving documents (Maintenance Error Decision Aid, 1994), have been implicated in 44% to 73% of maintenance errors (Veinott & Kanki, 1995; Nord & Kanki, 1999; Patankar, Lattanzio, Kanki, & Munro, 2003). Nord and Kanki (1999) found that errors were fairly evenly distributed across documents. The three most problematic areas identified were inspection and verification issues (34%), incompleteness of the documents (27%), and incorrectness of the documents (22%). Similarly, in a study of 458 ASARs reports submitted by AMTs, Lattanzio, Patankar, and Kanki (2008) found that the most frequently cited maintenance document deficiencies were missing information (48%), incorrect information (19%), difficult to interpret (19%), and conflicting information (19%).

Rogers, Hamblin, and Chaparro (2008) conducted a seminal study on the types of errors found in aircraft maintenance manuals published by four manufacturers. Their study was based on Publication Change Requests (PCRs) by AMTs. Results showed that the majority of PCRs related to procedures found in Flight Controls, Landing Gear, and Powerplant systems. These PCRs appear to mirror the most frequent problem areas identified in the previously mentioned ATA code segregation study (Civil Aviation Authority, 2009). The highest percentage of PCRs involved Procedural Errors (42.5%) followed by Language (29.9%), Technical (16.5%), Graphic (8.1%), and Effectivity (n/a). Common procedural errors were categorized as Step(s), Ordering, Alternate method, Check/Test/Inspection, Caution/Warning. Language errors included typographical errors (Typos), grammatical errors (Grammar), a need for clarification of the information (Clarity), and inaccurate information within a step (Incorrect) (pp. 301-302).

Most of the previous studies have focused primarily on line maintenance operations and thus there has been somewhat of a paucity in studies related to AMTs who work in other settings such as corporate and charter flight departments, fixed base operators (FBOs), and helicopter operators. This study is intended to provide insight into these areas of aviation maintenance that tend to be overlooked in current research. It is not the intent to infer the results are descriptive of the larger population of AMTs in non-airline operations. Instead, this informal study was conducted as a pilot test and possibly a starting point for future research in this area.

Method

Participants

The sample consisted of 27 AMTs (*N*=27) who were attending a human factors training course in Minneapolis, MN in March of 2009. Participants were assured anonymity and all in attendance participated voluntarily in the study. Participants' job position was the only demographic data captured. Positions were categorized as Supervisor/Inspector (67%), Line Mechanic (30%), and Director of Maintenance (3%). The majority of AMTs were employed in business/corporate aviation, FBOs, and helicopter operations.

Instrument

The MEC was distributed to all participants in the human factors training course. The MEC is a short, informal checklist that is used to capture AMTs responses to statements related to maintenance deviations. The checklist consists of seven statements with four response choices labeled; Never (1), Very Rarely (2), Occasionally (3), and Often (4). The complete MEC is shown in Table1.

Procedure

The MEC is typically distributed as a single survey and worded exactly as depicted in Table 1. However, in the current study, the MEC distribution was modified in that it was split into two duplicate checklists with the following exceptions. The first checklist asks the participant to estimate how often these items have been done by *themselves* in the last year. The second checklist asks the participant to estimate how often these items might have been done by *other* workers in their organization in the last year.

Table 1

Maintenance Events Checklist (Hobbs, 2002)

At work in the last year, how often would a *typical* maintenance worker have:
Done a job without the correct tool or equipment
Decided not to do a required functional check because of a lack of time
Done a job a better way than that in the manual or approved maintenance documents
Corrected an error by another maintainer, but not documented what they had done to avoid getting the person into trouble.
Done an unfamiliar job without being certain they were doing it correctly
Been misled by confusing documentation
Signed a job on behalf of another person without checking it

Results

A total of 28 sets of MECs were collected. However, one MEC had to be discarded due to incompleteness (more than half of the statements had no response). Data were analyzed with SPSS v. 15.0 (2006) software. Descriptive statistics were used for basic data analysis which included mean response scores, standard deviations, and percentages. Inferential statistics were not used due to sampling limitations which are described in the Discussion section.

The results were broken down into four graphic presentations. The first and second presentations show participants' responses about themselves and others (Table 2 and Table 3 respectively). The mean scores represent the average of the responses for all participants for each statement. The standard deviation (SD) shows how much variability there is from the mean. The smaller the SD, the less variability there is in the participants' overall answers. The third and fourth presentations show a breakdown by percentage of the participants' responses about themselves and others (Table 4 and Table 5 respectively).

Table 2

Participants' Responses About Themselves

Item No.	Description	*M*	*SD*
1	Done a job without the correct tool or equipment.	2.22	0.5773
2	Decided not to do a required functional check because of a lack of time.	1.59	0.6360
3	Done a job a better way than that in the manual or approved maintenance documents.	2.37	0.8835
4	Corrected an error by another maintainer, but not documented what they had done to avoid getting the person into trouble.	2.14	0.8182

5	Done an unfamiliar job without being certain they were doing it correctly.	1.96	0.7586
6	Been misled by confusing documentation.	2.51	0.7000
7	Signed a job on behalf of another person without checking it.	1.85	0.9488

Mean Score Key: Never (1), Very Rarely (2), Occasionally (3), Often (4).

Table 3

Participants' Responses About Others

Item No.	Description	M	SD
1	Done a job without the correct tool or equipment.	2.59	0.5723
2	Decided not to do a required functional check because of a lack of time.	1.88	0.5773
3	Done a job a better way than that in the manual or approved maintenance documents.	2.33	0.7337
4	Corrected an error by another maintainer, but not documented what they had done to avoid getting the person into trouble.	2.40	0.7472
5	Done an unfamiliar job without being certain they were doing it correctly.	1.96	0.6493
6	Been misled by confusing documentation	2.55	0.6405
7	Signed a job on behalf of another person without checking it.	2.03	0.9397

Mean Score Key: Never (1), Very Rarely (2), Occasionally (3), Often (4)

Table 4

Participants' Responses About Themselves

	Response (%)			
	Never	Very Rarely	Occasion-ally	Often
Done a job without the correct tool or equipment.	7	63	30	0
Decided not to do a required functional check because of a lack of time.	48	44	7	0
Done a job a better way than that in the manual or approved maintenance documents.	15	44	30	11

Corrected an error by another maintainer, but not documented what they had done to avoid getting the person into trouble.	18	56	18	7
Done an unfamiliar job without being certain they were doing it correctly.	26	56	15	4
Been misled by confusing documentation.	4	48	41	7
Signed a job on behalf of another person without checking it.	44	33	15	7
M =	23	49	22	5

**Percentages have been rounded and may not equal 100%*

Table 5

Participants' Responses About Others

	Response (%)			
	Never	Very Rarely	Occasion-ally	Often
Done a job without the correct tool or equipment.	4	33	63	0
Decided not to do a required functional check because of a lack of time.	22	67	11	0
Done a job a better way than that in the manual or approved maintenance documents.	15	37	48	0
Corrected an error by another maintainer, but not documented what they had done to avoid getting the person into trouble.	7	52	33	7
Done an unfamiliar job without being certain they were doing it correctly.	22	59	19	0
Been misled by confusing documentation	4	41	52	4
Signed a job on behalf of another person without checking it.	33	37	22	7
M=	15	47	35	3

**Percentages have been rounded and may not equal 100%*

Discussion

Although no formal hypotheses were posited or tested in this study, there is enough descriptive data to look at overall trends. Participants rating themselves tended to rate lower (deviate less) than when rating others. In fact, when looking at the mean percentages in Table 4 and Table 5, there is evidence that in almost all response sets participants believed that others have deviated from procedures more often than themselves. This could be due to AMTs perceptions that “other mechanics would do that but I never would.” The effects of the fundamental attribution error and locus of control, for instance, may influence these different perspectives between the two versions of the MEC. This social-psychological influence was expected and could be explored further in a future study. However, the rest of this section will focus specifically on participants’ responses about themselves.

