International Journal of Applied Aviation Studies

A publication of
FAA Academy
Oklahoma City, Oklahoma

REVIEW PROCESS

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

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

International Journal of Applied Aviation Studies
c/o Beverly Laughead
AMA-530
PO Box 25082
Oklahoma City, OK 73125

International Journal of Applied Aviation Studies
Volume 4, Number 2
ISSN Number: 1546-3214

1st Printing January 2005

POLICY AND DISCLAIMERS

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

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

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

PRODUCTION MANAGERS

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

EDITOR

Beverly Laughead	University of Oklahoma, Norman, OK

ASSOCIATE EDITORS

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

CONSULTING EDITORS

International

Robert Bor	Royal Free Hospital, London, UK
Finian Connolly	Executive & Professional Training Institute, Drogheda, Co Louth, Ireland
Hugh David	Eurocontrol, East Sussex, UK
Sidney Dekker	Linköping Institute of Technology, Linköping, Sweden
Alois Farthofer	Human Factors Research, St. Georgen, Austria
Gerard J. Fogarty	University of Southern Queensland, Toowoomba, Australia
Jarle Gimmestad	Braathens Airlines, Fornebu, Norway
Don Harris	Cranfield University, Cranfield, Bedfordshire, UK
Irene Henley	University of Western Sidney, Sidney,Australia
Graham Hunt	Massey University,Auckland, New Zealand
Magnus Jacobssen	Linköping Institute of Technology, Linköping, Sweden
K. Wolfgang Kallus	Karl Franzens University, Graz, Austria
Elizabeth Mathews	International Civil Aviation Organization, Montreal, Canada
Jens Rolfsen	Det Norske Veritas, Høvik, Norway
Claudia van de Wal	VALK Foundation, Netherlands
Lucas van Gerwen	VALK Foundation, Netherlands
Stephen Walsh	Interaction Training Associates, Netherlands
Rogier Woltjer	Linköping Institute of Technology, Linköping, Sweden
Beat Zimmermann	WX Systems, Thun, Switzerland

U.S.A.

R. Kurt Barnhart	Indiana State University,Terre Haute, IN
Ellen Bass	University of Virginia, Charlottesville, VA
Jeffrey Beaubien	American Institutes for Research, Washington, DC
Wendy Beckman	Parks College, St. Louis University, St. Louis, MO
James P. Bliss	Old Dominion University, Norfolk, VA
Deborah Boehm-Davis	George Mason University, Fairfax, VA
Brent Bowen	University of Nebraska at Omaha, Omaha, NE
John Cain	Florida Technology Institute, School of Aeronautics, Melbourne, FL
David Carmichael	Oklahoma City University, Oklahoma City, OK

CONSULTING EDITORS (con't.)

Stephen M. Casner	NASA Ames Research Center, Moffett Field, CA
David Conway	Southeastern Oklahoma State University, Durant, OK
Janet Cope	University of Oklahoma, Norman, OK
Key Dismukes	NASA Ames Research Center, Moffett Field, CA
Graham Elliott	FAA Academy, Oklahoma City, OK
Richard Fanjoy	Purdue University, West Lafayette, IN
Douglas R. Farrow	Federal Aviation Administration, Washington, DC
Kenneth H. Funk II	Oregon State University, Corvallis, OR
Trent Gabert	University of Oklahoma, Norman, OK
A. F. Grandt	Purdue University, West Lafayette, IN
Mavis Green	Middle Tennessee State University, Murfreesboro, TN
Jeffrey T. Hansberger	George Mason University, Fairfax, VA
James M. Hitt, II	Booz Allen Hamilton, McLean, VA
Judy Holcomb	FAA Academy, Oklahoma City, OK
Todd P. Hubbard	University of Oklahoma, Norman, OK
Michael Hynes	Hynes Aviation Services, Frederick, OK
Raymond E. King	Civil Aerospace Medical Institute, Oklahoma City, OK
Jefferson M. Koonce	IFR Flight Training School, Georgetown, TX
Mary N. Kutz	Oklahoma State University, Stillwater, OK
Florian Jentsch	University of Central Florida, Orlando, FL
Victor LaSaxon	The Boeing Company, Midwest City, OK
Rebecca Lutte	University of Nebraska at Omaha, Omaha, NE
Beth Lyall	Research Integrations, Inc., Tempe, AZ
Carol Manning	Civil Aerospace Medical Institute, Oklahoma City, OK
Harry Marks	University of Oklahoma, Norman, OK
Steven Marks	Oklahoma State University, Stillwater, OK
Royce Ann Martin	Bowling Green State University, Bowling Green, OH
Celeste Mayer	North Carolina State University, Raleigh, NC
H. C. McClure	MAC Air Consulting, Queen Creek, AZ
Jesse McMullen	FAA Academy, Oklahoma City, OK
Marjo Mitsutomi	University of Redlands, Redlands, CA
Richard H. Mogford	NASA Ames Research Center, Moffett Field, CA
John C. Morey	Dynamics Research Corporation, Andover, MA
Kathleen Mosier	San Francisco State University, San Francisco, CA
Gary J. Northam	Parks College, St. Louis University, St. Louis, MO
Manoj Patankar	Parks College, St. Louis University, St. Louis, MO
Mike Pearce	American Airlines
Donald A. Petrin	Purdue University, West Lafayette, IN
Julia Pounds	Civil Aerospace Medical Institute, Oklahoma City, OK
Stephen M. Quilty	Bowling Green State University, Bowling Green, OH
Esa M. Rantanen	University of Illinois, Urbana-Champaign, IL
Gail Reed	Pan Am International Flight Academy, Miami, FL
Robert Ripley	Auburn University,Auburn, AL
Art Saboski	Embry-Riddle Aeronautical University, Prescott, AZ
Nadine Sarter	The Ohio State University, Columbus, OH
Robert Schlegel	University of Oklahoma, Norman, OK
Gregory L. Schwab	Indiana State University, Terre Haute, IN
John W. Sheremeta, Jr.	Embry-Riddle Aeronautical University, Daytona Beach, FL

CONSULTING EDITORS (con't.)

PHILOSOPHY STATEMENT

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

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

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

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

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

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

B.S.L.

---

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

EDITOR'S NOTES

Formal Papers

Readers who are interested in the improvement of pilot vision under adverse weather may find this extensive report of studies conducted by Korn, Lorenz, Doehler, Többen and Hecker to be of interest. Built on three technological elements, the studies were conducted to examine the feasibility of a "Radar-PAPI" concept that may act in low visibility similar to that of the visual PAPI landing aids in good visibility.

Running third to safety and security, airline service is very important to most air travelers. Based on revisiting and reanalyzing four previous reports regarding Airline Quality Rating (AQR) as an objective mechanism of measuring airline service performance, Bowen, Headley, and Lu argue that an airline's operational deficiencies can be retrieved, while simultaneously opening a window for public review and service improvement.

Considering the potential danger of operating electronic devices on board aircraft, many readers will find the next article an interesting report. The article discusses a study conducted by Balfe and Head to establish the numbers and types of devices carried on board by passengers and the level of use during the flight, to determine passenger knowledge of hazards associated with operation on board and regulations governing operation, to determine passenger perceptions of level of risk and to determine level of passenger desire to operate devices on board.

Casner's article reports on a study conducted to determine how much practice is needed by pilots, who desire to fly IFR with GPS. With sixteen instrument-rated pilots who were recruited from local professional flight training schools, data gathered during the study were analyzed to determine: (1) whether or not the ground study and five practice flights were enough to allow pilots to master the skills; (2) how effective was self-study compared to dual instruction; and (3) which skills presented pilots with the most difficulty and accounted for the most errors.

Those interested in aircrew performance and aviation safety may find this next article to be of interest. As a continuation of ongoing research in pilot resource selection and based on the Big Five Personality Inventory, the Leadership Intelligence Quotient Self-Assessment and three select attribute scales, Hamilton and Ripley examined the effectiveness of binary logistic regression in screening candidates for instructor pilot upgrade.

The next article is another for those interested in the area of safety. Similar to prior studies to develop POD curves for detecting cracks in metal structural components of commercial aircraft, Erhart, Ostrom, and Wilhelmsen conducted a pilot study to determine what size dent the average person would detect.

Since previous research in the field of judgment and decision-making has been divided into two theoretical perspectives, researchers interested in the relevance of judgment and decision-making in the application to aviation may find this next article interesting. Using the proposal that pilot judgment and decision making can and should incorporate both theories, Jacobson and Mosier examined pilots' use of coherence- and correspondence-based decision-making processes as a function of variables such as phase of flight, weather conditions, type of event, and level of aircraft automation.

Based on audits, compliance with International Civil Aviation Organization (ICAO) requirements by many civil aviation authorities around the world is not adequate and often not sustained. Motevalli, Ansari, Lagos, and Novak discuss the role of oversight in the aviation industry worldwide, examine the critical elements of an effective civil aviation oversight authority, and discuss the key factors for sustaining such authority.

Readers who are interested in learning theories may find the Rhind and Head article to be of interest. Rhind and Head conducted a study to determine whether implicit learners perform under test conditions as well as explicit learners.

Training Development Reports, Studies, and Papers

The authors of the next article discuss and critique a clinical model that would advance the aviation education discipline at comprehensive universities. Marks and Vitek propose that flight centers be established as clinics, similar to those in the health profession. These centers would utilize instructors to accomplish the significant roles of flight training and minimize the involvement of research faculty.

Acknowledging that this study can be generalized only to settings highly similar to that at Indiana State University, Schwab conducted a study to determine whether the quality of service provided by the two contracted ISU flight schools was adequate to meet the needs of the flight school students. In hopes that this study will serve as a model for other aerospace departments to become more aware of, monitor, and make the necessary improvements in their flight training programs to deliver the safest and most high-quality flight training to their students, Schwab administered a survey to all 66 students enrolled in the ISU flight school program.

Aviation education and training instructors and developers may find this next article to be of interest. Robertson discusses the methods and strategies for teaching higher-order thinking skills in aviation and the learning theory supporting these methods. He argues that improving higher-order thinking skills will lead to better judgment, decision-making, and critical thinking; hence, fewer aviation accidents.

The Kutz, Brown, Carmichael, and Shandiz article may appeal to those who are interested in teaching/learning theories. The authors discuss research regarding the Myers-Briggs Type Indicator (MBTI) preferences of college students, specifically aviation majors and business majors, and compare the preferences of aviation Professional Pilot undergraduate majors with the preferences of Aviation Management undergraduate majors.

The Haley-Seikel and Hamm review of Human Factors in the Training of Pilots by Jefferson M. Koonce provides readers with a look at a "textbook" that is well worth reading and placing in each aviation enthusiast's library.

B.S.L.

International Journal of Applied Aviation Studies

Volume 4, Number 2 2004

Formal Papers

Training Development Papers

Book Review

International Journal of Applied Aviation Studies, 4, Number 2

Formal Papers

Passive "Radar-PAPI" Landing Aids for Precision Straight-in Approach and Landing in Low Visibility

Bernd Korn, Bernd Lorenz, Hans-Ullrich Doehler,
Helmut Többen and Peter Hecker

Institute of Flight Guidance
German Aerospace Center (DLR)
Lilienthalplatz 7
D-38108 Braunschweig
Deutschland

Abstract

Feasibility of a "Radar-PAPI" concept is examined that may act in low visibility like the visual PAPI landing aids in good visibility. This concept is built on three technological elements. First, the aircraft must be equipped with a forward-looking millimeter wave radar. Second, runway-marking retro-reflectors must be placed along the runway. Finally, another set of reflectors is needed to generate information on glide slope deviations. In a first field study, flight trials were completed to sample radar images. A subsequent low-fidelity part-task simulation was performed to answer the question as to how efficiently these kind of radar images can be utilized by pilots. Finally, an empirical study examined different options as to how to implement the Radar-PAPI concept in the cockpit. Results suggested that pilots learn the extraction of the visual cues from radar images surprisingly fast with the help of cue augmentation provided by retro-reflectors. Transformation of the usual plane-view display of the radar image to a three-dimensional perspective representation was associated with best pilot performance. In conclusion, 'Radar-PAPI' passive landing aids may complement other enhanced vision sensors, such as an infrared, to support low-visibility straight-in landings.

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

Introduction

Enhanced Vision Systems (EVSs) are currently developed to alleviate restrictions in airspace and airport capacity in low-visibility conditions. The demand for all-weather flight operations becomes most important for the approach and landing phase of a flight, when safety concerns resulting from low visibility require the nominal airport capacity to be significantly reduced. Consequently, there is a need to close the gap between capacity under low-visibility and good-visibility conditions. The EVS can help to fulfill this goal. The EVS relies on weather-penetrating forward-looking sensors that augment the naturally existing visual cues in the environment and provide a real-time image of prominent topographical objects that may be identified by the pilot. Infrared (IR) and millimeter-wave (MMW) sensors currently are envisaged as the most promising EVS support of pilot vision in low visibility. One important benefit of IR-sensors is that these sensors generate a perspective image, from which the human can derive the perceptual cues of depth to generate a three-dimensional interpretation of the outside world. This is an important feature of the IR-sensor, as such a perspective sensor image can be overlaid to the outside scene on a head-up display (HUD). The recently released rule of the FAA for Enhanced Flight Vision Systems (EFVSs) clearly acknowledges the operational benefits of such a technology by stating the following: "Use of an EFVS with a HUD may improve the level of safety by improving position awareness, providing visual cues to maintain a stabilized approach, and minimizing missed approach situations" (Federal Aviation Administration (FAA), 2004, p. 1621). Moreover, "The pilot would use this enhanced flight visibility … to continue the approach from DH [decision height] or MDA [minimum descent altitude] down to 100 ft above the touchdown zone elevation of the runway of intended landing" (FAA, 2004, p. 1621).

This rule change marks a significant token of confidence towards EVS technology. However, the penetration of bad weather (dense fog and light rain) in the infrared spectrum is remarkably poorer than the weather penetration that can be achieved by MMW-radar (Currie & Brown, 1987). Horne, et al. (1992) proposed a rather convincing method to define the all-weather characteristic of a sensor. They measured the reflected radar energy from the runway and the surrounding grass as a function of distance and various weather conditions. The ratio of the reflected energy received by the radar antenna between grass and runway was used as a measure for the contrast and they defined a 3 dB difference as a threshold for sufficient visibility. The all-weather capability of MMW-sensors, defined this way, can be considerably improved by introducing additional means that enhance the contrast between the runway and its adjacent terrain. When using MMW-radars like the HiVision sensor of the European Aeronautic Defense and Space Company (EADS) (Pirkl & Tospann, 1997), this can be accomplished by installing retro-reflectors along the runway, which are similar to the runway lights supporting visual landings. Furthermore, the use of reflectors enables an easy detection of the runway in MMW-radar images even if there is no contrast at all between the runway and its surroundings, as is the case for grass runways or simple forward-operating strips in military operations. In addition, passive retro-reflectors are easy to manufacture and to install. While radar reflectors provide the salient visual cues needed for runway identification, hence providing lateral guidance, there is still the need for vertical guidance. Active MMW radar primarily delivers information about the range and the angular direction (azimuth) of a certain object. This results in a flat plane-view representation, also referred to as a B-scope image of the scene, which does not provide any altitude or height information (Korn, Doehler & Hecker, 2000), i.e. it looks similar to a bird's eye view (see Figure 1, left picture). This range/angle information can be transformed into a perspective view "out-of-the-window" (C-scope), but there is still a lack of information about the objects' height or their vertical position. Thus, a perspective "out-of-the-window-view" (see Figure 1, right picture) can only be generated if the elevation of the surrounding terrain is known. To circumvent this requirement, the so-called "flat-earth-assumption" is often used (Pavel & Sharma, 1996). The only information needed under this assumption is the actual height above the flat earth which can be derived from the actual barometric altitude adjusted for the current air pressure (QNH) at the target airport. This introduces a height error of up to 20m to the resulting perspective image. This level of inaccuracy is too high for the vertical guidance needed in a low-visibility landing. A further possibility to solve the vertical guidance problem in the MMW radar image is to use special types of retro-reflectors where the amount of reflected energy is dependent on the aspect angle in the vertical plane.