When percentages were averaged, nearly 50% of the participants’ indicated they would *“very rarely”* deviate from the MEC content items. However, 22% indicated they would deviate *“occasionally”* and 5% indicated they would deviate *“often.”* Twenty three percent of the participants indicated they would *“never”* deviate from procedures. Although this was a quantitative, descriptive study and lacked qualitative input from participants, one may infer why these deviations might be occurring in the maintenance hangar. For example, 7% of the participants indicated that they had *occasionally* decided not to do a required functional check because of a lack of time. Time pressure can be a significant factor in aircraft maintenance operations but the root cause of this pressure tends to be propagated by a widespread organizational culture that condones inappropriate or unrealistic deadlines. Conversely, the omission of this important step may occur at the individual AMT level as a manifestation of complacency. The well-intentioned AMT may have done the job hundreds of times before and never had a problem. As such the AMT may decide to skip a functional check assuming everything has been done correctly. Coupled with a perceived time savings, it may become very tempting for the AMT to skip this required step. Thirty percent of the participants indicated that they have *occasionally* done a job without the correct tool or equipment. This could be due to a lack of resources occurring at both the organizational and individual levels. Poor quality documentation may account for 41% of the participants *occasionally* being misled by confusing documentation or the 30% that indicated that they have *occasionally* done a job a better way than that in the manual or approved maintenance documents. Eleven percent indicated that they *often* do a job in this manner. This could be due to the quality of the documentation itself and may correlate with the study by Rogers, Hamblin, and Chaparro (2008) discussed earlier.

Limitations

Although this study was conducted in an efficient and impromptu manner which facilitated a quick collection of data, there were a number of limitations that may have affected the results. First, the convenience sample was quite small (*N*=27) and was from one geographic location (Minnesota) in the United States. Thus the results may not be representative of the broader AMT population. Second, the study was conducted informally at a human factors course and lacked the rigor of the scientific method. Third, the human factors course itself may have acted as a confounding variable in that the MEC was distributed toward the end of the course. Due to these limitations one should interpret the results of this study with caution. This research was intended to be a pilot study; however, the results could be used as a foundation for additional, and more formal, studies relating to aircraft maintenance deviations.

Conclusion

The results of this study provide a look into deviations in aircraft maintenance procedures with a focus on non-airline operations. A person from outside the aviation domain might look at these results and be astonished that AMTs would deviate from required procedures at all. Yet, for those who work in the aviation maintenance domain, these results may not come as a surprise. Regardless, one must be careful not to assume that deviations are simply symptomatic of “bad apple mechanics.” There are a number of reasons why these deviations occur which may include organizational pressure, norms, complacency, and poor quality documentation. Each item on the MEC carries with it its own special nuances with corresponding complexities for reparation. Simple exhortations such as “make sure you always do a functional check” will not be effective. Organizations need to emphasize to their AMTs, in an ongoing manner, the importance of following approved procedures. Similarly, AMTs need to be aware of the consequences of deviating from procedures. Human factors training courses should emphasize that the majority of maintenance-related aircraft accidents and incidents have been the result of deviations from approved procedures.

Finally, researchers should continue to investigate aircraft maintenance documentation itself. Evidence has been found to indicate that there are a number of problems with manufacturers’ written procedures. Until documentation issues such as missing information, incorrect information, difficulty in interpretation, and conflicting information are resolved, it will be difficult to sell the idea of following approved procedures to AMTs on a far-reaching basis.

References

Civil Aviation Authority. (1992). *Flight Safety Occurrences Digest.* (92/D/12). London.

Civil Aviation Authority. (2009). *Aircraft maintenance incident analysis.* CAA Paper 2009/05. Retrieved April 1, 2009, from http://www.caa.co.uk/docs/33/2009_05.pdf.

Hobbs, A. (2002). *Maintenance Events Checklist.* Presented at the 1st Airbus Flight Operations Monitoring and Safety Development Conference, Hong Kong, 12-13 March, 2002.

Hobbs, A., & Kanki, B. G. (2008). Patterns of error in confidential maintenance incident reports. *The International Journal of Aviation Psychology, 18*(1), 5-16.

Hobbs, A., & Williamson, A. (2002). Human factor determinants of worker safety and work quality outcomes. *Australian Journal of Psychology, 54,* 151-161.

Hobbs, A., & Williamson, A. (2003). Associations between errors and contributing factors in aircraft maintenance. *Human Factors, 45,* 186-201.

Lattanzio, D., Patankar, K., & Kanki, B. G. (2008). Procedural error in maintenance: A review of research and methods. *The International Journal of Aviation Psychology, 18*(1), 17-29.

Maintenance Error Decision Aid (1994). [Results Form]. Collaborative effort of The Boeing Company, British Airways, Continental Airlines, United Airlines, the Federal Aviation Administration and the International Association of Machinists.

McDonald, N., Corrigan, S., Daly, C., & Cromie, S. (2000). Safety management systems and safety culture in aircraft maintenance organizations. *Safety Science, 34,* 151-176.

National Transportation Safety Board. (2009). *In-flight left engine fire: American Airlines flight 1400.* Retrieved April 1, 2009, from http://www.ntsb.gov/Publictn/2009/AAR0903.htm

Nord, K., & Kanki, B. G. (1999). Analysis of procedural errors in aircraft maintenance operations. In R. Jenson, *Proceedings of the 10th International Symposium on Aviation Psychology*. Dayton, OH: Ohio State University.

Patankar, K., Lattanzio, D., Kanki, B. G., & Munro, P. A. (2003). Identifying procedural errors in ASRS maintenance reports using MEDA and QUORUM. In R. Jensen, *Proceedings of the 12th International Symposium on Aviation Psychology*. Dayton, OH: Ohio State University.

Patankar, M. S. (2002). *Root cause analysis of rule violations by aviation maintenance technicians.* Grant No. 2001-G-001. Washington, DC: FAA Office of Aviation Medicine.

Rasmussen, J. (1983). Skills, rules and knowledge: Signals, signs and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics, 13,* 257-266.

Rogers, B. L., Hamblin, C. J., & Chaparro, A. (2008). Classification and analysis of errors reported in aircraft maintenance manuals. *International Journal of Applied Aviation Studies, 8*(2), 295-309.

SPSS for Windows. (2006). Rel. 15.0.0. Sept/2006. Chicago: SPSS Inc.

Veinott, E., & Kanki, B. G. (1995, September). *Identifying human factors issues in aircraft maintenance operations.* Poster session presented at the annual meeting of the Human Factors and Ergonomics Society, San Diego, CA.

International Journal of Applied Aviation Studies, Volume 9, Number 2

Workplace Preferences of Millennials In the Aviation Industry

Mary Niemczyk and Jon W. Ulrich
Arizona State University

Abstract

Previous research has determined that each generation has its own unique attitudes, work ethics, distinct, and preferred ways of managing and being managed. Today's workplace represents the largest diversity of generations than any other time in history. The examination of generational differences among workers is a critical and underdeveloped area of investigation, particularly in aviation. The purpose of this study was to determine the work environment preferences of the Millennial generation in the aviation industry. The Work Environment Scale – Form I (Moos & Insel, 1974) was administered to 290 aviation personnel. Results of this study portray a complex combination of relationship, personal growth, and organizational structures of their ideal workplace preferences, dominated by desires for greater personal freedoms with less managerial 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.

Return to Table of Contents

Workplace Preferences of Millennials
In the Aviation Industry

Today, like no other time in the history of the United States, the workforce is comprised of the greatest number of generational cohorts. Many organizations have employees representing four generations (Zemke, Raines, & Filipczak, 1999); all may be working concurrently, however, they may not be working collaboratively. Generational cohorts tend to view the world, and the workplace, from the standpoint of their distinct life experiences (Schuman & Scott, 1989), yet they are expected to act as a team to meet the goals of the organization. Unfortunately, their differences may often lead to miscommunication, employee conflict, work ethic debates, loyalty issues, varying wants and needs in terms of compensation, and training issues. A key to organizational success is to understand the perspectives and desires of each generation and be respectful of their differences (Gravett & Throckmorton, 2007).

The examination of generational differences among workers is a critical and underdeveloped area of investigation, particularly in aviation. Ultimately, all organizations are most strongly influenced by the values and preferences of their newest generation of employees. Failure on the part of management to understand and adjust appropriately to generational differences can result in misunderstandings and miscommunications. In time, this can affect employee productivity, performance, recruitment, retention, and safety.

The most recent group entering the workforce is referred to as the millennial generation, marked by having been born in the 20th Century while entering the workforce in the 21st Century. The purpose of this study is to determine the preferred workplace characteristics of the *millennial generation* in the aviation industry.

Literature Review

The distinction between where one generation ends and another begins is not rigidly defined, however, experts have found that individuals can be strongly united based on enduring shared social, economic, and political events. Members of all generations are likely to experience significant events during their developmental, adolescent years. These experiences tend to strongly impact individuals and form lifelong impressions affecting their outlook on life and work (Glass, 2007; Schuman & Scott, 1989). Individuals in each generational group seem to develop similar attitudes, ambitions, and a synergy that can provide them strength in society (Zemke, Raines & Filipczak, 1999).

Although several different terms have been used to categorize the various generations, the most common appear to be: Veterans, Baby Boomers, Generation X, and Millenials (Zemke, Raines & Filipczak, 1999). Additionally, the time frames used to define each of the generations vary throughout the literature. Generation-defined characteristics may not fit some individuals as they may adopt some of the values and attitudes of the previous or subsequent generation. Classifying generational members, therefore, is not an exact science (Johnson & Wilson, 2008). Generally speaking, and for purposes of this investigation, the generations currently in the workplace include: Veterans, born between 1922 – 1943; Baby Boomers, born between 1944 – 1960; Generation Xers, born between 1961 – 1980; and, Millennials, born between 1981 – 2000 (Lancaster & Stillman, 2002; Smola & Sutton, 2002; Strauss & Howe, 1997; Zemke, Raines & Filipczak, 1999).

Generations tend to form a *persona* or set of characteristics by which they are defined. These characteristics may include attitudes towards work, technology, gender roles, religion, race, and family. Their value systems may also deviate from other generations. These characteristics not only provide commonality within each generational group, but also the distinction between generational groups making each group somewhat unique. These characteristics seem to permeate the generation and become influential elements in the mind-set of its members (Schuman & Scott, 1989). Interestingly, these generational characteristics tend to remain fixed as the members age and seem to be lifelong traits (Strauss & Howe, 1997).

Each generation also tends to have distinct preferences regarding organizational business structures and behaviors (Glass, 2007). In stark contrast to previous generations, the Millennial generation is unwilling to dedicate much of their daily life to their work. They instead prefer to have more of a balance between work and their other interests (Smola & Sutton, 2002), and seem to exhibit a ‘work to live’, not ‘live to work’ attitude (Ryan, 2007).

When they are at work, however, the Millennial generation has a strong preference for structure and organization. This may be the result of inordinate amounts of time spent in the highly structured and controlled media environments of their electronic games, such as *Gameboys* and *Nintendos*. They prefer orderly work environments and dislike ambiguity of any kind. This cohort tends to prefer clear expectations and has a strong desire for a well-defined career path (Epstein & Howes, 2009). For some members, the desire for structure is so unyielding that, if it is not provided, they may quit their job and search for it at another organization (Westerman & Yamamura, 2007).

This generation is accustomed to using all types of technology, and incorporates it into many aspects of their lives. Their experience with various technologies has provided them with unprecedented freedoms, as well as immediate gratification. Unlike previous generations, Millennials are accustomed to instantaneous access to money (ATM), entertainment (iPod), information (Internet), communication (computers and smart phones), and even dating (online dating services) (Teaching the Millennials, 2007; Zemke, Raines & Filipczak, 1999). They have spent countless hours on the Internet and have hundreds of friends via social-networking sites such as *Facebook* and *MySpace*. Because they have been able to explore the world via the Internet, they tend to enjoy extensive freedoms and the desire to make their own decisions (McGlynn, 2005; Zemke, Raines & Filipczak, 1999).

This generation tends to be self-confident and may appear to have an entitlement attitude. Raised by *Baby-Boomer* parents, they were placed at the center of their families’ existence during their formative years. Unlike previous generations who only received a trophy for finishing in either first, second, or third place in competitions, members of this generation received a trophy for simply being a member of the team. They are accustomed to receiving praise and accolades and having their parents being quite involved in their daily lives (Zemke, Raines & Filipczak, 1999).

Generational differences have been the force behind many societal shifts. As newer generations emerge and replace older generations, a "generational replacement" tends to occur. Society is likely to transform and begin to reflect the attitudes and values of the upcoming generation (Johnson & Wilson, 2008; Mitchell, 1995). The Millennial generation represents the second largest of the current generations, following the Baby Boomers. Approximately 81 million Americans are members of this cohort, which is about one-fourth of the entire population (US Census Bureau, 2009). Due to their size, it is almost certain that they are, and will continue to impact work environments.

The Millennial generation has recently entered the workforce and many aviation employers are wondering how to recruit, manage, motivate, and communicate with these individuals. In order to assist aviation managers and supervisors in becoming more adept in coordinating the efforts of this cohort, it is important to determine the interpersonal desires, goal orientations, supervisory methods, and organizational structures that may work best with them. The purpose of this study, therefore, was to determine the work environment preferences of the Millennial generation in the aviation industry.

Method

Subjects

Participants in this study held various roles within aviation, including flight student, professional pilot, air traffic controller, aviation maintenance technician, and aviation administrator. There were 290 participants, ranging in age from 18 to 27 years, with mean and median ages of 20.5 and 20.0 years, respectively. There were 219 male and 24 female respondents, with 47 respondents choosing to not indicate their gender. All participants were born and raised in the United States. Participation was voluntary and uncompensated.