Figure 1. MMW radar image of a runway, B-scope presentation (left image), C-scope presentation (right image). The perspective C-scope presentation only can be derived from the B-scope when additional information about the terrain elevation is known or the flat earth assumption is used together with the barometric altitude.

Albeit different in nature, a similar problem exists for visual approaches as the mere visibility of the runway provides only weak vertical guidance cues. Moreover, pilots' assessment of the appropriate glide-path angle is prone to visual illusions and is extremely variable, particularly during night time and/or when pilots misinterpret the width of the runway (Mertens & Lewis, 1982). Therefore, visual landing aids such as Precision Approach Path Indicators (PAPIs) or Visual Approach Slope Indicators (VASIs) have been developed to provide vertical guidance for visual approaches. Divided into three sections, this paper describes how such a concept can be transferred to the MMW-radar domain. First, the MMW Radar-PAPI/VASI concept will be outlined and the results of flight trials will be reported. These flights were completed in a research aircraft equipped with the HiVision MMW radar sensor of EADS (Pirkl & Tospann, 1997). Flight trials involved a series of straight-in approaches on a runway equipped with the above-mentioned radar retro-reflectors.

Second, a low-fidelity part-task simulation of a straight-in landing involving fourteen pilots was performed to answer the question as to how efficiently these kinds of radar images can be utilized by pilots to extract the visual PAPI categories required for successful landings. Two radar display concepts were developed that differ in the spatial arrangement of the retro-reflectors that provide vertical guidance. These were compared to a familiar color-coded PAPI baseline condition. Third and finally, a radar-image generator was interfaced with a high-fidelity generic cockpit simulator.An empirical study involving six pilots of study 2 examined different options as to how to implement the Radar-PAPI concept in the cockpit. In this study, pilot performance differences between the untransformed plane-view (B-scope) and the transformed perspective (C-scope) radar image were investigated also addressing crew coordination issues.

Study 1: Passive Aids for MMW Radar Based Approach and Landing

Today there are different types of visual landing aids. First, approach and runway lights assist pilots in identifying the runway providing support mainly in lateral guidance. Second, Precision Approach Path Indicators (PAPI) or the older Visual Approach Slope Indicators (VASI) assist pilots in the landing approach by providing color-coded visual glide path information. Figure 2 illustrates the color-coding principle of PAPI and VASI landing aids, which differ with regard to the spatial arrangement of the color lights and the number of guidance categories.

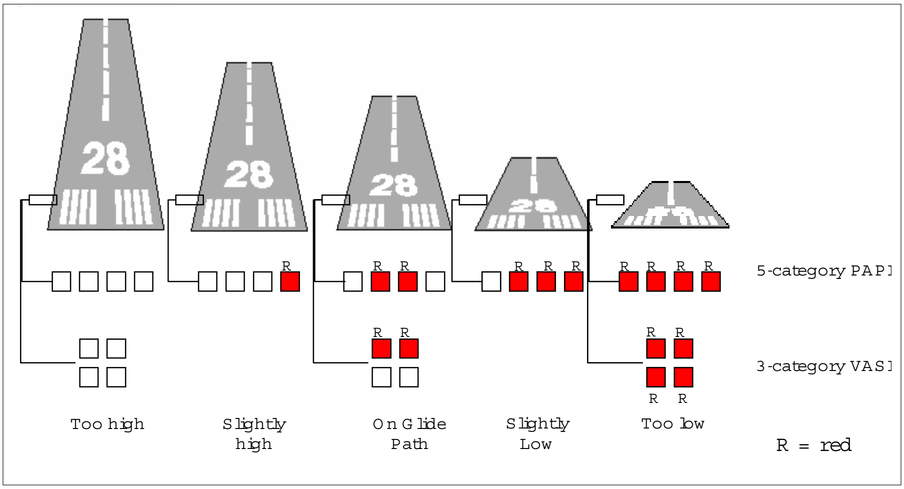


Figure 2. Red (R) and white color-coding principle of glide path information of typical 5-category PAPI and 3-category VASI visual landing aids.

To transfer this concept to the application of radar reflectors as landing aids for a MMW-radar based EVS approach, two different types of reflectors are needed. For marking the runway border, the amount of reflected energy should be independent of the aspect angle for both the horizontal and the vertical plane to enable an easy detection of the runway structure in the radar image. For vertical guidance, the amount of reflected energy should be a sharp and well-defined function of the aspect angle in the vertical plane, but nearly independent of the aspect angle in the horizontal plane.

Technical details of a variety of different reflectors and their most important properties can be found in Currie and Brown (1987). A trihedral corner reflector (see Figure 3 left) was selected as a runway marker as this reflector has the advantage over other reflectors because it has wide lobes in both planes while also providing a relatively large radar cross section (RCS), which is a measure for the amount of energy reflected back from an object to the radar sensor (Currie & Brown, 1987). Diplanes (see Figure 3 right) have a large RCS, too. They have a broad beam in the plane perpendicular to the radar seam and a very narrow one in the plane along the seam. Therefore, the amount of energy reflected back to the sensor will be dependent on the aspect angle in the vertical plane. Thus, precise installation of a multiple of these reflectors can provide a range of guidance cues to the pilot as to where the pilot is flying in relation to the target descend path. In other words, these reflectors can be applied to build up a "radar-PAPI" or "radar-VASI" system. Instead of using colors for visual glide slope indicators, now the glide slope will be intensity coded.

Figure 3. Corner reflector (left) and Diplane reflector (right) used in the flight trials.

Figure 4 shows the principle of a "radar-PAPI" configuration with 4 diplane reflectors providing image augmentation for four glide-path angles (α_1 to α_4) optimized for a nominal descend path. In principle, α_1 and α_2 provide maximum-intensity reflections at angles below and α_3 and α_4 at angles above the nominal glide path. Following the coding rules of Figure 4 pilots get categorical information about their position relative to the nominal glide path (e.g. "slight above").

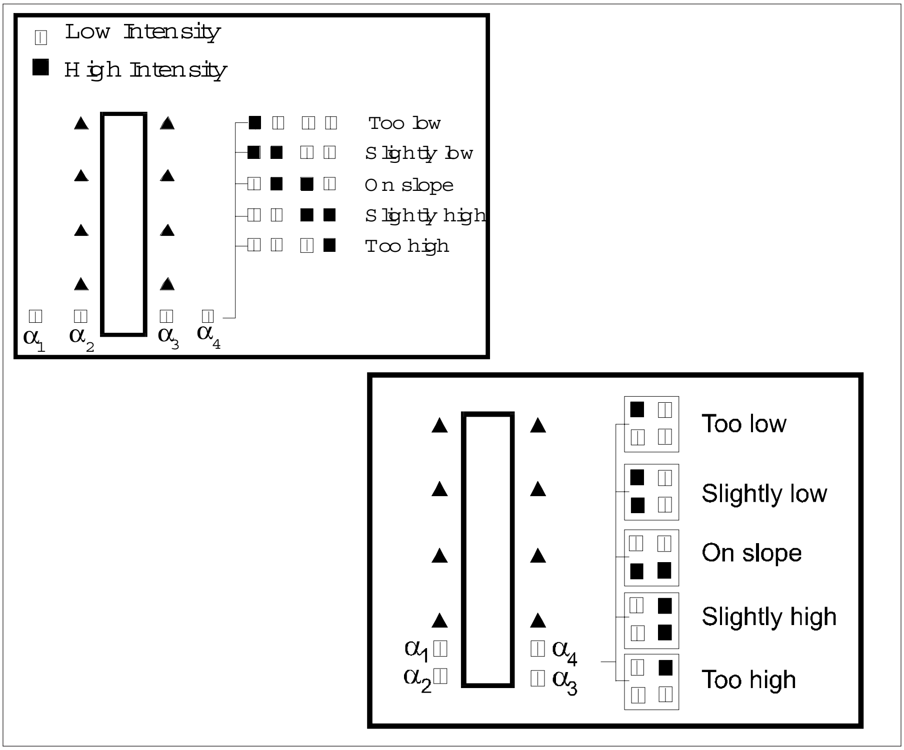


Figure 4. Intensity-coding principle of a 5-category "Radar-PAPI" (left) and 5-category "Radar-VASI" (right) obtained with four diplane radar retro-reflectors providing image augmentation for four glide-path angles (α_1 to α_4).

Method

Procedure. Figure 5 shows the configuration of corner reflectors (C-1, ..., C-8) and diplane reflectors (D-1, ..., D-4) around the grass runway 26 at the research airport in Braunschweig, which were installed for a series of flights testing the functional feasibility of the Radar-PAPI concept. Unfortunately, as some of the diplanes would have been too close to the asphalt runway, the local airport authorities could not allow the one-row line-up of a PAPI configuration due to safety reasons. Therefore, we used only a two-over-two configuration, subsequently referred to as "radar-VASI" configuration, despite the fact that this concept still provides five guidance categories and not three like the visual VASI landing aid concept.

About 25 approaches were flown to the grass runway 26. There were no displays of the radar images to the pilots. The pilots were instructed to perform visual approaches to the runway varying the vertical approach angle from steep approach of about 4½° down to flat 2° approaches to acquire a variety of radar images that can be used to analyze how sensitive the reflects of the different diplanes are to small glide-angle variations.

Figure 5. Installation of radar reflectors along grass runway 26 in Braunschweig, Germany, Corner reflectors C-1, …, C-8 and Diplanes D-1, …, D-4 with α_1=2.3°, α_2=2.9°, α_3=3.5°, α_4=4.1° supporting a precision straight-in landing with a 3.2° nominal glide path.

Results and Discussion of Study 1

Here, it is focused on the main results obtained in Study 1. Those types of analysis done with the image data that were of a technical nature around purely engineering issues will be omitted. For these type of results, the reader is referred to Korn and Doehler (2003). Figure 6 shows a radar image from an approach to the grass runway 26 in Braunschweig. The signals of the 8 corner reflectors indicate the grass runway very clearly. In Braunschweig the nominal glide path angle for the runway 26 is 3.2°. There are only high intensity values coming from the left pair of reflectors with α =2.3° and α =2.9°. That means, in this case, the aircraft was below the nominal glide slope.

Figure 6. Radar image of the "Radar-VASI" system. The aircraft is below glide slope, because only the left pair of Diplanes with α=2.3° and α=2.9° have high intensities.

In conclusion, the flight tests were promising in demonstrating the functional feasibility of a radar-based EVS approach supporting vertical and lateral guidance that can be provided by the two chosen types of retro-reflectors. First, corner reflectors offer a very simple way to provide lateral guidance by augmenting the salience of the runway structure in the radar images. Especially in the case of grass runways or airports where there is no contrast at all in the radar image between the runway and its vicinity, such reflectors may play an important role. Second, the additional use of special reflectors (e.g. of diplanes) which have a narrow main-lobe in elevation can provide vertical guidance information on glide slope deviations similar to the existing VASI- or PAPI- visual landing aids. Especially for MMW sensors that measure only range and azimuth but no elevation, this kind of information is essential for pilots.

Study 2

Having established the all-weather capability of the MMW sensor, and moreover, the improvements in the detection of the runway and the glide-path angle with the help of retro reflectors, raises the important question as to whether pilots are able to take advantage of raw radar images displayed without any automatic target extraction. Thus, the question is: can pilots utilize these images as visual cues for a precise straight-in landing in low visibility. This human performance issue was addressed in a laboratory study that is presented now.

Method

Experimental Task. The participants were required to control a two-axis compensatory tracking task. The axes represented the lateral and the vertical position of a simulated aircraft. Speed was held constant. The task was designed to provide a low-fidelity simulation of the final stage of a straight-in landing with a 3° glide slope on a target runway. The task, however, did not simulate the system dynamics of a real aircraft, but used a pure first-order system in both axes. Deviations from target values in both axes were driven by a random-like disturbance input obtained by the sum of five (lateral) or four (vertical) sine waves varying in frequency and amplitude. The chosen frequencies and amplitudes are depicted in Table 1. The resulting input was added as noise to the position of the aircraft during approach. Consequently, the disturbance impact on the visual cues in the display increased as the aircraft approached the threshold. While an amplitude of 50 m meant only a displacement of 5 pixels at the starting distance of 3000 m it increased to a displacement of e.g. 30 pixels at a distance of about 500m to the threshold. Generally, high amplitudes were chosen for low frequencies and vice versa. Initial phases of sine waves were randomized to provide a different disturbance input for each trial yet maintain an equal task difficulty across trials. Control dynamics remained the same under all display conditions that will be described. Participants used a joystick to control deviations in both axes. Aircraft position relative to both the lateral target position and the nominal glide path were stored at 20 Hz from which the average root-mean square error (RMSE) was calculated off-line.

Display Conditions

The display of deviations from the lateral and vertical target differed and resulted in three display conditions (Figure 7).

Table 1

Disturbance input to the first-order system dynamics

	Amplitude [m]	Frequency [Hz]
Vertical	5.0	.02
	2.0	.04
	.5	.10
	.25	.333
Horizontal	50.0	02
	9.0	.04
	5.0	.083
	3.0	.166
	1.5	.333

Visual PAPI (baseline condition). Participants viewed the landing aid imagery of a standard 4-box color PAPI located left to the runway trapezoid along with a HUD-like superimposed primary flight display (PFD). Only altitude valued varied as speed was held constant at 97 knots. Heading also remained fixed at 360°. No pitch and roll input was possible. Deviation from the 3-degree glide slope angle could be derived from the PAPI principle, with which all pilots were familiar, i.e. the on-slope state was indicated by two red and two white lights (see Figure 7, right), with three red lights being a slightly below-slope, and four red lights a very below-slope indication. Above-slope indications corresponded to the reciprocal pattern of three white (slightly above) and four white (high above) lights. Visual cues for lateral guidance had to be derived from the displacement of the runway symbol. Like in real flights, the geometry of the runway does also convey information on the height once the pilots become familiar with how it should look like when they are on the nominal descent. Further cues on the rate of descend could be derived from the speed of altimeter changes.

Radar-PAPI. Ten pairs of corner reflector images were displayed to mark the detected runway. Pairs were equally spaced with a fixed distance corresponding to 200 m. Four diplanes arranged in a horizontal row became visible at the front end of the corner reflector pairs (see Figure 7, left). The distance of this row to the front pair of corner reflectors was smaller than the distance between each pair of corner reflectors to alleviate the difficulty to distinguish between diplanes and corner reflectors. A vertical line served as a visual target for lateral tracking and had to be kept within the corner reflector image pairs. A further vertical reference line on the left side of the screen provided an indication of the distance between runway and aircraft (major tick marks at every 500 m, minor tick marks at every 100 m without exact numerical values given). Glide slope deviations were indicated following the coding scheme illustrated in Figure 4. Since this imagery does not provide any visual cue for the rate of descend, a vertical bar indicating the rate of altitude changes was located in the center of the screen. The size of the upward or downward move of this vertical bar gave an indication as to the climb or the descend rate, respectively. Participants were instructed and trained to avoid climbing to capture the glide slope from below, instead, should try to intercept the glide slope with lowering the descend rate or with horizontal flight.