Procedures

The Work Environment Scale – Form I (WES) (Moos & Insel, 1974) was administered to individuals in various aviation entities; airlines, manufacturing, flight schools, and air traffic control facilities. The survey was paper-based. Participants were provided a written description of the study, along with the survey question booklet and response sheet. Scores were manually tabulated using a scoring template provided with the assessment materials.

Materials

The WES – Form I provides individuals the opportunity to describe what they consider to be their expected or ideal work setting. It has been used extensively in a variety of clinical and research practices, as well as by managers and consultants attempting to determine employee workplace goals and value orientations. The instrument has shown validity in predicting outcomes in various occupational settings, including the military, education, government, and health care. It has been used extensively throughout the US and internationally and has been translated into seven languages. By design, the WES is descriptive rather than evaluative (Moos, 1994b).

The WES consists of three forms: the Real Form (Form R) which measures the perceptions of employees in their current work environment; the Ideal Form (Form I) which measures employees' perceptions of their ideal workplace; and the Expectations Form (Form E) which measures prospective employees' expectations about a future work setting (Moos, 1994a; Moos, 1994b). Since its development, the WES has been found to provide significant insight into employee workplace perceptions. Examples of uses include determining employee satisfaction in accounting and business organizations (Westerman & Cyr, 2004: Westerman & Simmons, 2007; Westerman & Yamamura, 2006), determining nurse's perceptions of their real and ideal work environments (Baker, Carlisle, Riley, Tapper & Dewey, 1992; Kotzer, Koepping, & LeDuc, 2006; Long, Williams, & Hollin, 1995), as well as workplace satisfaction of mental health practitioners (McRae, Prior, Silverman, & Banerjee, 2007), to name a few.

The WES consists of 10 subscales that assess three underlying sets of dimensions: Relationship, Personal Growth or Goal Orientation, and System Maintenance and Change. The Form I subscale internal consistencies (Cronbach's Alpha) and intercorrelations range from 0.55 to 0.74. Norms have been developed for each WES form and for each country of use. In total, there are 90 True/False statements, 9 items for each of the 10 subscales. Possible scores for each subscale range between 0 and 9. The WES Form- R, Form-I, and Form-E are parallel in that each has 90 items that focus on the same work setting elements but are worded appropriately to assess current, ideal, and future perspectives. The scoring keys and answer sheets are identical for each of the three forms (Moos & Insel, 1974; Moos, 1994a; Moos, 1994b).

Within each of the three dimensions are subscales assessing various aspects of the particular dimension. The Relationship dimension consists of Involvement, Coworker Cohesion, and Supervisor Support subscales. The Personal Growth and Goal Orientation dimension consists of Autonomy, Task Orientation, and Work Pressure subscales. The System Maintenance and System Change Dimension consists of Clarity, Managerial Control, Innovation, and Physical Comfort subscales (Moos & Insel, 1974; Moos, 1994a; Moos, 1994b).

Data Analysis

Adopting the method developed by the survey developers (Moos & Insel, 1974; Moos, 1994a; Moos, 1994b), a scoring template was used to determine raw scores for each individual. Descriptive statistics for centrality and relative variation were generated from these data. The data for each individual were then summed for each of the 10 subscales. Chi-square analyses were then performed to determine which factors, if any, dominated the Millennial generation's workplace preferences.

Results

Descriptive Analysis

Subscale mean, median, modal, standard deviation and coefficient of variation (CV) values are reported in Table 1. The interpretation of the centrality measures for each subscale are explained by Moos (1989) and vary based on each subscale, e.g., *considerably below average* scores for the Involvement subscale are in the 0-3.5 range, while they are in the 0-2 range for the Innovation subscale. Interpretations for each of these statistics follows.

Table 1

Subscale Descriptive Statistics

	Relationship			Personal Growth / Goal Orientation			System Maintenance / System Change			
	Involvement	Peer Cohesion	Supervisor Support	Autonomy	Task Orientation	Work Pressure	Clarity	Control	Innovation	Physical Comfort
Mean	7.7	7.4	7.3	7.3	7.8	3.6	7.0	5.4	6.4	7.1
Median	8	8	8	8	8	3	7	6	7	8
Mode	9	9	8	8	9	4	8	7	7	8
Std Dev	1.6	1.8	1.5	1.6	1.5	2.1	1.4	2.1	2.2	1.7
CV	20%	24%	21%	22%	19%	57%	20%	39%	34%	23%

The Relationship Dimension

The first dimension measured by the WES is *Relationship*. The items on its subscales address personal relationships in the workplace, focusing on employee commitment, collegiality, and support. The three subscales of this dimension are *Involvement*, *Peer Cohesion*, and *Supervisor Support* ((Moos & Insel, 1974; Moos, 1989; Moos, 1994a; Moos, 1994b).

Involvement Subscale. The mean score on this subscale is 7.7, which is considered well above average, while the median and modal scores are 8 and 9, respectively. A CV of 20% indicates low relative variation within the respondents' scores.

This subscale measures the employee's commitment to their job. The items on this subscale ask respondents about their desire for challenging and interesting work, the effort they exert in performing their work, and whether employees assist each other in accomplishing tasks. Based on their responses, the survey respondents indicate that they would be highly committed to their jobs.

Peer Cohesion. The mean score on this subscale is 7.4, which is well above average, while the median and modal scores are 8 and 9, respectively. While a CV of 24% indicates more relative variation than the Involvement subscale, it is still relatively low.

This subscale measures the amount of support employees provide each other. The items on this subscale ask respondents about the depth and loyalty of the relationships people will develop within the workplace. Based on their responses, the survey respondents indicate a preference to work in a very supportive, cohesive environment.

Supervisor Support. The mean score for this subscale is 7.3, which is considered to be well above average. The median and modal scores are each 8, while the CV is 21% which is more in line with the Involvement subscale for relative score variation.

This subscale measures the extent by which management facilitates a supportive work environment. Items on this subscale ask respondents about how often supervisors compliment employee performance, provide positive feedback, and provide open lines of communication. With a well above average score, the respondents indicate a desire for a substantial amount of support and recognition from management.

Personal Growth or Goal Orientation Dimension

This second set of WES dimensions consists of personal growth and goal-oriented subscales. This dimension includes the autonomy, task orientation, and work pressure items. Overall, this set of items focuses on the ways in which an environment encourages or prevents personal growth (Moos & Insel, 1974; Moos, 1989; Moos, 1994a; Moos, 1994b).

Autonomy. The mean score for this subscale is 7.3, which is considered to be well above average. The median and modal scores are each 8, while its CV is 22%, again in line with the Involvement score subscale relative variation.

This subscale measures the extent by which employees' desire self-governance. Items on this subscale ask respondents about their preferred levels of empowerment in decision-making, initiative, innovation, and independence. The mean score for this subscale seems to indicate that these respondents would prefer to have a significant amount of freedom and ability to make their own decisions regarding their work.

Task Orientation. The mean score for this subscale is 7.8, which is again well above average. The median and modal scores are 8 and 9, respectively. The subscale's CV is 19%, indicating a stronger cohesion (less relative variation) within the respondents than the previous subscales. This subscale measures the levels of emphasis placed on efficiency, focus, and task completion. The mean score for this subscale indicates that the respondents would be focused on accomplishing tasks in a timely manner.

Work Pressure. The mean score for this subscale is 3.6, which is below average. Of the ten subscales, this average is the lowest. In concert with this ranking, its median and modal scores are 3 and 4, representing a relative direction away from the scores of the other nine subscales. Further, its CV is the largest at 57%, indicating considerable scoring diversity within the respondent group. Scores ranged from 0 to 9.

This subscale assesses employees preferred levels of urgency and pressures that exist in the work environment. The mean score for this subscale indicates that this generation would prefer to work in a relaxed environment, free of most job pressures.

System Maintenance and System Change Dimensions

This third set of WES dimensions assesses the preferred clarity of expectations, enforcement of rules, opportunities for innovation, and comfort in the work environment. The four subscales in this dimension are Clarity, Control, Innovation, and Physical Comfort (Moos & Insel, 1974; Moos, 1989; Moos, 1994a; Moos, 1994b).

Clarity. The mean score for this subscale is 7.0, which is considerably above average. The median and modal scores are 7 and 8, respectively, with a CV of 20%. Whereas the median and modal scores for this subscale are lower than most of the previous subscales, its relative variation is commensurate with the previous values.

This subscale assesses the extent to which corporate rules, regulations, and job expectations need to be defined. The mean score for this subscale indicates that the respondents would prefer to have significant details regarding the expectations of their job and work environment.

Control. The mean score for this subscale is 5.4, which is at the high end of *average* (on this subscale, 5.5-6.0 is considered as *above average*). The median and modal scores are 6 and 7, respectively. While these scores are lower than the previous subscales, a CV of 39% (second highest) indicates a fair degree of score variation within the respondent group.

This subscale assesses the firmness with which management enforces rules and policies. The mean score for this subscale indicates that while this generation would prefer to work in an environment that enforces rules, but it does not want an overly restrictive management.

Innovation. The mean score for this subscale is 6.4, which is considered to be well above average. The median and modal scores are each 7, while its CV is 34%, similar in value and interpretation as the prior *Control* subscale.

This subscale measures the extent that employees are encouraged to be creative in developing new methodologies and approaches, and are allowed to test new ideas. The mean score for this subscale indicates that the respondents strongly prefer a workplace that considers personal initiative and creativity to be of value.

Physical Comfort. The mean score for this subscale is 7.1, which is again considerably above average. The median and modal scores are each 8, placing this subscale equivalent to the first few subscales presented above. Similarly, its CV is 23%, indicating score variation equivalency to the first subscales.

This subscale measures the characteristics of the workplace that influence feelings of contentment and positive well-being. The mean score for this subscale indicates that this generation holds a strong preference for an environment that is pleasing and comfortable.

Chi-Square Analyses

Several chi-square tests were performed to determine which of the subscales, if any, indicates dominance or weakness relative to the remaining subscales. First, the subscales were tested within their respective dimensions, and then overall across all dimensions.

When examining the three subscales within the Relationship dimension, none were found to score significantly differently than any other, χ^2 (2, N = 6505) = 4.00, $p > 0.136$. This indicates that the respondents view the Involvement, Coworker Cohesion, and Supervisor Support subscales equally.

An examination of the three subscales within the Personal Growth and Goal Orientation dimension reveals a highly significant weakness for the Work Pressure subscale over the Autonomy and Task Orientation subscales, χ^2 (2, N = 5444) = 484.6, $p << 0.0001$. The contribution for the Work Pressure subscale represents roughly two-thirds the total chi-square statistic, indicating a considerable dislike of workplace pressures by this group. Factoring out the Work Pressure subscale reveals a less significant difference between the remaining two subscales with greater preference being given to Task Orientation than Autonomy, χ^2 (1, N = 4390) = 5.263, $p < 0.022$.

An examination of the four System Maintenance and System Change subscales reveals a significant *weakness* for the Managerial Control subscale, which accounts for roughly 60% of the overall chi-square statistic, χ^2 (3, N = 7516) = 79.533, $p << 0.001$. By "weakness," we mean that the respondents as a whole scored this factor much lower than expected, thereby contributing a large chi-square statistic value. Since the respondents seemingly do not favor Managerial Control, we removed it and reanalyzed the remaining three subscale factors. Consequently, the Innovation subscale is found to be weakly regarded while the Physical Comfort demonstrates some dominance with the Clarity subscale scores about as expected, χ^2 (2, N = 5938) = 14.470, $p < 0.001$.

Finally, all ten of the subscales were compared against one another. As is indicated by the descriptive results above, the Work Pressure subscale is very weakly regarded (again, "weakly" is in terms as we describe above), accounting for over 63% of the total chi-square statistic, χ^2 (9, N = 19,464) = 644.8, $p << 0.0001$. Managerial Control assumes a distant second contributing only 11% of the overall chi-square statistic. Nonetheless, these two subscales indicate a considerable dislike for these factors by this Millennial group. In contrast, the Clarity and Innovation subscales appear to be moderately regarded whereas the remaining six subscales appear to be favored on a somewhat equal footing.

Discussion

The purpose of this study was to determine the ideal work environment preferences of members of the Millennial generation currently in the aviation industry.

The results of this study portray a complex combination of the relationship, personal growth, and organizational structure of the ideal workplace as indicated by these respondents.

Overall, results indicate that these respondents view themselves as being highly committed to their jobs, and prefer a workplace environment that is very supportive and cohesive. It appears that this generation would work best in an inclusive environment, where managers utilize a more engaged approach incorporating coaching and mentoring strategies instead of authoritative directives. Millennials prefer being treated as partners and may work best in organizations with flattened hierarchies (Earle, 2003).

Unlike many other businesses, aviation is an extremely fast-paced industry requiring many time-sensitive decisions. While Millennials want the expectations of the workplace and of their job functions communicated in explicit detail, they do not want to be micro-managed. Respondents in this study indicated that they are focused on getting their work completed in a timely manner; however, they would also like the freedom to be creative and innovative. Though this may not be appropriate in many facets of the industry, inventive ideas could actually prove to be beneficial. Novel thinking that provides new products, methodologies, and ways of doing business could advance aviation performance worldwide. Management may wish to cultivate this ingenuity and provide Millennial workers with the opportunity to utilize their creativity in looking at old problems in new ways.

Aviation is very unique in that is operationally structured on a vast array of rules, regulations, and time schedules. Without these, the industry would not be able to function as effectively, efficiently, and safely as it does. Establishment and enforcement of these requirements comes not just from management, but also from local, state, and federal governments. Although the respondents of this study indicated they prefer a work environment in which they are informed in great detail of the rules, they also want flexibility in their decision-making. This establishes an apparent conflict that aviation managers may have to confront. Previous research has found that providing members of this generation the background of why particular rules exists tends to expand their understanding and acceptance of the particular policy (Martin & Tulgan, 2006). The aviation industry could experience dire effects if the uniformity and consistency provided by rules, regulations, and schedules are not upheld. Haphazard disregard of this structure could prove disastrous; therefore, more time should be spent explaining the rationale for the rules and regulations or, perhaps by providing employees with a company website to access resources which include readily available explanations.

The aviation industry is quite dynamic and is impacted by diverse factors ranging from politics to weather to economics, to name a few. The compounding interaction of these and other issues complicated by the continuous struggle to provide good, safe products and services, while also striving for financial success, causes aviation personnel to continually feel a sense of urgency in their work. The respondents in this study expressed a strong desire to work in an environment free from such job pressures. Again, this preference is in strong opposition to the nature of the industry.