Radar VASI. This condition only differed from the Radar-PAPI condition by the spatial arrangement of the diplane reflector images. As shown in Figure 7 (middle), a two-over-two arrangement was chosen according to the VASI landing aid concept already illustrated in Figure 4 (right).

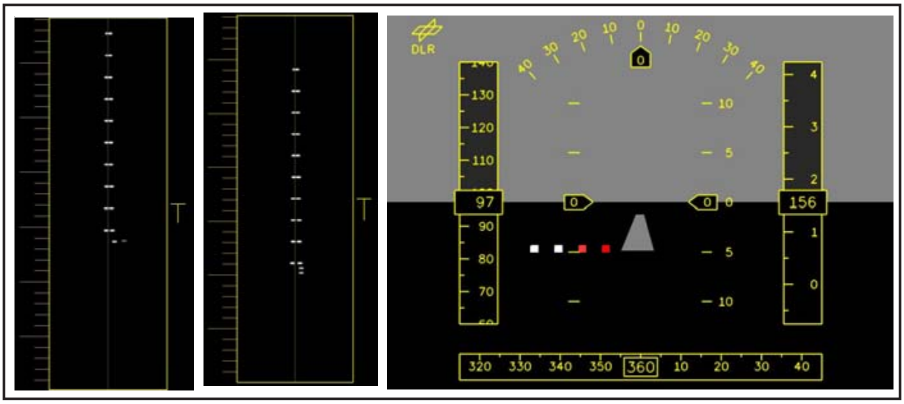


Figure 7. Display conditions: Radar-PAPI (left), Radar-VASI (middle) and visual PAPI (right)

Participants

Fourteen male volunteers participated in the study.Thirteen participants were licensed pilots at the Private Pilot (PPL) level (N=6) or higher (N=7). One pilot was in PPL training. All higher-rated pilots were in possession of a valid IFR rating. The VFR and the IFR groups did not differ with regard to age. The age in the VFR group that included the pilot in PPL training ranged from 25 to 45 years (M = 32.43, SD = 6.27) and the age in the IFR group ranged from 24 to 46 years (M = 31.71, SD = 7.87). Total flight experience ranged from 17 to 550 hours (M = 150.65, SD = 180.02) in the VFR group and from 220 to 4000 hours (M = 1000.00, SD = 1392.47) in the IFR group

Apparatus

The task was programmed in Visual C++ using Open-GL graphic libraries for the display and was implemented on a laptop computer equipped with a 15-inch TFT screen.

Procedure

Testing sessions started with a short introduction that provided some background information on technical aspects of the MMW forward-looking sensor and associated EVS issues. This introduction included a video sequence recorded during the functional field test of Study 1. This video showed sequences of real radar images sensed while the research aircraft intercepted and captured a 3-degree glide path from below and passing a grass runway upon which four diplanes in VASI arrangement and eight pairs of corner reflectors had been installed. This demonstrated the transition through the full state range of glide path angles covered by the diplanes. Next the participants were demonstrated how to control the experimental task, which they subsequently trained for about 45 min. Each straight-in landing trial lasted about 70 s and ended upon touchdown. In all Radar-PAPI and Radar VASI trials, an indication (the letters "DH" for decision height) appeared in the bottom of the screen when an altitude of 100 feet was reached. At that time the participants were instructed to press the space-bar of the laptop keyboard, which switched the screen to the perspective visual PAPI view obscured by heavy fog. The participants completed the trial with this view.An overall of 30 training trials was completed by each participant, 10 for each display condition in the order of first visual PAPI, second Radar-PAPI, and finally radar VASI. The following test session included 15 trials performed in blocks of 5 trials varying the three display conditions. Block order was counter-balanced across participants. The test session ended with a debriefing during which the pilots described their experience with the task and their preference with regard to the two radar display conditions.

Results

The average RMSE scores for both axes were calculated for each test trial. A three-way split-plot analysis of variance (ANOVA) design (Kirk, 1982) was used for the statistical evaluation. The design involved a between-subject factor varying pilot rating (VFR vs. IFR), a within-subject factor of Display Condition (Visual PAPI, Radar-PAPI, Radar VASI), and a second within-subject five-level factor of Trial Number (one through five). The number of degrees of freedom of the F-ratios was corrected for sphericity violations using Greenhouse-Geisser's epsilon and significant effects involving more than two factor levels were further evaluated by means of multiple comparison tests using Bonferoni adjusted significance levels. Separate ANOVAs were computed for lateral and vertical average RMSE scores using the GLM procedure of the SPSS statistical software package.

Vertical Tracking. The ANOVA indicated that the average vertical RMSE did not vary significantly as a function of any of the experimental variables (all $p > .05$). This also applies for the slight performance superiority of the IFR group over the VFR group. As Figure 8 (upper panels) demonstrates, performance was remarkably stable across trials in both groups, with only the IFR group showing a tendency of better performance using the visual PAPI display. In fact, a separate ANOVA computed on the data of the IFR group only, revealed a reliable main effect of Display Condition ($F(2,12) = 7.14$; $p < .01$) that confirms this.

Lateral Tracking. The ANOVA revealed a highly reliable main effect of Display Condition ($F(2,24) = 50.08$; $p < .001$). Best performance, i.e. lowest RMSE scores, in both groups were obtained under the visual PAPI condition, which is confirmed by the outcome of the multiple comparison tests that revealed a significant difference between the visual PAPI condition as compared to both Radar-PAPI and radar-VASI conditions (both p < .001). A superiority of the IFR group as indicated by smaller average RMSE scores was marginally significant (F(1,12) = 3.59, p = .08). However, a significant Pilot Rating by Display Condition interaction effect was found (F(2,24) = 5.21; p < .05). The source of this effect can be seen in Figure8 (lower panels). The performance in the VFR group was poorer under the Radar VASI condition as compared to the Radar-PAPI condition, whereas no such difference occurred in the IFR group. As already stated for the vertical tracking error, the performance was rather stable across trials with only one outlier seen in the fifth trial of the VFR group where the performance in the Radar-PAPI condition became erratically worse.

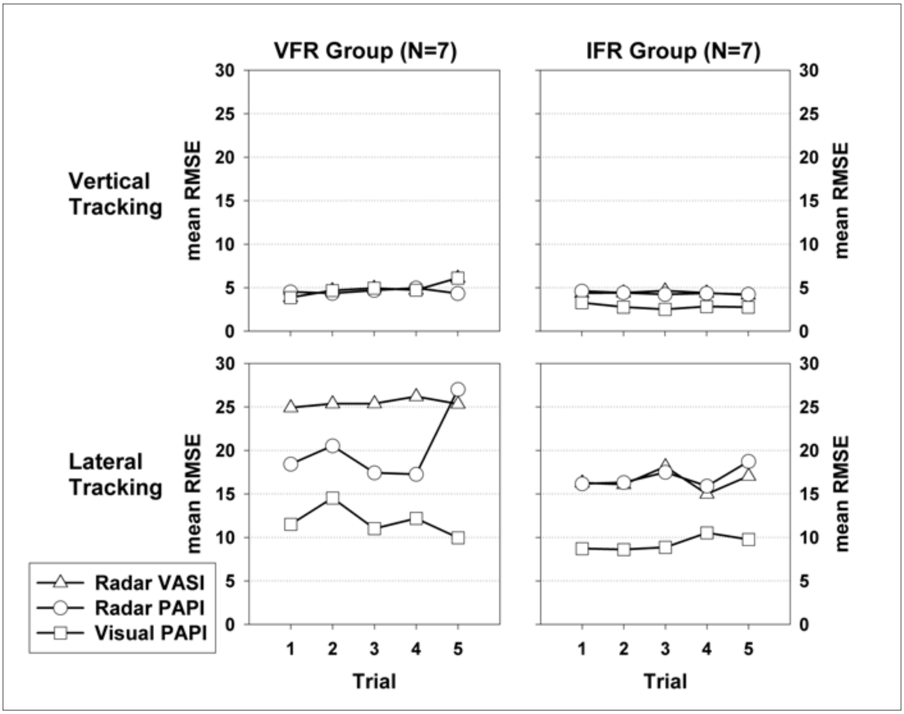


Figure 8. Average tracking error (upper panels: vertical tracking, lower panels: lateral tracking) as a function of trial and separated for three display conditions. Left panels: VFR-pilots, right panels: IFR-pilots.

Subjective Preference. In the VFR group all but one pilot preferred the Radar-PAPI display over the radar VASI display. In the IFR group three pilots preferred the Radar-PAPI, three the radar VASI condition and one pilot indicated no preference. Reasons for Radar-PAPI preference were that this display enabled an easier separation of diplanes from corner reflectors and provided a better indication of extreme glide path deviations. Pilots who preferred the radar VASI, on the other hand, liked the close proximity of the group of diplanes to the corner reflectors and indicated that they felt it easier to simultaneously control both axes this way. In fact, this statement agrees with the better radar VASI lateral tracking performance in these pilots. Almost all pilots expressed feelings of surprise about their good learning progress during training. Most pilots complained about not having any cue as to where exactly to expect diplane images as long as they were not visible. They would prefer to be supported by some kind of marker cues. Rather than having difficulties with the radar images they expressed to have the most trouble with the unfamiliar system dynamics, which clearly was not one of a real aircraft and required constant and at times heavy control input.

Discussion of Study 2

The analysis of the average RMS tracking error revealed some interesting findings. Deviations from the vertical glide path were small and remarkably stable across trials and display conditions in both pilot groups. The average RMSE was markedly higher in lateral than in vertical tracking. A superiority of the visual PAPI conditions emerged in the efficiency of lateral tracking. Furthermore, the pilots with IFR rating outperformed the VFR pilot in lateral tracking, in general, and under the radar VASI condition in particular.There are a variety of possible explanations for this pattern of results. First, the lateral disturbance input was larger in amplitude as compared to the disturbance amplitude used for the vertical tracking loop. This explains the upward shift in average tracking error, but not the differential pattern. It is very likely that vertical tracking was emphasized with higher priority over lateral tracking. On the one hand, this emphasis, per se, is stressed by the central research question under scrutiny, i.e. the question whether MMW radar images can provide visual glide path cues. The higher proficiency of the IFR pilots may indicate that this group (1) has more spare capacity to control the secondary lateral tracking task, (2) use more efficient visual scanning strategies to sample both corner and diplane images, hence, showed better divided attention (Wickens & Hollands, 2000) (3) produced less crosstalk between control over both axes, i.e. the control input to one axis interfere less with the input to the other axis (Fracker & Wickens, 1989), or (4) were less dependent on the visual geometry cues provided by the perspective runway trapezoid and can better use the vertical target line, which behaved similar to a localizer.A mix of these effects is most likely to have taken place. Regardless of what exactly was the origin of the performance advantage of the IFR group, the fact indicates some promising criterion validity of the artificial part-task simulation with regard to more realistic straight-in landing scenarios where an IFR education would most likely be an advantage as well.

The slight advantage of Radar-PAPIs over radar VASIs, which agrees by and large with subjective preference statements of the pilots, corresponds to the respective advantage found with the existing visual color-light landing aids (Castle, 1983).

Study 3

Study 3 sought an implementation of the radar-PAPI concept in a more realistic cockpit environment. The study had three major goals:

1. Interfacing a MMW radar image generator with a high-fidelity fixed-based cockpit simulator,

2. Exploration of the display concept for radar reflector images. and
3. Investigation of task allocation issues of pilot flying (PF) and pilot non-flying (PNF).

With the implementation of the Radar-PAPI concept in a high fidelity simulation it was expected to broaden the empirical basis of the human performance aspect of this concept to a more realistic environment. The results of Study 2 indicate a critical weakness of the plane-view radar settings for efficient visual information sampling to control lateral path deviations in comparison to the perspective view provided by the traditional PAPI. As already outlined above the transformation of the radar image data to create a perspective C-scope image is possible but will always convey unreliable height information to the pilot. Therefore, a potential superiority of the C-scope over the B-scope display on pilot performance as well as the potential performance impact of the implied height error in the C-scope display were assessed in Study 3. For this purpose, three altitude error conditions were nested within the C-scope display condition (0m, 15m 25m). Addressing the third goal in the list above, another issue was the question of how to allocate tasks between the two-pilot crew during a straight-in landing in low visibility. Therefore, we implemented a third display condition by varying procedural aspects of crew coordination. The PNF's role was to derive localizer and glide-slope information from a B-scope display and to guide the PF in a similar manner as a military controller does using a Precision Approach Radar.The PF maintained outside scene monitoring using a HUD. The evaluation is based on objective performance data and subjective workload and situation awareness ratings.

Method

Participants. Six volunteers participated in the study.Two pilots were licensed at the Private Pilot (PPL) level and four were commercial pilots (CPL: N=2; ATPL: N=2). All pilots were in possession of a valid IFR rating. Age ranged from 25 to 46 years (M = 33.00, SD = 7.77). Total flight experience ranged from 230 to 4000 hours (M = 1130.00, SD = 1478.11). All pilots had already participated in study 2.

Experimental Setup. The experiments were completed using a fixed-base cockpit simulator, which is a flight simulator based on the Airbus A320 aircraft supported with 180° width x 40° height collimated outside view.

Simulation Scenario. The task for the flight crews was to perform an approach and landing to the runway 28 of Zurich International Airport under CAT II conditions. This runway is only equipped with a VOR. Furthermore, the final phase of the approach to runway 28 consists of a long 3° segment followed by a steep 3.7° final to avoid mountainous terrain (see Figure 9). Thus, the standard IFR approach to runway 28 is to follow the VOR/DME guidance until MDA (minimum descend altitude) of 904 feet above ground and then perform a visual approach using the PAPI lights for vertical guidance. For the tested scenario under CAT II visual conditions the flight crew was instructed to follow the standard approach procedure until MDA and then to use the radar image for further lateral and vertical guidance to the new decision height (DH) of 100ft. At DH, the flight crew had to confirm visual contact to the runway to complete the landing.

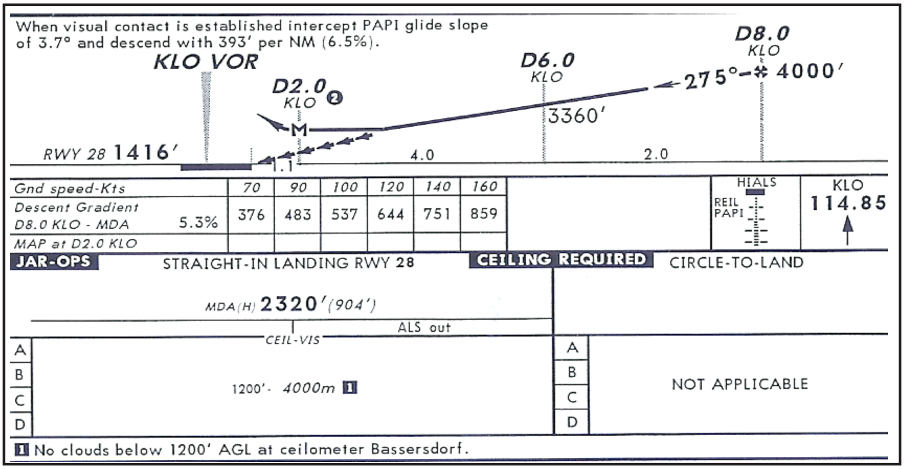


Figure 9. Approach chart (vertical profile) to Zurich 28.