As with any group, the Millennial generation exists across a continuum of likes and dislikes. This is supported by the relatively large variances in preferences indicated on the *Work Pressure* and *Control* subscales (57% and 39%, respectively). Consequently, it is incorrect to conclude that all Millennial generation respondents in this study shy away from work environments in which continual demands exist. By contrast, some respondents indicated no particular discomfort with such pressures. Because of this, it may be prudent for management to appropriately screen candidates for various aviation positions.

Respondents in this investigation stated that they preferred a physical workspace that was pleasing and comfortable. Because of their concurrent desire to work in partnership with others, it may be beneficial to redesign corporate workspaces to encourage collaboration. For many employees, an organization that provides them with a workplace environment that makes them feel energized and valued can encourage them to be more productive and perhaps work longer (Earle, 2003). In a previous study focusing on workplace design, employees indicated they would work an extra hour a day and felt their companies would be more competitive if they developed a better environment in which to work (Pfeffer, 2007). For members of the millennial generation, not only would a comfortable workspace be more conducive to their preferences, but it may be advantageous for management to also provide them with mobile technologies that would allow them to work anytime, anywhere.

Results of this study tend to agree with previous investigations focusing on the characteristics of the Millennial generation. Based on these findings, it is apparent that this generation has very distinct preferences for their ideal workplace and strong expectations of their employers. They are a generation that knows what they want and are used to getting it (Epstein & Howes, 2006; Martin & Tulgan, 2002). The arrival of this generation into the aviation workplace may present some challenges, but it also provides many opportunities. Success will be achieved by those organizations that are aware of the Millennial generation's workplace preferences.

Limitations

Potential limitations of this study may be the relatively small sample size; therefore, generalizability of the results may be restricted. This study also relied upon self-report surveys to assess the participants' work environment preferences. Consideration should be given to utilization of other data collection measures, such as interviewing respondents, as well as collecting data from managers and supervisors of this cohort.

This study is the first in a series focusing on the implications of the Millennial generation working in the aviation industry, and was meant to be exploratory in nature. Future studies will include analyses of differences between generations currently at work in the aviation industry, as well as investigating generational preferences of individuals within functional areas.

Conclusion

The primary contribution of this investigation was the determination of the ideal workplace preferences of the Millennial generation currently in the aviation industry. Overall, results indicate that these respondents have very strong and distinct preferences for their ideal workplace. For aviation managers, it is important to understand the uniqueness of this generation so as to allow for a smooth assimilation of these workers into the workplace.

References

Baker, G., Carlisle, C., Riley, M., Tapper, J., & Dewey, M. (1992). The work environment scale: A comparison of British and North American nurses. *Journal of Advanced Nursing*, 17, 692-698.

Epstein, M., & Howes, P. (2006). The millennial generation: Recruiting, retaining, and managing. *Today's CPA*, September/October, 24-26.

Glass, A. (2007). Understanding generational differences for competitive success. *Industrial and Commercial Training*, *39*(2), 98-103.

Gravett, L., & Throckmorton, R. (2007). Bridging the generation gap: How to get radio babies, boomers, gen xers, and gen yers to work together and achieve more. Franklin Lakes, NJ: The Career Press, Inc.

Johnson, B., & Wilson, S. (2008). Responding to generational differences in business education: Challenges and opportunities. *Business Education Forum, 62*(3), 40-44.

Kotzer, A.M., Koepping, D.M., & LeDuc, K. (2006). Perceived nursing work environment of acute care pediatric nurses. *Pediatric Nursing, 32*(4), 327-332.

Lancaster, L., & Stillman, D. (2002). *When generations collide: Who they are. Why they clash. How to solve the generational puzzle at work*. New York, NY: HarperCollins Publishers.

Long, C., Williams, M., & Hollin, C. (1995). Staff perceptions of organization change of treatment delivery on a addition unit. *Journal of Advanced Nursing*, 21, 759-765.

Martin, C., & Tulgan, B. (2002). *Managing the generation mix: From collision to collaboration.* Amherst, MA: HRD Press, Inc.

Martin, C., & Tulgan, B. (2006). *Managing the generation mix: From collision to collaboration.* 2nd Ed. Amherst, MA: HRD Press, Inc.

McCrae, N., Prior, S., Silverman, M, & Banerjee, S. (2007). Workplace satisfaction in a mental health service for older adults: An analysis of the effects of setting and professional status. *Archives of Psychiatric Nursing, 21*(1), 17-24.

McGlynn, A.P. (2005). Teaching millenials, our newest cultural cohort. *The Education Digest, 71*(4), 12-16.

Mitchell, S. (1995). *The official guide to the generations*. 1st Ed. Ithaca, NY: New Strategist Publications, Inc.

Moos, R.H. (1989). *Work environment scale – form r interpretive report form*. Palo Alto, CA: Consulting Psychologists Press, Inc.

Moos, R. H. (1994a). *The social climate scales: A user's guide*. 2nd Ed. Palo Alto, CA: Consulting Psychologists Press, Inc.

Moos, R. H. (1994b). *Work environment scale manual.* 3rd Ed. Palo Alto, CA: Consulting Psychologists Press, Inc.

Moos, R. H., & Insel, P.N. (1974). *Work environment scale form i*. Mountain View, CA: Consulting Psychologists Press, Inc.

Pfeffer, J. (2007). Thinking outside the cube. Business 2.0. Retrieved November 29, 2009 from, http://www.lexisnexis.com.ezproxy1.lib.asu.edu/us/lnacademic/search/omesubmitForm.do

Raines, C. (2002). Managing millennials. Retrieved April 13, 2009 http://www.generationsatwork.com/articles/millenials.htm

Ryan, R. (2007). *Live first, work second*. Madison, WI: Next Generation Consulting, Inc.

Schuman, H., & Scott, J. (1989). Generations and collective memories. *American Sociological Review, 54*(3), 359-381.

Smola, K.W., and Sutton, C.D. (2002) Generational differences: Revisiting generational work values for the new millennium. *Journal of Organizational Behavior*, *23*(4), 363 -382.

Strauss, W., & Howe, N. (1997). *The fourth turning*. New York, NY; Broadway Books.

Teaching the millennials (2007, March 20). Center for Digital Education. Retrieved September 16, 2008, from http://www.centerdigitaled.com/story.php?id=104514.

US Census Bureau. (2009). Census 2000 summary file 1. [Data file]. Available from http://factfinder.census.gov/servlet/DTTable?_bm=y&-geo_id=01000US&-ds_name=DEC_2000_SF1_U&-_lang=en&-redoLog=false&-mt_name=DEC_2000_SF1_U_P012&-format=&-CONTEXT=dt

Westerman, J.W. & Cyr, L. (2004). An integrative analysis of person-organization fit theories. *International Journal of Selection and Assessment 12*(3), 252-261.

Westerman, J.W. & Simmons, B. (2007). The effects of work environment on the personality-performance relationship: An exploratory study. *Journal of Managerial Issues, 19*(2), 288 – 303.

Westerman, J.W., & Yamamura, J. H. (2007). Generational preferences for work environment fit: Effects on employee outcomes. *Career Development International, 12*(2), 150-161.