In the simulated scenario 13 pairs of corner reflectors were used to mark the runway.The Radar-PAPI system consisted of 6 diplane reflectors located between the first and second corner reflector pairs. For the experiments a MMW radar simulator was used to generate radar images from our Zurich Airport database including the reflections of the respective radar reflectors. The basic parameters of the simulated radar were quite similar to those of the existing MMW radar sensor HiVision of EADS (Pirkl & Tospann, 1997). The azimuth resolution was about 0.2° with a coverage of 40°. The maximum scanning distance of the simulated radar was about 5000 m with a resolution of 10 m. An example of the resulting radar B-scope image is depicted in Figure 10 (left).

Display Conditions. In this study three different display conditions were investigated. In the first two conditions the PFD of the PF was replaced by a display including the radar image in B-scope representation (Figure10 left) and C-scope (Figure10 right) representation, respectively.The most important aircraft state information, like speed, heading, altitude, radar height, and sink rate were added to these displays, too. In both display conditions the task of the PF was to follow the standard VOR/DME approach using these new PFDs and the standard Navigation Display until the radar reflectors became visible in the radar images. Then the pilot had to use the radar images to derive the necessary lateral and vertical guidance to complete the final 3.7° segment of the approach to runway 28. The primary task of the PNF was to monitor the outside scene and to inform the PF when runway structures became visible (at DH of 100ft) and where the runway was located with respect to the aircraft's position.

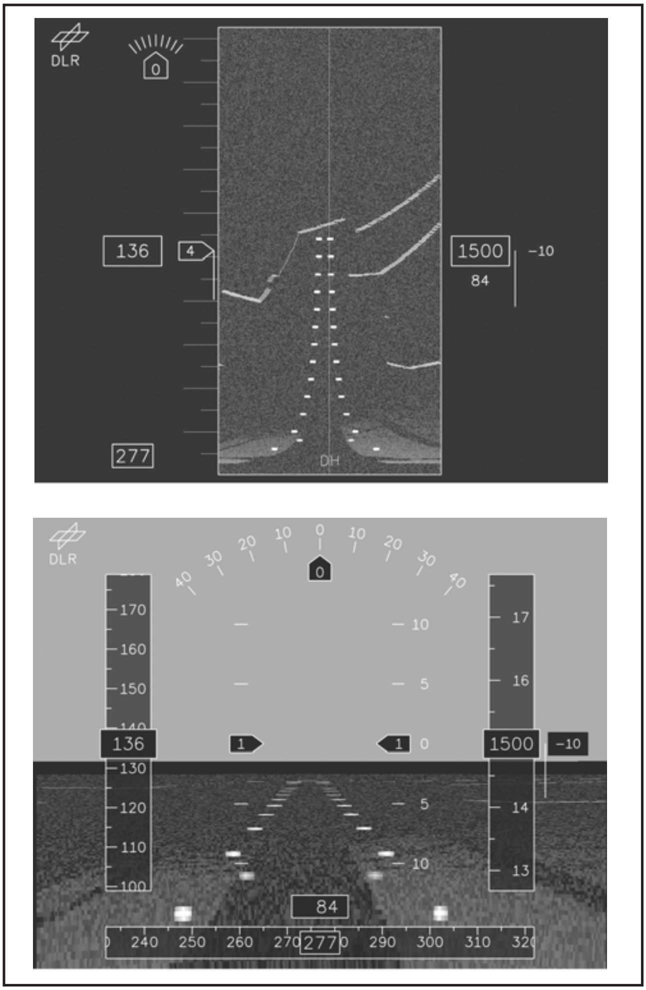


Figure 10. Radar image on a modified PFD. Top image: B-scope representation together with aircraft state data such as speed (136 kn), heading (277) altitude (1500 ft), radar height (84 ft), and vertical speed (1000 ft/min). Bottom image: Perspective C-scope representation with an overlay of PFD symbology. Within both radar image representation the information for lateral (corner reflectors) and vertical (diplane reflectors) guidance information can be identified easily: Here the aircraft is on centerline and on slope.

For the third condition the B-scope display was presented to the PNF and the PF used a HUD. The task of the PNF was to derive guidance information from the B-scope image and to generate guidance commands to the PF. For vertical guidance the PNF generated sink rate commands such as "sink rate: 900" to the PF; for the lateral guidance heading commands such as "heading: 275" were generated in the first segment of the final approach. For smaller corrections during the last 2000 m of the approach commands such as "to the left" or "to the right" were used to guide the PF until he identified the visual contact to the runway and completed the landing visually. In Figure 11, the outside scene at a height of 90 ft (shortly after DH) through the HUD is depicted. In the HUD only aircraft state data like speed, vertical speed (VS), altitude, radar height, or heading, but no guidance information (e.g. ILS) was displayed.

Testing sessions started with a short introduction to the tasks and to the cockpit simulator itself. In a training phase of 30 to 60 minutes the pilots had the chance to familiarize themselves to the simulator and the different display conditions. After that each pilot had to perform three approaches for each display condition. The sequence of display conditions were counterbalanced across pilots. Within the C-scope display conditions, different height errors (0 m, 15 m, and 25m) were introduced and pilots were held naïve as to what height error they would encounter. For the objective performance quantification the aircraft state data was recorded at 20 Hz. From this data lateral and vertical deviations from the target approach path were computed.

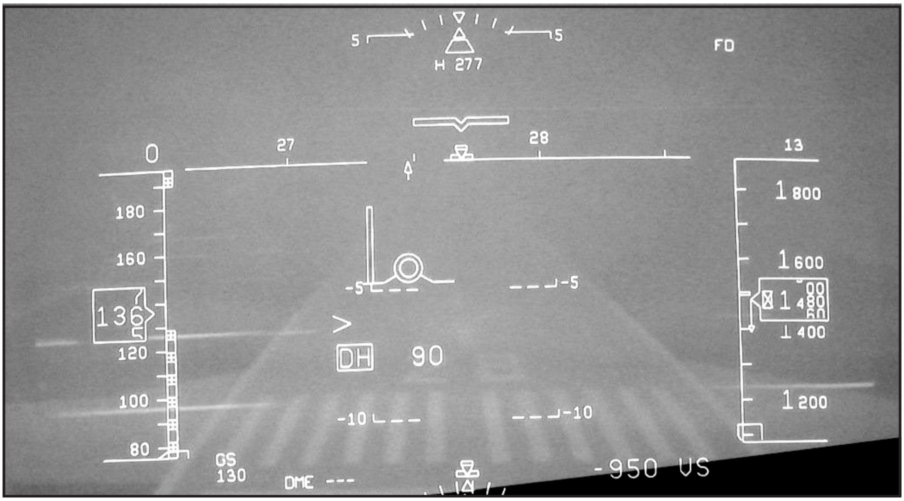


Figure 11. View through the HUD to the outside scene under CAT II visual conditions at a height of 90 ft above the runway.

Subjective Questionnaires

After each block of three simulation trials per display condition, the pilot completed two subjective questionnaires.

Mental workload. The NASA-Task Load Index (NASA-TLX) (Hart & Staveland, 1988) was used to assess subjective workload.

Situation Awareness. The Situational Awareness Rating Technique (SART) (Taylor, 1990) was used to examine potential between-display differences in subjective situation awareness associated with the approach-and-landing task performed. We used the original SART scales consisting of the three primary

SART rating dimensions: (1) demand on attentional resources (D), (2) supply of attentional resources (S), and (3) understanding (U) with 10 secondary rating dimensions nested within the three primary ones (see Eurocontrol, 2003)

Results

The straight-in approach-and-landing scenarios were divided into two segments with regard to distance to threshold. Segment 1 ranged from 4000 to 2000 meter to threshold and segment 2 from 2000 to 500 meter to threshold. Lateral deviations and vertical angle deviations from the target precision glide slope were scored as root-mean-square errors (RMSE). Two separate 2 * 3 univariate analyses of variance (ANOVA) were computed for lateral and vertical RMSEs, respectively. Scores were averaged across the three scenario repetitions within each display condition prior to ANOVA computation. ANOVA factors were display condition (three levels: B-scope, C-scope, HUD) and flight segment (first vs. second). Both factors were varied within subjects. Degree-of-freedom adjustments of significance levels were performed using Greenhouse epsilon, when a heterogeneous variance-covariance matrix was detected by Mauchley's sphericity test following the GLM repeated measures procedure of the Statistical Package of the Social Sciences (SPSS). In a subsequent ANOVA the three C-scope trials were analyzed for the influence of the experimentally induced altitude error (no error, 15 meter, 25 meter) on the vertical and lateral RMSE indicators of flight path performance.

Vertical Tracking. The ANOVA revealed a reliable main effect due to the flight segment ($F(1,5) = 8.274$; $p = .035$) caused by a significant increase in average vertical RMSE in the second as compared to the first flight segment. A main effect of display condition was not found ($F < 1$), however, a reliable interaction effect between flight segment and display condition was found ($F(2,10) = 12.27$; $p = .007$). As Figure 12 (right panel) illustrates, this interaction pattern resulted from the finding that an increase of average RMSE from the first to the second flight segment occurred only in the B-scope and the HUD conditions, whereas an opposite effect of an average RMSE reduction occurred in the C-scope condition. The subsequent ANOVA based on the three C-scope trials to examine the impact of altitude error and flight segment on vertical RMSE scores did not yield any significant results (all $p > .30$).

Lateral Tracking. The pattern of significant ANOVA effects found with the lateral RMSE score were the same as described above for the lateral scores, i.e. a main effect of segment was reliable ($F(1,5) = 9.17$; $p = .029$), a main effect of display condition was lacking ($F < 1$), and an interaction effect between both factors was found ($F(2,10) = 7.01$; $p = .016$). However, the meaning of these effects was quite different to that found with the vertical scores. First, the average RMSE was smaller in the second compared to the first flight segment indicating a performance improvement closer to the threshold. Figure 12 (left panel) illustrates, also that, the interaction effect had a quite different cause than the one found with vertical path tracking. Performance was stable across both flight segments with the C-scope display. Little improvement occurred using the B-scope display; however, a marked improvement occurred in the HUD condition.

Taking both flight segments together points to an advantage of the C-scope display. The lack of statistical confidence for an effect of display conditions was mainly a result of a huge standard error found in the B-scope condition (three times as high as in the other two conditions) pointing to individual differences. In fact, two pilots performed extremely poor, and two other pilots extremely well in this condition. The other two display conditions produced much more homogeneous results. The ANOVA based on the three C-scope trials also did not yield any significant effect of different altitude errors for the lateral RMSE scores (all p > .20).

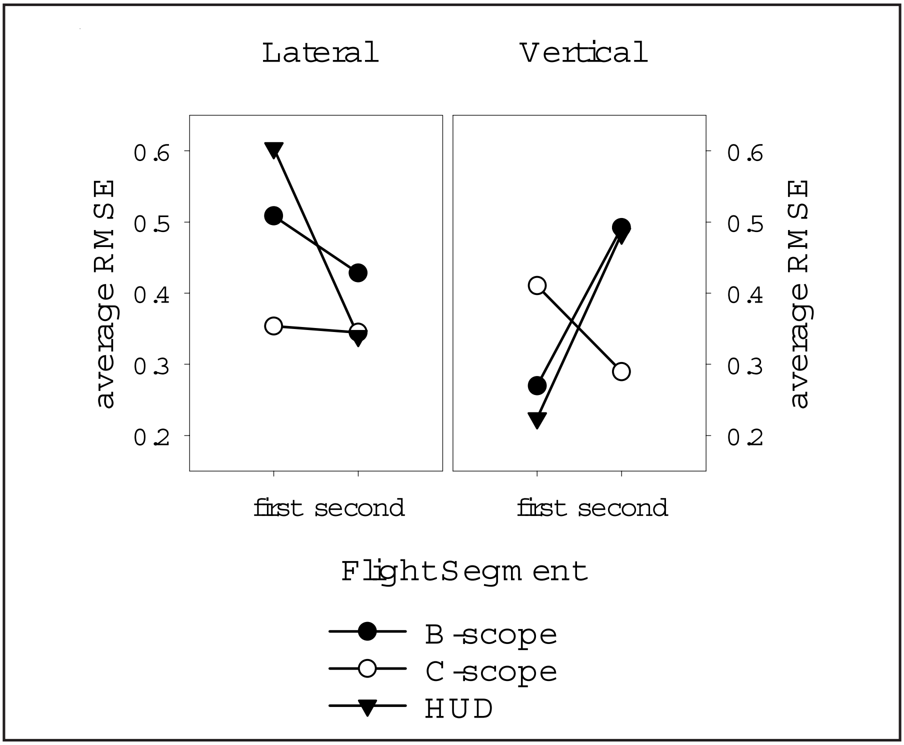


Figure 12. Lateral and vertical flight path deviation error as a function of first and second flight segment separated for three display conditions

Subjective Questionnaires

All subjective questionnaire data was analyzed in one-factorial ANOVAs involving the three-level factor of display condition only.

Mental workload. The evaluation of the untransformed NASA-TLX ratings averaged across all six sub-scales slightly failed to reveal a reliable effect due to display condition (F(1,5) = 4.63; p = .061). Numerically, the results suggest higher subjective workload experienced in the B-scope (M = 11.19) as compared to the C-scope (M = 7.39) and to the HUD condition (M = 8.00) on the scale ranging from 0 (no workload) to 20 (high workload).

Situation awareness (SA). First, all three primary scales, often referred to as the 3D-SART (Eurocontrol, 2003) were analyzed in three separate ANOVAs. Average *resource demand* was rated highest for the B-scope (5.5), intermediate for the C-scope (4.0), and lowest for the HUD (3.5) condition on the 7-point Likert scale labeled Low (rating point 1) to High (rating point 7). These differences caused a significant main effect of display condition (F(1,5) = 11.47; p = .003). Statistical post-hoc comparison of means (Bonferroni-adjusted) indicated a significant difference between B-scope and HUD (p = .035), a marginal difference between B-scope and C-scope (p = .052) and no difference between C-scope and HUD (p = .61). The results obtained with the *resource supply* scale had almost the same pattern as the *resource demand* scale (highest score in the B-scope condition) indicating the fact that the perceived higher demand in gaining SA is responded to by a higher amount of attention and effort supplied to the situation. The effect is statistically a little bit smaller, yet sufficiently reliable (F(1,5) = 5.35; p = .029). No significant differences were observed in the third *understanding* scale (F < 1) suggesting the notion that regardless of the demand perceived and the effort needed to gain SA, no differences between the display conditions were found in the resulting quality of SA. A *global SA-score* was obtained by averaging the secondary scales nested within each of the three primary scales (often referred to as the 10-D SART) and subsequently using the simple algorithm computed on these average values: SA (global) = Understanding – (Demand – Supply). The ANOVA performed on this score did not reveal a difference between display conditions (F < 1).

Discussion of Study 3

The study investigated the question whether radar-PAPIs can replace traditional visual PAPIs in low visibility.A straight-in landing scenario was chosen that involved a VOR/DME approach that would usually require the pilots to establish visual contact at MDA (here at about 900 feet above ground) in order to intercept a visual PAPI glide slope. In the present study, CAT-II visibility (< 200 feet above ground) was simulated and pilots were instructed to continue landing by intercepting the Radar-PAPI glide slope. This scenario was chosen as it represents the conditions, for which EVS may make flight operations less weather dependent and which is addressed by the above-mentioned FAA rule change. The study compared three display formats and associated crew coordination issues: (1) the PF views a head-down B-scope display and switches to visual landing upon PNF's call-out that runway is in sight; (2) the PF views a head-down C-scope display and switches to visual landing upon PNF's call-out that runway is in sight; (3) the PF maintains outside viewing through a head-up display (HUD) that displays primary flight guidance information. The PF receives vertical and lateral guidance from the PNF who views a head-down B-scope. The PNF guidance terminates upon PF's call-out that runway is in sight.

The display with the highest similarity with the existing PAPI landing aids based on color lights is the C-scope display. Here, the radar images are transformed to enable a presentation in a perspective view.The other two display conditions use a B-scope view, i.e. a flat plane-view representation of the radar image. In general, the results show that all pilots could extract the glide path information from the radar images to perform a successful landing under all three experimental conditions. Subjective workload and situation awareness ratings suggest that the first condition (PF flies B-scope) was experienced as most demanding. Two almost equally long path segments, an early segment (4000 m – 2000 m to threshold) and a late segment (2000 m – 500 m to threshold) were differentiated for the analysis of the lateral and vertical flight path deviations. This analysis revealed an interesting pattern. Lateral tracking was best supported with the C-scope display, particularly during the first segment, whereas the conditions 'PF flies B-scope' and 'PF flies HUD' promoted poor lateral tracking in the first segment, however, improving in the second segment. This interaction pattern between display format and flight segment reversed in meaning when the vertical tracking performance is considered. Here, C-scope performance in the first segment was poor but improved in the second segment, whereas initial performance in the other two conditions supported with the B-scope format was good but deteriorated markedly in the second segment. Thus, a consistent vertical and lateral performance improvement from the first to the second flight segment only occurred with the C-scope display. In the other two conditions it appeared as if improvements in lateral tracking from the first to the second segment occurred at the simultaneous expense of deteriorating vertical tracking efficiency suggesting the notion of a performance trade-off associated with a shift in attention emphasis.

In order to explain this pattern of results there are several display phenomena to take into account. First, as already mentioned, one clear advantage of the C-scope display is its familiarity due to the fact that it mostly resembles the known visual PAPI concept. Furthermore, the geometry generated by the perspective runway-marking reflector images convey external environmental cues for both lateral and vertical flight path deviations despite the fact that the latter might not be accurate because of the inherent potential height error. Experimentally induced height errors, however, did neither affect vertical nor lateral flight path tracking at either flight segment. Pilots were made aware of the potential height error inherent in the C-scope display beforehand and were instructed to focus on the diplane reflectors to control the descent path. Disregarding the existence of an absolute height error, the figural changes in the geometry of runway markers produced during the descent behave much like a configural or integrative display (Bennett & Flach, 1992) allowing parallel processing of vertical and lateral guidance cues. It can be assumed that pilots used the vertical deviation cues provided by the diplanes to calibrate the altitude cues in the perspective runway trapezoid generated by the corner reflectors. It is further suggested that once this calibration was achieved, path and altitude deviation cues could be derived from the perspective runway object of the C-scope display allowing their parallel sampling. The additional observation of further improvements with the C-scope display closer to the runway threshold is consistent with this suggestion for two reasons. First, the perspective runway image becomes clearer in the final segment, enhancing the altitude and path maintenance cues. Second, the requirement for their parallel sampling increased in the final segment, hence, the performance cost of the B-scope was greatest in the final segment as it did not allow parallel sampling. The attention issues involved will be dealt with in some more detail below in the general discussion.

General Discussion and Conclusions

The new FAA rule on Enhanced Flight Vision Systems (EFVS) requires pilot's ability to unambiguously identify runway structures in sensor images to continue approaches below decision height (DH) in low visibility. Millimeter-wave (MMW) radar sensors provide the best penetration of adverse weather. However, due to low resolution and the underlying sensing principle of MMW radar the quality of the radar images is rather poor in comparison to infrared or TV-camera images. Moreover, MMW radar sensors can only detect the structure of the runway lighting system and the runway itself if there is enough contrast between the asphalt or the concrete and the surrounding grass area. More often than not these structures cannot be identified reliably in radar images.

The identification of the runway can be improved dramatically by the installation of radar retro-reflectors at the landing site, which augment the crucial elements in the radar images. Radar-reflectors, hence, can act under IMC similar to the runway lighting system under VMC. This may become an even more important issue in military operations for landings in low visibility on unequipped air fields, e.g. on grass runways that do not provide any contrast cues in radar images.

Based on this rationale and in analogy to visual approach aids we developed a concept of a so-called "radar-PAPI" system. This system relies on two different types of retro-reflectors. A set of corner reflectors served to provide runway markings required for lateral flight path guidance and a set of glide-angle sensitive diplanes served to provide vertical guidance. Proof-of-concept was examined in three steps. First, a series of straight-in landings was completed with a research aircraft equipped with a MMW radar sensor on a runway equipped with the above-mentioned radar retro-reflectors. Second, a low-fidelity part-task simulation of a straight-in landing was performed to answer the question as to how efficiently these kind of radar images can be used by pilots to extract the PAPI categories required for successful landings. Finally, a radar-image generator was interfaced with a high-fidelity generic cockpit simulator.An empirical study involving pilots of the second study examined different options as to how to implement the Radar-PAPI concept in the cockpit and also address crew coordination issues. Results of the first study were promising in demonstrating that sufficiently reliable lateral and vertical guidance information can be extracted from the recorded radar images. This points out the general functional feasibility of the radar-PAPI concept. The second study suggests that pilots learn the extraction of the visual cues from comparatively noisy radar images surprisingly fast with the help of cue augmentation provided by retro-reflectors. This conclusion could be confirmed in the third study involving a much more realistic straight-in approach and landing simulation scenario. In general, Radar-PAPIs may in fact replace the traditional visual color PAPIs and can, therefore, enable safe landings on airports not equipped with ILS navigation support when no visual contact can be established at MDA. Both human performance studies revealed some further interesting findings on how to develop a display format for the implementation of the radar-PAPI concept. Particularly for the final phase of the approach (< 1nm) the performance benefit of a perspective display format emerged. In the flat plane view (B-scope) a flight path/altitude performance trade-off occurred, which indicates the difficulty of pilots to access vertical and lateral guidance information in parallel. The underlying mechanism of this deficiency is most likely due to fundamental attention processes. A similar flight path/altitude performance trade-off was observed in human performance studies at NASA-AMES on the *attentional tunneling* phenomenon when pilots use a HUD. Foyle, McCann, Sanford and Schwirtzke (1993) and McCann and Foyle (1994) used a low-fidelity flight task simulation and instructed pilot subjects to maintain 100 ft altitude while simultaneously flying along a flight path trajectory defined by a set of pyramid markers located on a virtual terrain. This had to be accomplished in the presence of vertical and lateral turbulence. In a HUD condition, a digital altitude reading was superimposed, which resulted in improved altitude maintenance in comparison to conditions without altitude information. The interesting finding, however, was that this improvement occurred at the cost of simultaneously increased flight path deviations. Thus, in these studies, a flight path/altitude performance trade-off also occurred, which point to impaired parallel sampling of information from the outside scene and the superimposed HUD symbology. This performance trade-off again occurred in a further study (Foyle, McCann & Sheldon, 1995) in which several HUD formats were compared using the same flight task as in the earlier studies. In complete confirmation with the earlier finding, the flight path/altitude performance trade-off occurred in a condition where the altitude reading (both digital and analog) had a fixed location on the HUD. In another so-called scene-linked format, the altitude gauge on the HUD became a virtual object of the outside scene, i.e. the gauge moved in a conformal manner with outside objects during the flight. By means of such a scene-linking of HUD and outside-scene objects the performance trade-off was absent. Now, the HUD improved both altitude and flight path maintenance. Foyle et al. (1995) relate this improvement to a performance advantage of scene-linked HUD display formats obtained by reduced attentional tunneling meaning by better divided attention. It is suggested here that the underlying attention mechanism responsible for this performance benefit is the same as the one that produced the benefits of the traditional color PAPI display in study 2 and of the C-scope display in study 3. The crucial element is the formation of a common visual object from which information for both flight path and altitude deviations can be derived. In the present study, this was given by the geometry of the perspective runway structure. In the study by Foyle et al. (1995), scene-linking supported a perceptual integration of the altitude information in the outside terrain imagery. These interpretations are also in good agreement with Treisman's (1982) feature integration theory.

To control the vertical and lateral glide path sequentially was even more necessary in the third condition of study 3 during which the PF received the corresponding guidance commands sequentially by the PNF who monitored a B-scope display. No marked differences, however, were observed to the condition during which the PF viewed the B-scope display directly. The cost of additional transition lags caused by the indirect access to the required information may have been compensated by the fact that the PNF acting in this role was quite efficient.

Another display phenomenon that needs to be considered here is display clutter. Using the C-scope display the reflector images lumped closely together when they first became visible. The closer the aircraft got to the airport the more they spread apart. This clutter can be expected to be less critical for lateral rather than for vertical flight path tracking as the latter requires the separation of both reflector types. In fact, the whole pattern of results obtained with the C-scope display confirms this explanation in that lateral tracking performance remained rather stable, whereas vertical tracking improved in the transition from the first to the second flight segment. The B-scope display did not produce these clutter problems, hence provided a much better spatial separation of the reflector images. This promoted a better identification of guidance cues particularly affecting vertical tracking in the first flight segment. The observed display superiority of a less clutter-prone plane-view display of the radar-PAPIs more distant (> 1nm) to the threshold and its inferiority due to less efficient parallel cue sampling to control altitude and flight path in closer proximity (< 1 nm) to the threshold poses a quite puzzling design problem to be solved in the future. Despite the fact that the inherent height error in the transformed C-scope radar image had no adverse impact on pilot performance, it should still be born in mind that superimposing a C-scope image with the outside scene on a HUD cannot be achieved satisfactorily.This remains a weakness of the MMW-sensor as compared to the IR-sensor.An enhanced vision concept that uses the MMW-sensor in combination with an IR-sensor may be an option to compensate for the weakness of both sensors. Consider a landing under light fog or haze. The IR sensor may provide an augmented perspective vision, but may not provide enough vision needed to perceive color PAPIs. The MMW sensor used in combination with passive retro-reflectors could complement the type of landing with a Radar-PAPI glide slope. With the addition of synthetic vision derived from terrain databases adapted to the actual aircraft position state, further improvement in pilot vision under adverse weather can be gained.

References

Bennett, K. B. & Flach, J. M. (1992). Graphical displays: Im-plications for divided attention, focused attention, and problem solving. *Human Factors, 34*, 513-533.

Castle, B. (1983). *Evaluation of Precision Approach Path Indicator (PAPI).* Federal Aviation Administration Technical Report No. DOT/FAA/CT-82/153. Washington, DC: Department of Transportation.

Currie, N. C. & Brown, C. E. (Eds.) (1987). *Principles and applications of millimetre-wave radar*. Norwood, MA: Artech House.

EUROCONTROL (2003). *The development of situation awareness measures in ATM systems*,Technical Report HRS/SHSP-005-REP-01. Brussels: Author.

Federal Aviation Administration (2004). *Federal Register 14 CFR Parts 1, 91. Enhanced Flight Vision Systems; Final Rule*. Washington, DC: Department of Transportation.

Foyle, D. C., McCann, R. S., Sanford, B. D. & Schwirtzke, M. F. J. (1993). Attentional effects with superimposed symbology: Implications for head-up displays (HUD). *Proceedings of the 37th Annual Meeting of the Human Factors and Ergonomics Society*, 1340-1344. Santa Monica, CA: Human Factors and Ergonomics Society.

Foyle, D. C., McCann, R. S. & Sheldon, S. G.. (1995). Attentional issues with superimposed symbology: formats for scene-linked displays. In R.S. Jensen & L.A. Rakovan (Eds.), *Proceedings of the Eighth International Symposium on Aviation Psychology*, 98-103. Columbus, OH: Ohio State University.

Fracker, M. L. & Wickens, C. D. (1989). Resources, confusions, and compatibility in dual axis tracking: displays, controls, and dynamics. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 80-96.

Hart, S. G.. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati. (Eds). *Human mental workload (pp. 139-183)*. Amsterdam: North Holland.

Horne, W.F., Ekiert, S., Radke, J., Tucker, R. R., Hannan, C. H. & Zak, J. A. (1992). Synthetic vision system technology demonstration methodology and results to date. *SAE International Technical Paper Series #921971*, Aerotech '92; 1-19.

Kirk, R. E. (1982). *Experimental design (2nd ed.)*. Belmont: Brooks/Cole.

Korn, B. & Doehler, H.-U. (2003). Passive landing aids for precision EVS approach and landing, *Proceedings of the 22nd Digital Avionics Systems Conference (DASC 2003)*, Indianapolis, IN, paper 9.D.1

Korn, B., Doehler, H.-U. & Hecker, P. (2000). MMW Radar Based Navigation: Solutions of the Vertical Position Problem. In J.G. Verly (Ed.), *Proceedings of the 2000 SPIE Conference, Orlando, FL (Vol. 4023, Enhanced and Synthetic Vision)*, 29-37.

McCann, R. S. & Foyle, D. C. (1994). Superimposed symbology: attentional problems and design solutions. *SAE Transactions: Journal of Aerospace, 103*, 2009-2016.

Mertens, H. W. & Lewis, M. F. (1982). Effect of different runway sizes on pilot performance during simulated night landing approaches. *Aviation, Space, and Environmental Medicine, 53*, 463-471.

Pavel, M. & Sharma, R. K. (1996). Fusion of radar images - rectification without the flat earth assumption. In J.G. Verly (Ed.*), Proceedings of the 1996 SPIE Conference, Orlando, FL (Vol. 2736, Enhanced and Synthetic Vision)*, 108-118.

Pirkl, M. & Tospann, F. J. (1997). The HiVision MM-Wave Radar for Enhanced Vision Systems in Civil and Military Transportation Aircraft. In J.G. Verly (Ed.), *Proceedings of the 1997 SPIE Conference, Orlando, FL (Vol. 3088, Enhanced and Synthetic Vision*), 8-18.

Taylor, R. M. (1990). Situational Awareness Rating Technique (SART): The development of a toll for aircrew systems design. *In AGARD Conference Proceedings No 478, Situational Awareness in Aerospace Operations*. Aerospace Medical Panel Symposium, Copenhagen, October, 2-6. Neuilly Sur Seine, France: Advisory Group of Aerospace Research and Development.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 194-214.

Wickens, C. D. & Hollands, J. (2000). *Engineering psychology and human performance (3rd ed.).* New York: Prentice Hall.