Zemke, R., Raines, C., & Filipczak, B. (1999). *Generations at work: Managing the clash of veterans, boomers, xers, and nexters in your workplace*. New York: AMACOM, 1999. Electronic reproduction. Boulder, CO: NetLibrary, 1999.

International Journal of Applied Aviation Studies, Volume 9, Number 2

A Review of General Aviation Accidents in Pacific Ocean Operations

AJ de Voogt

American Museum of Natural History
New York, NY
USA

B Heijnen

Maastricht University
Maastricht
The Netherlands

Abstract

In general aviation, flight emergencies above the open ocean occur for both helicopters and airplanes. They are particularly damaging due to the absence of a landing area when making an emergency landing at sea. Engine failure appears to be the primary cause of open ocean accidents for both single engine and twin engine aircraft. There is no significant relation between injury and aircraft damage in Pacific Ocean operations, partly because controlled landings on water damage the aircraft but are rarely fatal. Helicopter operations above open water result in significant increases in severe or fatal injuries when compared to airplanes. All general aviation accidents with small aircraft occurring above the Pacific Ocean and reported by the National Transport Safety Board were analyzed and the differences between helicopter and airplane operations were highlighted. The results of the analysis indicate that the main preventive measure to mitigate general aviation emergencies above the open ocean is the elimination of engine failures that result from fuel exhaustion due to navigation or fuel management errors.

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.

Return to Table of Contents

Background

When General Aviation (GA) aircraft cross the open ocean, their possibilities for a safe emergency landing are limited. Still, ferry flights and observation flights are regularly conducted in these circumstances (Sacchie, 1979).

Strategies for reducing the number of General Aviation accidents have received increasing attention. In 2007, Li and Baker reported 228,000 active private pilots and 220,000 registered General Aviation aircraft in the United States. In the years prior, 91% of all aviation crashes and 94% of all aviation fatalities were in GA. This striking difference between General Aviation and major airline operations has been in place for many years. Yet accident investigation models with the help of taxonomies (Shappell & Wiegmann, 1997) and formal causal analysis (Ladkin, 2000) were not designed to address this difference since GA is both too diverse and its accident reports too limited in content. The organizations behind each GA operation are limited to few people, and flight data are rarely recorded that may allow an analysis of the communication and actions prior to an accident.

Instead, the mitigation of General Aviation accidents is assisted by aircraft-specific and operation-specific analyses, such as those for helicopter emergency flights (Baker et al., 2006) or for homebuilt aircraft (Hasselquist & Baker, 1999). Flights above the open ocean are conducted by both helicopters and airplanes, and the characteristics of ocean operations require specific accident mitigation strategies.

Emergency landings by large non-GA aircraft (above 12,000 lbs) on the open water are rare, but when they do happen, they are rarely without fatal consequences. In contrast, the smaller general aviation aircraft frequently survive landings at sea or have taken measures that limit the necessity for such a landing when they are required to operate above open water. It is this set of measures and the characteristics of the accidents that are of interest for both the pilots involved in regular ocean operations and for pilots occasionally traversing expanses of open water.

Each ocean has its particular general aviation operations. The Gulf of Mexico mainly features helicopter air-taxi operations to oilrigs, while the Atlantic Ocean has hardly any regular helicopter operations present. The Pacific and the Atlantic Ocean both feature ferry flights with general aviation airplanes, but the former has also seen an increasing number of helicopter flights operating from fishing vessels in recent years. The accidents occurring in the Pacific Ocean allow an exploratory comparison of helicopter and airplane accidents above open water.

The possibilities for mitigating the number and the severity of these accidents were explored with an analysis of 67 accidents recorded for the Pacific Ocean region with aircraft registered in the United States.

Method

Data for the study were extracted from the National Transportation Safety Board (NTSB) online database, which contains probable cause statement and factual reports of Civil Aviation accidents and incidents since 1962 (NTSB, 2008). Accidents are reported to the database if aircraft sustained substantial damage or were destroyed and/or if at least one of the occupants or ground crew suffered severe or fatal injuries. The database query included all accidents occurring in the Pacific Ocean region for aircraft lighter than 12,000 lbs and part of general aviation operations. Since denominator data on Pacific Ocean flights are not available, no information about the risk of helicopter and airplane flights can be provided.

Probable cause statements provide a summary of the accident as well as the causes and contributing factors as determined by the NTSB examiner. The factual reports contain a narrative statement based on interviews with witnesses and pilot(s) as well as details about the aircraft, the pilot and the weather circumstances. The data provided for accidents occurring before 1982 are limited in their description. For example, weather and personal data of the pilot as well as technical data of the aircraft may be missing. Despite these limitations, the reports provide sufficient data for an analysis. Categorical data, manifested in numbers of observations (counts) in different categories reported by the NTSB investigator, were related using Pearson's Chi Square analysis at 95% statistical reliability.

Since the total number of flights with general aviation aircraft above the Pacific Ocean is unknown, it is not possible to present a risk analysis in terms of accidents per number of flight hours. The dangers of Pacific Ocean operations are found in the proportionate number of fatal accidents in the dataset. Also, when aircraft disappear in the Pacific Ocean it is often not possible to determine the cause or even the occurrence preceding the crash and this limits the dataset.

Results

From 1964 until 2004, 67 accidents occurred in the Pacific Ocean with general aviation aircraft involving 28 helicopters and 39 airplanes. Airplane accidents averaged one accident per year throughout this time frame. Helicopters did not appear in these accident statistics until 1978, and averaged two accidents per year from 1991 onwards.

Injuries and damage

The number of accidents with severe or fatal injuries versus those with only minor or no injuries was significantly smaller for airplanes compared to helicopters ($\chi^2 = 4.0$, df = 1, $p < 0.05$) (see Table 1). Twelve helicopter accidents included fatalities, with a total of 17 occupants killed; 14 occupants were killed in 13 fatal airplane accidents.

Table 1

Damage and injury patterns per aircraft type.

	total number of accidents	substantial damage	destroyed	severe injury	fatal injury
helicopter	28 (100%)	8 (29%)	19 (68%)	8 (29%)	12 (43%)
airplane	39 (100%)	4 (10%)	35 (90%)	4 (10%)	13 (33%)

There was no significant relationship between the damage of the aircraft and the severity of the injury sustained by the occupants ($\chi^2 = 1.02$, df = 1, $p > 0.3$). See table 2.

Table 2

Relation of damage and injury

	none/minor/ substantial damage	destroyed
none/minor injury	8	25
severe/fatal injury	4	30

Aircraft engines

One helicopter with twin-engines was involved in an accident, while 17 out of 39 airplanes had two engines reported. There was no significant difference between the number of fatal accidents with airplanes with or without twin engines ($p > 0.8$). There were four substantially damaged airplanes with a single engine and none with twin engines compared to the rest that were destroyed, but this difference was also not significant.

Pilot characteristics

The average age of the general aviation pilots in this dataset was 39.4 years, with 12 pilots younger than 30, 20 pilots between 30 and 40, 22 pilots between 40 and 50 and 11 pilots who were 50 years or older of which the oldest were 60 and 66. The age of two pilots was unrecorded. Only 11 pilots had less than 1000 hours flight hours, 25 (37%) pilots had between 1,000 and 5,000 hours, and 21 (31%) pilots had over 5,000 hours of flight experience in either helicopter or airplane.