International Journal of Applied Aviation Studies, 4, Number 2
Copyright © 2004, FAA Academy, Oklahoma City, OK

Developing a Standardized Mechanism for Measuring Airline Service Performance: A Preparation for Airlines and the Flying Public

Brent D. Bowen

The University of Nebraska at Omaha
6001 Dodge Street
Allwine Hall 422
Omaha, NE 68182.0508
Voice 402.554.3424
Fax 402.554.3781
bbowen@mail.unomaha.edu

Dean E. Headley

Wichita State University
W. Frank Barton School of Business
Wichita, KS 67260.0084
Voice 316.978.3367
Fax 316.978.3276
Dean.Headley@wichita.edu

and

Chien-tsung Lu

Central Missouri State University
Department of Aviation
Thomas R. Gaines 210
Warrensburg, MO 64093
Voice 660.543.8092
ctlu@cmsu1.cmsu.edu

Abstract

The Airline Quality Rating (AQR) was first announced in the early 1990s as an objective mechanism of measuring airline service performance based on multiple selected criteria important to air

---

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

travel consumers. Since then, the AQR results have been available to assist government in decision-making as well as to help airlines identify service gaps. Initially, the national AQR reports used weighted averages and monthly performance data in the areas of on-time arrivals, involuntary denied boardings, mishandled baggage, and a combination category of customer complaints from which major airline comparative performances were reported. In order to further strengthen the significance of the AQR results and prepare the comparison of major airlines' service for aviation enthusiasts (airline managers and the flying public), this paper revisited AQR data sets and reanalyzed four previous reports (2000, 2001, 2002, and 2003). Benchmarking techniques were utilized to assess airlines' AQR performance and track the comparative service quality for major U.S. airlines' domestic operations. Based on the benchmarked outcomes, an airline's operational deficiencies can be retrieved, while simultaneously opening a window for public review and service improvement.

The initial objective of the AQR was to develop a method for accurately measuring airline service quality. Since 1991, the Airline Quality Rating (AQR) reports have successfully contributed insightful measurements to the air transportation industry, and their influence with the aviation business has continuously grown (Goodman, 1992, April 29; Mann, 2000). To date, the AQR reports have been widely recognized and are available for airlines to promote service quality and attract potential passengers (Spencer, 1999). Statistically speaking, most air carriers are seeking to better control their service quality by quantitative methods (Bowen, Headley, & Luedtke, 1992), and the results of the annual AQR are convenient for airline industry to better approach their business goals. By further benchmarking airline service quality, air carriers can use benchmarked findings to reveal the existing weaknesses of their past and current services. In so doing, airlines are able to compare their service with business rivals and therefore wisely prepare to bolster future performance. The benchmarked findings are also useful to the flying public for reconsidering airlines who could not satisfy passengers' needs and interests.

Service Quality— A Brief Review of Previous AQR Measurements

Since the first report in 1991, the AQR has revealed some up and down years in airlines' service quality. From 1991 through 1994, the AQR scores showed declining performance for the industry (See Table 1). During the financially turbulent period between 1995 and 1997, airline quality slightly improved each year. However, after 1997, AQR scores again began rapidly declining. The AQR has given the flying public a means to quantify airline service quality.The AQR chronicles the airline service quality with a system that is fractured and near a breaking point.

Table 1

Average Reports of AQR – 10 Major Airlines

Year	1991	1992	1993	1994	1995	1996	1997	1998	1999
Industry Average	-.017	-.031	-.070	-.110	-.0853	-.0686	-.0001	-1.609	-1.608

Note. The higher the average is, the better the service performance would be. Please refer to the subsection, "Operational Mechanism of the AQR," herein.

The Air Carriers' Commitment to Passengers

In late 1999, in the wake of the urgent requirement from the Department of Transportation (DOT) and the FAA regarding the protection of airline passengers' rights, fourteen (14) air carriers gave their service commitment to the aviation community to better treat their customers. The commitment to better passenger service includes notification of flight delays, satisfying the needs of waiting passengers due to delays, promptly awarding passengers refund airfare, and being more responsive to passengers' complaints (DOT Aviation Consumer Protection Division, n.d.). To accomplish the commitment, large airlines launched the "Customer First Plan." Although it was slightly different from airline to airline, the aspects of the Customer First Plan concerned a multidimensional commitment to pamper passengers. For instance, Northwest Airlines imposed several operational strategies (based on the fundamental requirements from the DOT and the FAA) into its Customer First Plan. The updated passenger service standards included:

1) offering the lowest fares
2) notification of a known delay, cancellation, and diversions
3) on-time baggage delivery
4) increase in baggage liability limits
5) the right of the passenger to cancel reservations and to ask for ticket refunds
6) adequate accommodations for passengers with disabilities and special needs
7) satisfaction of customers' essential needs during long delays
8) fairness and consistency regarding the handling of denied-boarding passengers
9) disclosure of travel itineraries, rational cancellation policies, and fair frequent flyer rules
10) aircraft configuration information
11) standardization of service among code-share partners, and
12) a prompt response to customer's complaints (Northwest Airlines, n.d.)

Yet the outcome of the airline's service commitment was not available to the flying public.

Measurement of Former AQR Projects

Bowen, Headley, and Kane (1998) first applied benchmarking techniques in evaluating annual AQR reports between 1991 and 1997. They stated that benchmarking airline performance was a functional way to monitor "overall industry performance and the resulting effects of situational environment changes" for regulatory officials, financial investors, and interest groups (p.9). By benchmarking service, airlines can locate specific service indicators that need to be enhanced. Although benchmarking airline quality does not provide solutions for the airlines, benchmarked airline service can identify critical issues for airline operators, who can then seek remedies for these deficiencies (Bowen, Headley, & Kane, 1998). Bowen, Headley, and Lu (2003) conducted another review regarding major airlines' AQR scores between 1998 and 2001. They further suggested that airlines should apply benchmarking techniques to continuously oversee their operational performance and promptly resolve abnormalities by providing corrective actions.

Over the years, the AQR has provided objective measurements in airline service quality for the public, government, and airline industry. In the meantime, the project of benchmarking the AQR has helped to promote airline quality (Bowen, Headley, & Kane, 1998; Bowen & Lu, 2003). This article attempts to: (a) pose a follow-up benchmarked assessment of the annual AQRs from 2000 to 2003 for aviation enthusiasts, and (b) to continue a benchmarking analysis tailored specifically for the airline industry.

A Brief Introduction of AQR Measurement

The following section illustrates the concepts of conducting the AQR analysis before benchmarking airline services.

Selecting elements of AQR. The intent of the AQR reports was to offer some observations in the areas of most concern to consumers, such as ontime arrivals, mishandled baggage, involuntary denied boardings, and 12 customer complaint areas. This information can be useful in helping less familiar consumers gain a perspective on issues of interest in the airline industry. Elements considered for inclusion in the rating scale were screened to meet two basic criteria: 1) elements must be obtainable from published data sources for each airline; and 2) elements must have relevance to consumer concerns regarding airline quality. Data for the elements used in calculating the ratings represent performance aspects of airlines that were important to consumers. Factors included in the rating scale were taken from an initial list of over 80 potential factors (Bowen, Headley, Kane, & Lutte, 1999).

Raw data collection and individual weights of the AQR. The raw data of each selected measuring element was originally retrieved from the annual Air Travel Consumer Report published by the Department of Transportation (DOT), which is very credible and reliable; this data is believed by the industry to be bias-free. Weights of selected areas were established by surveying 65 airline industry experts regarding their opinions as to what consumers would rate as most important (scaled from 0 [lowest] to 10 [highest] on the Likert scale) in judging airline performance. The value of individual weight was assigned by the average score of the associated survey question in the returned questionnaire. Survey experts were objectively selected as key informants in the AQR; the group of experts comprised airport executives, airline managers, members of aviation interest groups and organizations, customers, aviation authorities, academia, and aircraft manufacturers. Such a range allows for the highest objectivity and lowest homogeneity in units of analysis. Meanwhile, each weight and element was assigned a plus or minus sign to reflect the nature of impact based on consumers' perception. Weights reflect the importance of the criteria in consumer decision-making, while signs (+ or -) reflect the direction of impact that the criteria should have on consumers' rating of airline quality. Table 2 shows the average-weights and the effect of impact retrieved from key informants.

Table 2

Airline Quality Rating Criteria, Weight, and Impact

Criteria		WEIGHT	(+/-)	IMPACT
OT	**On-Time**	8.63	+	**Positive**
DB	**Denied Boardings**	8.03	-	**Negative**
MB	**Mishandled Baggage**	7.92	-	**Negative**
CC	**Customer Complaints**	7.17	-	**Negative**

Note. The twelve consumer complaint areas include Flight Problems; Oversales; Reservations, Ticketing and Boardings; Fares; Refunds; Baggage; Customer Service; Disability; Advertising; Tours; Animals; Other.

Operational Mechanism of the AQR. The AQR calculating mechanism permits the comparison of major domestic airlines on a regular basis using a standard set of quality factors (See Appendix 1). The calculating formula takes multiple weighted objective factors into consideration in arriving at a single rating for an airline (Bowen, Headley, & Luedtke, 1992). Before the AQR, there was effectively no consistent method for monitoring the quality of airlines on a timely, objective, and comparable basis (Headley & Bowen, 1997). With the introduction of the AQR, a multifactor, weighted-average approach became available for the public about "how well airlines meet consumer concerns" (Spencer, 1999, p. 49).

The AQR is a weighted average of multiple elements important to consumers when judging the quality of airline services (Goodman, 1992, April 29; Bowen, Headley, & Lutte, 1993; Mann, 2000). The formula for calculating the AQR score is posed as the following:

$$AQR = \frac{(+8.63 * OT) - (8.03 * DB) - (7.92 * MB) - (7.17 * CC)}{(8.63 + 8.03 + 7.92 + 7.17)}$$

When raw data of all selected factors (e.g., OT, DB, MB, and CC), weights (e.g., 8.63 for OT), and impacts (e.g., "+" for OT and "-" for DB, MB, and CC) are combined for calculation, it yields a single-interval scaled value for an airline's level of quality. The value of the AQR is comparable across airlines and time periods as well (Headley & Bowen, 1997).

In this paper, benchmarking techniques, combined with first-hand data analysis, were used and important results provided. This section outlined the introduction of benchmarking techniques, procedures of benchmarking, and display of benchmarked findings and results.

Introduction to the Importance of Benchmarking Techniques

It is crucial for commercial industry to make progress, meet customer requirements, become more competitive, and generate profits. The benchmarking technique is one of the prevailing aids for achieving the desired business goals (Camp, 1989; Patterson, 1996). For example, in order to improve productivity and enhance service quality, many airlines (such as American Airlines and Southwest Airlines) conducted benchmarking analysis with success (Camp, 1989; Fitz-enz, 1993; Patterson, 1996; Tucker, 1996). Southwest Airlines used benchmarking to accelerate both its aircraft refueling process and the turnaround time of its ground luggage cars; this helped to shorten the carrier's ground operating time and improve its on-time performance (Fitz-enz, 1993). Conversely, a lack of benchmarking was among a laundry list of problems plaguing Trans World Airlines (TWA). Focusing on price wars instead of its customer service, financial, or management performance, TWA's financial statement had been in the red ink for five years without the company making any increase in market share in late 1980s (Fitz-enz, 1993). An independent benchmarking evaluation revealed that customer satisfaction in TWA's service had declined dramatically during that period, which led to a loss of customers as well as skilled employees between 1988 and 1991 and to its bankruptcy in 1992 (Fitz-enz, 1993). TWA's case highlights the importance of accurately targeting and benchmarking business performance.

Benchmarking Procedures

Planning operational procedures plays an important role in benchmarking methodology (Camp, 1989). Many researchers provided various sets of benchmarking procedures based on their unique disciplines and research purposes. The benchmarking remarks contributed by quantitative researchers like Camp, Tucker, Patterson, and other researchers had synthesized a workable set of benchmarking processes used in this paper, as outlined in the following steps: (a) calculate AQR scores from 2000 to 2003 reports and display general findings; (b) decide specific elements to be benchmarked; (c) identify benchmarking criteria; (d) conduct data analysis and compare results; and (e) find the problems (Fitz-enz, 1993; Tucker, 1996; Patterson, 1996; Neufville & Guzmán, 1998). The research framework of this paper had been constructed at this point, and the authors took benchmarking actions after AQR reports had been calculated. Because this paper was seeking to benchmark AQRs between 1999 and 2002, (AQR reports of 2000 and 2003) and the database was ready for use, a secondary data analysis was implemented combining benchmarking procedures.

Benchmarking Airline Quality Based on Selected Criterion

The benchmarking technique is useful to target the performance difference among units of analysis. Consequently, an enhancement program tailored to bridge a specific performance gap can thus be framed (Keehley, Longmire, Medlin, & MacBride, 1997). Airline operators should proactively and continuously search for the best practices for promoting customer service through the usage of benchmarking skills (Bowen, Headley, and Kane, 1998). To this end, the criteria of benchmarking should be defined beforehand and then compared to the variance among performances. In this project, the benchmarking criterion was defined as the annual industry average performance followed by a cross-section and horizontal analysis. Researchers can easily observe the difference between airlines, as well as the deviation compared to the industry average, thus enabling a detailed analysis. By following the benchmarked analysis of annual AQR scores, readers could also identify specific airlines that either performed well or needed to reform their poor service practices.

Reliability and Validity

The governmental information databases could help researchers secure data reliability and validity. The DOT's database contains highly reliable and valid information that can be adopted to satisfy both reliability and validity criteria (Berg & Latin, 1994; Creswell, 1998; Lincoln & Cuba 1985). In addition, in this study, the reliability of the rating scale was tested through determination of Cronbach's Alpha calculation, which resulted in a reliability coefficient of 0.87 (where 1.0 is perfect) (See Table 3). This suggests that the AQR is very reliable and that factor determination results would be similar for other comparable samples (Bowen, Headley, & Luedtke, 1992).

Table 3
Reliability Coefficient

Measure	Score	Scale	Result
Cronbach's Alpha	0.87	0 ~ 1.0	Extremely high validity

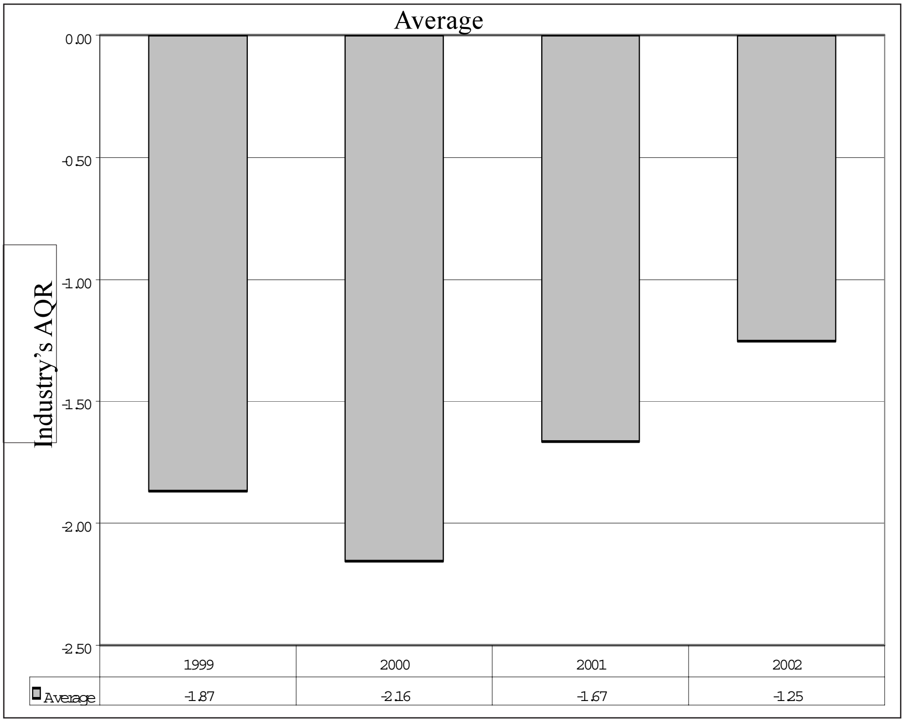
Findings

AQR Score Display and General Findings

Based on 10 major airlines, the AQR industry average score shows the entire industry declined in service quality from 1999 (-1.87) to 2000 (-2.16), but improved in 2001 (-1.67) and 2002 (-1.25) (See Table 4).

Table 4
Average AQR Scores

Score	FY 1999	FY 2000	FY 2001	FY 2002
Industry	- 1.87	- 2.16	- 1.67	- 1.25



Note. The higher the score, the better the service. Average AQR scores are based on monthly AQR score calculations using the AQR weighted average method as reported in the cited research monographs. The calendar year is used to arrive at the average AQR score for the year.

The average industry AQR score was -1.74, which was used as the criterion of annual benchmarking analysis in measuring the overall AQR score containing on-time performance, mishandled baggage, involuntary denied boardings, and customer complaints.

America West Airlines (-3.43) and United Airlines (-3.01) seriously struggled in providing a qualified service in FY 2000 report, yet their service performance turned upward in the FY 2001 and 2002 reports (See Appendix 2). Although most airlines faced a service downturn in FY 2000, all of them made progress in FY 2001 and FY 2002. In general, the industry has fiercely competed with each other since FY 2000 according to the convergent trend among airlines' AQR scores.

Table 5 shows the descriptive analysis of four years of performance by the airline industry. Because of the consistent progress of AQR scores from FY 2000 to FY 2002, it is reasonable that the median shifted from -1.83 (FY 2000) to -1.10 (FY 2002) while the mean (average) AQR score increased from -2.09 in FY 2000 to -1.12 in FY 2002. The standard deviation of FY 2002 was smaller than that of year 2000 and 2001. This situation accurately mirrored a smaller range of AQR scores in 2002 than in 2000, which indicated a higher level of service competition.

Table 5

*Descriptive Analysis of AQRs**

	FY 1999	FY 2000	FY 2001	FY 2002
Average	-1.83	-2.09	-1.53	-1.12
Standard Error	0.11	0.23	0.09	0.06
Median	-1.85	-1.83	-1.48	-1.10
Standard Deviation	0.32	0.68	0.26	0.17
Sample Variance	0.10	0.47	0.07	0.03
Range	1.11	1.96	0.78	0.54
Minimum	-2.39	-3.43	-1.97	-1.39
Maximum	-1.28	-1.47	-1.19	-0.85

* Data does not contain T WA and American Eagle Airlines

Furthermore, the worst individual AQR score (-3.43) was found in FY 2000, and the best individual AQR (-.85) was observed in FY 2002. The descriptive analysis of AQR scores in FY 2000 indicated that the service performance of targeted major airlines varied dramatically (comparing the reports of FY 1999, 2001, and 2002), which was also reflected by its larger standard deviation (.68) as to that of the other years.

Finding Operational Deficiencies by Comparing Airline Quality

Our findings revealed that the monthly AQR scores for the whole industry declined from FY 1999 to FY 2000, but showed promising performance in the 2002 and 2003 reports (See Appendix 3). The overall industry-wide performance was capricious across a four-year period on a monthly basis. The entire industry maintained fair performance for most of 1999, but that performance dropped substantially near the end of the year. Service performance in most months during FY 1999 and 2000 were below average, and overall airline service was observed to be at the lowest point in FY 2000. Fortunately, airlines started to improve in FY 2001 and achieved the highest AQR score in September 2002 (-0.8). This improvement began in July 2001, while a missing section in 2001 resulted from the September 11th attacks. All selected airlines in 2002 performed above the overall four-year annual average (-1.74). Interestingly, airlines seemed routinely to have difficulty in providing good service during Christmas (December and January) and in the summer (June, July, and August).

An in-depth analysis should be conducted to locate the causes behind the overall performance. Other correlated facts uncovered by the benchmarking analysis should also interest passengers, airlines, and to some extent, the government authorities. Southwest, Northwest, Alaska, and US Airways had noteworthy performances over the period in question, and indeed a selected benchmarked analysis comparing five airlines and the industry average from 1999 to 2002 is discussed in the appendix section. Four benchmarked evaluations of airline service over the past four years also revealed significant service deviations compared with annual industry averages.

On-Time Rates. The benchmarked AQR chart shows that the overall on-time arrival percentage declined from FY 1999 to 2000 (72.6 percent in 2000, compared to 76.1 percent in 1999) and then increased to 77.4 in FY 2001 and 82.1 percent in 2002 (See Appendix 4-1 & 4-2). Alaska Airlines constantly performed poorly in on-time service throughout the four years. Northwest Airlines encountered difficulty maintaining its on-time service in early 1999; after September 1999, its on-time rates improved and remained above average. Southwest Airlines had superior on-time rates (although it experienced two difficult months - March and October, 2000).

Involuntary Denied Boarding Rates. The cases of involuntary denied boardings increased between the third quarter of FY 1999 and the second quarter of 2000 (See Appendix 5). Northwest Airlines and US Airways performed extremely well between FY 1999 and 2002. Alaska Airlines appeared to significantly increase their numbers of denied boarding cases at the first quarter of FY 2002. The industry average remained consistently low (0.89). The denied boarding rates of Southwest and Alaska Airlines remained higher than average throughout the year from FY 1999 to 2002.

Mishandled Baggage Rates. The mishandled baggage rates were unstable throughout the four years—FY 2002 (4.72 per 1,000 passengers); 2001 (4.55); 2000 (5.29); and 1999 (5.08) (See Appendix 6-1 & 6-2). Most cases of mishandled baggage happened during the Christmas and summer period. In particular, the highest peak level of mishandled baggage occurred between December and January of FY 1999, 2000, 2001, and 2002. Benchmarked results revealed that Northwest Airlines had poor results in handling baggage in FY 1999 and 2002 but good results in FY 2000 and 2001. Alaska Airlines had poor performance in FY 1999, but remained outstanding in this category between FY 2000 and 2002. US Airways maintained stable progress in this category, and its mishandled baggage rate declined from January 2001 to December 2002 on a monthly basis. Southwest Airlines seemed more adept at handling passengers' baggage than Northwest Airlines.

Customer Complaint Rates. Overall customer complaint rates increased from 2.47 per 100,000 passengers in 1999 to 2.98 in FY 2000, but reduced to 2.11 in 2001 and finally 1.22 in 2002 (See Appendix 7-1 & 7-2). Northwest Airlines received the most complaints in 2001 while Southwest received the least across the four-year period. Alaska Airlines showed poor performance in the category of customer complaints between December 2001 and December 2002, but Alaska Airlines made a great progress in 2002. Most passengers filed their complaints to airlines between: (a) July and December in 1999, 2000, 2001, and 2002; or (b) during three extremely poor service periods in September 1999, January 2001, and December 2001. The overall performance of this category in FY 2002 improved.

Discussion and Conclusion

The findings of this study helped airlines to: 1) identify the operational difference as to that of their business counterparts, 2) enhance on-going activities and operating structures, and 3) provide better service to its customers. Furthermore, based on the findings of this paper, the flying public could identify the airlines that provide better customer service. Although the industry's service quality struggled in FY 1999, the industry had, on the whole, accomplished its service commitment as outlined in its Customer First Plan after 1999 (FY 2000, 2001, and 2002).

While the overall service had been enhanced in FY 2000, 2001, and 2002 in relation to the categories of on-time rate and involuntary denied-boarding, rates of mishandled baggage and customer complaints increased. In particular, the service trend downturned substantially in FY 2001. Both involuntary denied boarding and mishandled baggage cases remained high during the peak season—summer and New Years. This status quo has not been changed since the first version of AQR was published in 1991. In order to reduce the cost of upholding the "Customer First" commitment, airlines should sincerely consider solving the root causes of involuntary denied boarding and mishandled baggage. On the other hand, potential passengers should preview airline's service performance based on the findings of this study before booking their next flight. By doing so, passengers will bear a smaller probability of encountering an unexpected "service deficiency" during their air travel.

On-time rate is the most critical factor in this study.Although it is wise for the airline to gain a higher AQR by promoting on-time performance, on-time departure/ arrivals are affected by many uncontrollable factors. When just the more controllable elements are considered, the ten largest U.S. carriers maintained an 82.1% on-time arrival record for FY 2002. This was better than the 77.4% on-time arrival record for the industry in FY 2001. Across the industry, 0.72 passengers per 10,000 boardings were bumped from their flight involuntarily in FY 2002. This is a 16% improvement in the industry rate of 0.86 denied boardings per 10,000 passengers in FY 2001. On average, the Department of Transportation received 1.22 customer complaints per 100,000 passengers for the ten largest carriers in FY 2002. The volume of complaints in FY 2002 represents a 42% decrease in the rate of complaints over the 2001 rate. The complaint rate for the industry has stayed under 1 case per 100,000 passengers consistently since August 2002. These complaints represent a wide range of areas, such as cancellations, delays, oversales, reservation and ticketing problems, fares, refunds, customer treatment, unfair advertising, and other general problems.

In addition, no one died in a commercial airline accident in FY 2002, making it the third year in the past decade (1993, 1998, and 2002) to have been a fatality-free year of operations. In contrast, the 2001 fatality total of 530, reflects the deaths (265) on the four airplanes used in the terrorist attacks of September 11, 2001, and the 265 deaths associated with the November 12, 2001, crash in Belle Harbor, New York. These fatal airline tragedies produced the worst year for fatalities in the past 20 for American commercial aviation (Part 121 scheduled). According to the AQR 2003 (FY 2002) report, approximately 470 million people boarded one of the 10 largest U.S. carriers to fly somewhere inside the U.S. in 2002 (down from 488 million in 2001). This does not consider those international travelers who departed from U.S. airports. Regional and commuter carriers accounted for an additional estimated 83 million passengers flying domestic routes as well (also 83 million in 2001). This totals an estimated 553 million people flying to destinations within the U.S. in 2002 (down from 571 million in 2001). With the economic downturn since late 2000 and the 9/11/01 terrorist attacks, passenger volumes have dropped by about 20%. Fortunately, the passengers are returning regardless of the slow revenue recovery. While industry-wide service quality is making progress, passenger growth and a return to previous volumes are predicted at a moderate pace over the next several years. Yet as long as the economy remains strong and potential threats of terrorism can be remedied, the flying public will use air travel at previously seen volumes (Bowen & Headley, 2003).

Final Thought

To date, rather than consulting the regular Aviation Rulemaking Advisory Committee (ARAC) members as mandated by the Federal Advisory Committee Act and Administrative Procedure Act, the FAA's rulemakers are joyfully communicating with industry-driven Aviation Rulemaking Committee (ARC). With the authority given by the Federal Aviation Administrator, the FAA's rulemakers can overcome legislative bottleneck and improve governmental productivity and efficiency.The usage of benchmarking skills cannot be limited to airlines and the flying public; it also provides useful insights to the FAA's ARC in examining an airline's effort (Customer First Plan) in promoting passenger service.

References

Berg, K. E. & Latin, R. W. (1994). *Essentials of modern research methods: In health, physical education, and recreation.* Englewood Cliffs, NJ: Prentice-Hall.

Bowen, B. D. & Headley, D. E. (2000). *Airline Quality Rating 2000*. University of Nebraska at Omaha, Omaha, NE.

Bowen, B. D. & Headley, D. E. (2001). *Airline Quality Rating 2001*. University of Nebraska at Omaha, Omaha, NE.

Bowen, B. D., & Headley, D. E. (2002). *Airline Quality Rating 2002*. University of Nebraska at Omaha, Omaha, NE.

Bowen, B. D. & Headley, D. E. (2003). *Airline Quality Rating 2003*. University of Nebraska at Omaha, Omaha, NE.

Bowen, B. D., Headley, D. E., & Luedtke, J. (1992). A quantitative methodology for measuring airline quality. *Journal of Aviation/Aerospace Education and Research, 2* (2), 27-33.

Bowen, B. D., Headley, D. E., & Lutte, R. (1993). The airline quality rating: Developing an industry standard. *The Journal of Aviation/Aerospace Education & Research, 4* (1), 33-39.

Bowen, B. D., Headley, D. E., & Kane, K. (1998). Enhancing global competitiveness: Benchmarking airline operational performance in highly regulated environments. *Proceeding of the ATRG Extension Symposium, 10pps*, University of Nebraska at Omaha.

Bowen, B. D., Headley, D. E., Kane, K., & Lutte, R. (1999). Enhancing global competitiveness: Benchmarking airline operational performance in highly regulated environments. *Collegiate Aviation Review, 17* (1), 9-17.

Bowen, B. D., Headley, D. E., & Lu, C-t. (2003). Benchmarking airline service performance: A reference point for airline and government evaluation. *Public Works Management & Policy, 7* (3), 188-204.

Camp, R. C. (1989). *Benchmarking: The search for industry best practices that lead to superior performance*. White Plains, NY: Quality Press.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.

D O T Aviation Consumer Protection Division (n.d). *Airlines consumer service plans*. Data retrieved August 31, 2004 from http://airconsumer.ost.dot.gov/customerservice.htm

Federal Aviation Administration (FAA). (2001). *FAA aerospace forecast – fiscal years 2001-2012*. Washington, DC: Author.

Fitz-enz, J. (1993). *Benchmarking staff performance: How staff departments can enhance their value to the customers*. San Francisco, CA: Jossey-Bass.

Goodman, J. (1992, April 29). Airline watchdog is fly-by-night (and day) scholar. *Chronicle of Higher Education, 38* ,A5.

Headley, D. E. & Bowen, B. D. (1997). International airline quality measurement. *Journal of Air Transportation World Wide ,2* (1), 53-63.

Keehley,P., Longmire, L., Medlin, S. M., & MacBride, S. (1997). *Benchmarking for best practices in the public sector.* San Francisco, CA: Jossey-Bass.

Lincoln, Y. S. & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage

Mann, P. (2000). Airlines upbraided for self-defeating practices. *Aviation Week & Space Technology, 153* (12), 61.

Neufville, R., & Guzmán, J. R. (1998, July/August). Benchmarking for design of major airports worldwide. *Journal of Transportation Engineering, 124* (4), 391-395.

Northwest Airlines (n.d.). *Customer service plan*. Data retrieved August 31, 2004 from http://www.nwa.com/plan/

Office of Inspector General, Department of Transportation. (2001, February 12). *Final report on airline consumer service commitment*. Washington, DC: Author.

Patterson, J. (1996). *Benchmarking basics: Looking for a better way*. Menlo Park, CA: Crisp.

Spencer,P. (1999). Airline quality ratings. *Consumers' Research Magazine, 82* (5), 43.

Tucker, S. (1996). *Benchmarking: A guide for educators*.Thousand Oaks, CA: Corwin Press.

Appendix 1

Airline Quality Rating Criteria Overview

The individual criteria used to calculate the AQR scores are summed up in four basic categories that reflect customer-oriented areas of airline performance. Definitions of the four categories used in this AQR 2001 (2000 data) are outlined below.

OT:On-Time Performance (+8.63)

Regularly published data regarding on-time arrival performance is obtained from the U.S. Department of Transportation's *Air Travel Consumer Report*. According to the DOT, a flight is counted on-time if it is operated within 15 minutes of the scheduled time shown in the carriers' Computerized Reservations Systems. Delays caused by mechanical problems are counted as of January 1, 1995. Canceled and diverted operations are counted as late. The AQR calculations use the percentage of flights arriving on time for each airline for each month.