Environmental conditions

Weather and daylight conditions were not consistently reported, but for the period after 1982 three accidents took place at night and one at dusk. The helicopter flight at dusk also reported fog and one night flight with an airplane occurred during rainy conditions. Two other day flights were conducted in rain and one in hazy weather. No other special weather conditions were reported.

Activities during accidents

Twenty-one (75%) helicopters involved in Pacific Ocean accidents were conducting aerial observations, two conducted other commercial operations, one conducted a ferry flight and four were unrecorded or had ‘other work use’ reported. In contrast, 22 (56%) airplanes were conducting ferry flights, 11 conducted personal or pleasure flights, two conducted instructional flights and four had noncommercial or unreported activities.

Six helicopter accidents took place on the deck of a ship. About 25% of airplanes landed on water, also known as ditching, and a similar percentage was found for helicopters. The remaining aircraft crashed in the water, meaning that no landing was attempted or intended according to the NTSB investigator. The crash landing was fatal in 21 out of 47 cases, a ratio significantly different from the controlled landing or ditching in which only 1 out of 14 cases was fatal ($\chi^2 = 5.06$, df = 1, $p < 0.03$).

Helicopters that were part of Pacific Ocean accidents reported seven (25%) airframe/component failures that included one rotor failure and one engine failure. Five (18%) losses of engine power were reported in addition to seven (25%) accidents traced to a loss of control in flight. In comparison, the airplane occurrences included 15 (38%) engine failures, one airframe failure, one (3%) loss of control in flight, and ten (26%) losses of engine power. In short, there was a significant difference between helicopters and airplanes in that only one engine failure was reported for helicopters and 25 engine (power) failures were reported for airplanes ($\chi^2 = 22.66$, df = 1, $p < 0.01$).

Causes of accident

In at least ten (36%) cases a helicopter pilot was the cause of the accident, and in at least 16 (40%) cases it was an airplane pilot. In nine (32%) cases the cause of the helicopter accident could not be determined, and this was also the case for 17 (43%) airplane accidents. In all remaining cases, the accident was caused by maintenance or by other personnel or the cause was not attributed to anyone, such as the occurrence of a bird strike. Accidents with multiple causes were only found in part of the data and secondary causes identified by the NTSB examiner were not included in the analysis.

Discussion

The proportionate number of fatal accidents and destroyed aircraft in flights above the Pacific Ocean is higher than the number for general aviation flights elsewhere (Li & Baker, 1999). The information provided by the NTSB accident reports is, however, severely limited. Both the absence of data on the organization that oversees a particular flight as well as the lack of flight data in the minutes prior to the accident prevent detailed causal analyses. The reports by the NTSB also changed over time so that some information could not be collected consistently for all accidents in the dataset. Information on incidents and comparisons with ocean operations elsewhere may contribute to future studies on this subject.

In Pacific Ocean general aviation accidents, the severity of the injuries is not related to the damage of the aircraft. In other words, a destroyed aircraft does not always lead to fatal or severe injuries. In ocean operations, the relation was not inverse, as has only been reported for ballooning (de Voogt & van Doorn, 2006), but unlike general aviation as a whole, the pilots survived most forced landings. Yet the water may still destroy the aircraft after a successful landing is completed. While the ocean rarely provides an ideal landing area, it also does not have obstacles that could harm the passenger and aircraft before it reaches sea level, with the occasional exception of a ship. These characteristics of ocean landings partly explain the absence of the relation between occupant injury and aircraft damage.

A notable difference was found between helicopter accidents and airplane accidents in Pacific Ocean operations. This difference should for an important part be attributed to the type of operation, although other relatively risky helicopter operations, such as sling loads and aerial application, have a lower proportionate number of fatalities in their dataset (de Voogt, Uitdewilligen & Eremenko, 2008; de Voogt & van Doorn, 2007). Operations from ships provide helicopters with at least one safe landing spot, but at the same time accidents on a ship create different kinds of damage and injury. Helicopter operations on ships were rare in the NTSB database, which only contains Civil Aviation accidents, and no significant number of such accidents is available to allow a separate study of this type of operation. Still, if landings on ship decks are excluded from the analysis, the difference between helicopters and fixed-wing aircraft accidents remains, suggesting an increased danger of helicopters above open sea.

In comparison with helicopters, airplanes suffered a higher percentage of engine failures or loss of engine power during ocean flights. At the same time, airplanes were more commonly equipped with more than one engine as evidenced by the study sample. The main cause of an accident for multi-engine aircraft was nonetheless engine failure. This indicates that despite multiple engines, a higher percentage of airplanes involved in an accident suffered from engine trouble when compared to helicopters. At least eight engine failures were caused by the pilot according to the NTSB investigator; while eleven had undetermined causes.

Discipline in the application of regulations, in particular for ship deck operations, is a general recommendation that applies to all ocean flights. Rotor strikes and lines attached to the helicopter during take-off are clear indications that the few accidents taking place on ship decks could also have been prevented with better oversight. The same holds true for engine failures due to incorrect fuel management that affected both helicopters and fixed-wing aircraft.

Controlled landings on water in this dataset were, with one exception, not fatal, which means that the absence of a landing place was not the main risk for Pacific Ocean accidents. At best, it is the perceived absence of a landing area rather than a controlled landing on the water surface that resulted in an accident. Due to their freqency in the dataset, the main preventive measure to mitigate accidents above open water remains the elimination of engine failures that can result from fuel exhaustion due to navigation or fuel management errors.

References

Baker S. P., Grabowski J.G., Dodd R. S., Shanahan D. F., Lamb M. W., & Li G. (2006). EMS helicopter crashes: What influences fatal outcome? *Annals of Emergency Medicine, 47*(4), 351–356.

NTSB (2008). Accident Database & Synopses. Retrieved from http://www.ntsb.gov/ntsb/query.asp on May 3, 2008.

Hasselquist A., & Baker S. (1999). Homebuilt Aircraft Crashes. *Aviation, Space, and Environmental Medicine 70*, 543-547.

Ladkin, P. B. (2000). Computer Safety, Reliability, and Security. *Proceedings of the 19th International Conference, SAFECOMP 2000, Lecture Notes in Computer Science No. 1943.* Springer-Verlag.

Li G., & Baker S. P. (1999). Correlates of pilot fatality in general aviation crashes. *Aviation, Space, and Environmental Medicine, 70***,** 305-309.

Li G., & Baker S. P. (2007). Crash risk in general aviation (Editorial). *Journal of the American Medical Association, 297*, 1596-1598.

Sacchie L. (1979). *Ocean flying: A pilot's guide*. New York: McGraw-Hill.

Shappell S. A., Wiegmann D. A. (1997). A human error approach to accident investigation: The taxonomy of unsafe operations. *International Journal of Aviation Psychology, 7*(4), *269–291.*

de Voogt A. J., van Doorn R. R. A. (2006). Balloon crash damage and injuries: An analysis of 86 accidents, 2000–2004. *Aviation, Space, and Environmental Medicine, 77*(5), 556–558.

de Voogt A. J., Uitdewilligen S., Eremenko N. (2008). Safety in high-risk helicopter operations: The role of added crew. *Safety Science, 47,* 717-721.

de Voogt A. J., van Doorn R. A. (2007) Helicopter accident analysis: datamining the NTSB database. *Proceedings of the European Rotorcraft Forum, Kazan, Russia, 2007, September 11–13.*