DB:Involuntary Denied Boardings (-8.03)

This factor includes involuntary denied boardings. Data regarding denied boardings is obtained from the U.S. Department of Transportation's *Air Travel Consumer Report*. The data includes the number of passengers who hold confirmed reservations and are involuntarily denied boardings on a flight that is oversold. These figures include only passengers whose oversold flight departs without them onboard. The AQR uses the ratio of involuntary denied boardings per 10,000 passengers boarded by month.

MB: Mishandled Baggage Reports (-7.92)

Regularly published data regarding consumer reports to the carriers of mishandled baggage is obtained from the U.S. Department of Transportation's *Air Travel Consumer Report*. According to the DOT, a mishandled bag includes claims for lost, damaged, delayed, or pilfered baggage. Data is reported by carriers as to the rate of mishandled baggage reports per 1,000 passengers and for the industry. The AQR ratio is based on the total number of reports each major carrier received from passengers concerning lost, damaged, delayed, or pilfered baggage per 1,000 passengers served.

CC: Customer Complaints (-7.17)

The criteria of customer complaints is made up of 12 specific complaint categories (outlined below) monitored by the U. S. Department of Transportation and reported monthly in the *Air Travel Consumer Report*. Customers can file complaints with the DOT in writing, by telephone, via e-mail, or in person. The AQR uses complaints about the various categories as part of the larger customer complaint criteria and calculates the customer complaint ratio on the number of complaints received per 100,000 passengers flown.

Flight Problems

Data are available by the total number of customer complaints pertaining to cancellations, delays, or any other deviations from schedule, whether planned or unplanned for each airline each month.

Oversales

This complaint category includes all bumping problems, irrespective of whether the airline complied with DOT oversale regulations. Data is available by the total number of customer complaints pertaining to oversales for each airline each month.

Reservations, Ticketing, and Boardings

This category includes airline or travel agent mistakes in reservations and ticketing; problems in making reservations and obtaining tickets due to busy telephone lines or waiting in line or delays in mailing tickets; and problems with boardings the aircraft (except oversales). Data is available by the total number of customer complaints pertaining to ticketing and boardings for each airline each month.

Fares

As defined by the DOT, customer complaints about fares include incorrect or incomplete information about fares, discount fare conditions and availability, overcharges, fare increases, and level of fares in general. Data is available by the total number of customer complaints pertaining to fares for each airline each month.

Refunds

This category includes customer complaints about problems in obtaining refunds for unused or lost tickets, fare adjustments, or bankruptcies. Data is available by the total number of customer complaints pertaining to refunds for each airline each month.

Baggage

Claims for lost, damaged, or delayed baggage; charges for excess baggage, carry-on problems, and difficulties with airline claim procedures; are included in this category. Data is available by the total number of customer complaints pertaining to baggage for each airline each month.

Customer Service

This category includes complaints about rude or unhelpful employees, inadequate meals or cabin service, and treatment of delayed passengers. Data is available by the total number of customer complaints pertaining to customer service for each airline each month.

Disability

Previously included as part of the Reservations, Ticketing and Boardings Category (thru 6/99), this category includes complaints about civil rights violations from air travelers with disabilities. Data is available by the total number of customer complaints pertaining to disabilities for each airline each month.

Advertising

These are complaints concerning advertising that is unfair, misleading, or offensive to customers. Data is available by the total number of customer complaints regarding advertising for each airline each month.

Tours

This category includes complaints about problems with scheduled or charter tour packages. Data is available by the total number of customer complaints pertaining to tours for each airline each month.

Animals

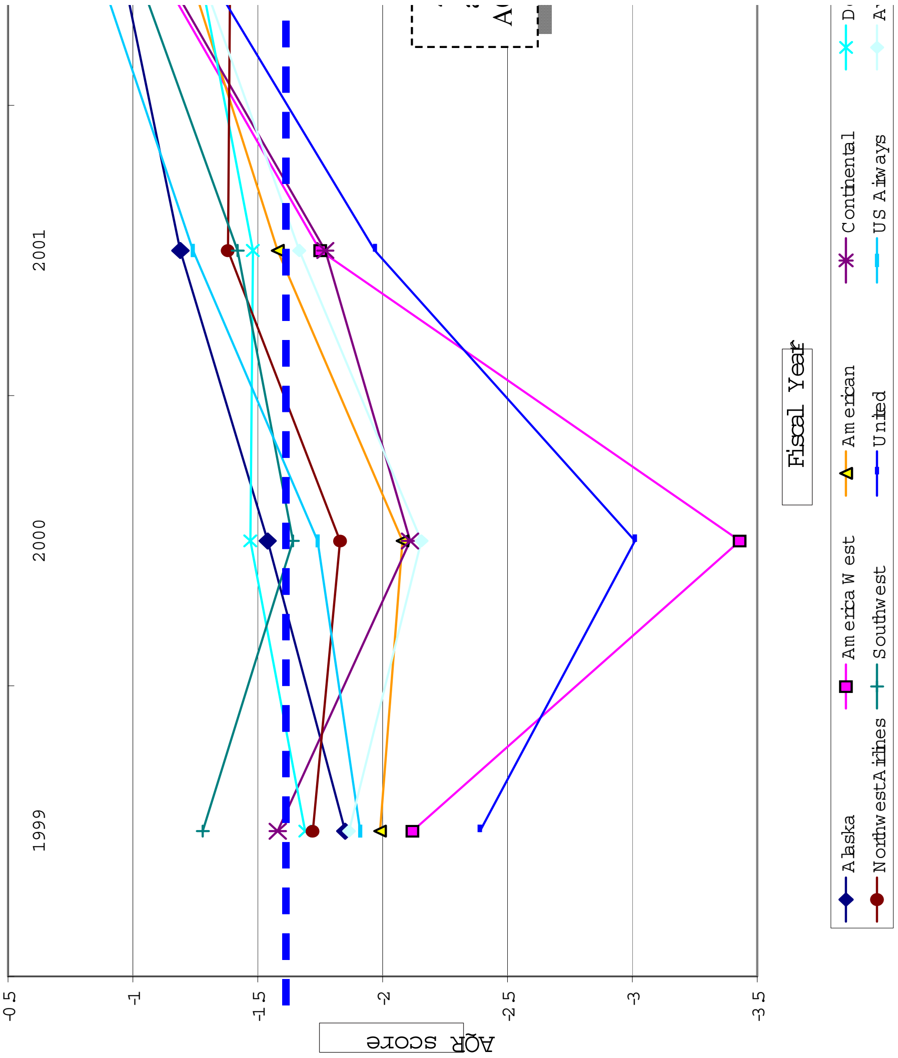
This category, added in October 2000, tracks customer complaints about loss, injury, or death of an animal during air transport by an air carrier. Data is available by the total number of customer complaints regarding animals for each airline each month.

Other

Data regarding customer complaints about frequent flyer programs, smoking, credit, cargo problems, security, airport facilities, claims for bodily injury, and other problems not classified above are included in this category. Smoking and credit, previously separate elements, were added to this general category as of 9/99. Data is available by the total number of customer complaints regarding other problems for each airline each month.

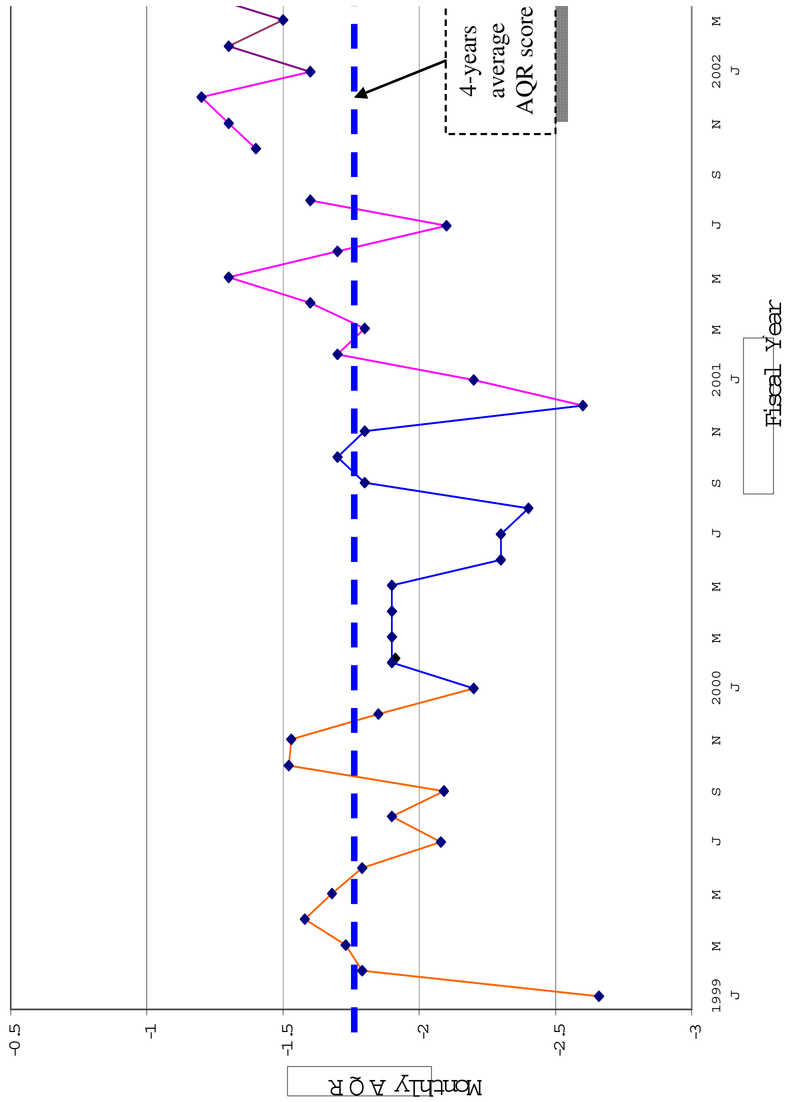
Appendix 2

Benchmarked AQR Scores – 2000 ~ 2003 Reports



Appendix 3

AQR Industry Scores (by month) – 1999 - 2002



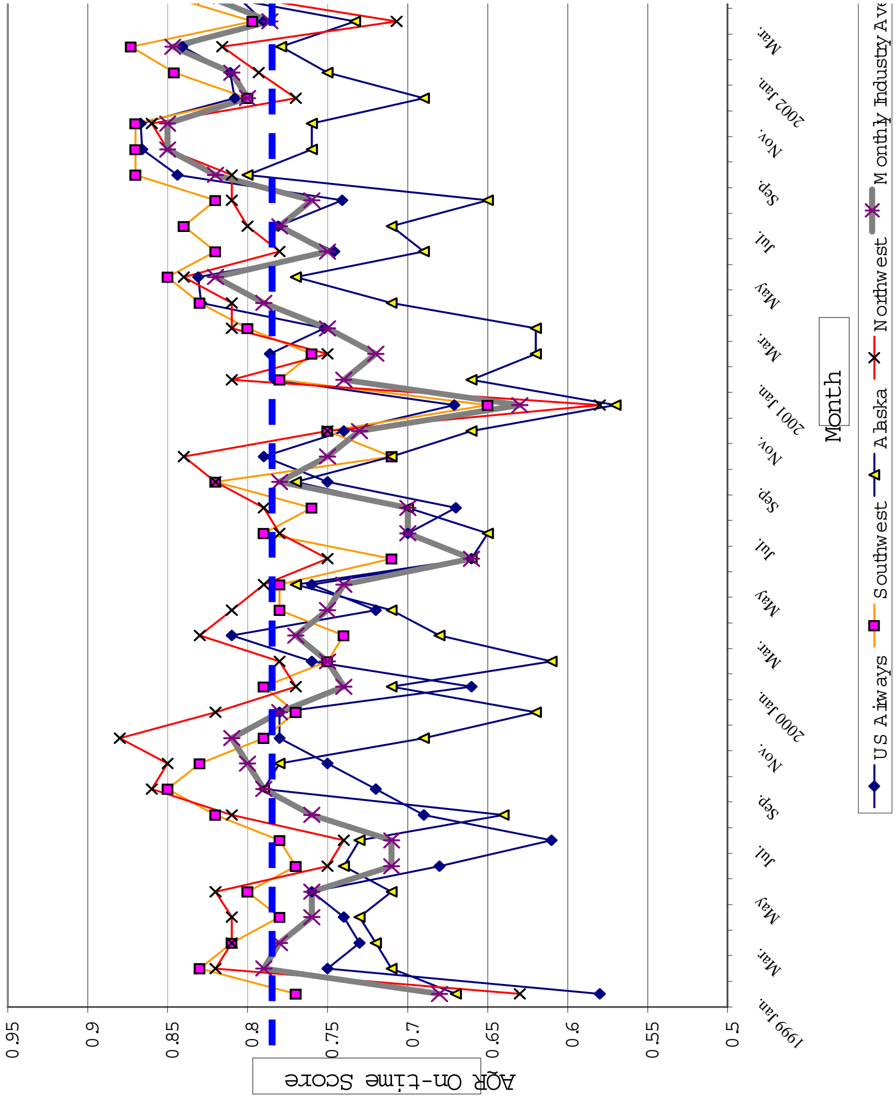
Appendix 4-1

On-Time Rate – Four-Year Benchmarking Analysis (1999 - 2002)

	1999 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oc
US Airways	0.58	0.75	0.73	0.74	0.76	0.68	0.61	0.69	0.72	0.7
Southwest	0.77	0.83	0.81	0.78	0.80	0.77	0.78	0.82	0.85	0.8
Alaska	0.67	0.71	0.72	0.73	0.71	0.74	0.73	0.64	0.79	0.7
Northwest	0.63	0.82	0.81	0.81	0.82	0.75	0.74	0.81	0.86	0.8
Monthly Industry Average	0.68	0.79	0.78	0.76	0.76	0.71	0.71	0.76	0.79	0.8
	2000 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oc
US Airways	0.66	0.76	0.81	0.72	0.76	0.66	0.70	0.67	0.75	0.7
Southwest	0.79	0.75	0.74	0.78	0.78	0.71	0.79	0.76	0.82	0.7
Alaska	0.71	0.61	0.68	0.71	0.77	0.66	0.65	0.70	0.77	0.7
Northwest	0.77	0.78	0.83	0.81	0.79	0.75	0.78	0.79	0.82	0.8
Monthly Industry Average	0.74	0.75	0.77	0.75	0.74	0.66	0.70	0.70	0.78	0.7
	2001 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oc
US Airways	0.783	0.79	0.75	0.83	0.83	0.75	0.78	0.74	0.84	0.8
Southwest	0.78	0.76	0.80	0.83	0.85	0.82	0.84	0.82	0.87	0.8
Alaska	0.66	0.62	0.62	0.71	0.77	0.69	0.71	0.65	0.80	0.7
Northwest	0.81	0.75	0.81	0.81	0.84	0.78	0.80	0.81	0.81	0.8
Monthly Industry Average	0.74	0.72	0.75	0.79	0.82	0.75	0.78	0.76	0.82	0.8
	2002 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oc
US Airways	.811	.841	.790	.807	.813	.810	.832	.859	.909	.87
Southwest	.846	.873	.797	.850	.825	.787	.779	.815	.886	.82
Alaska	.750	.779	.733	.820	.809	.746	.771	.760	.851	.80
Northwest	.793	.816	.707	.806	.811	.774	.779	.818	.870	.86
Monthly Industry Average	.810	.847	.786	.826	.828	.786	.798	.829	.880	.84

Appendix 4-2

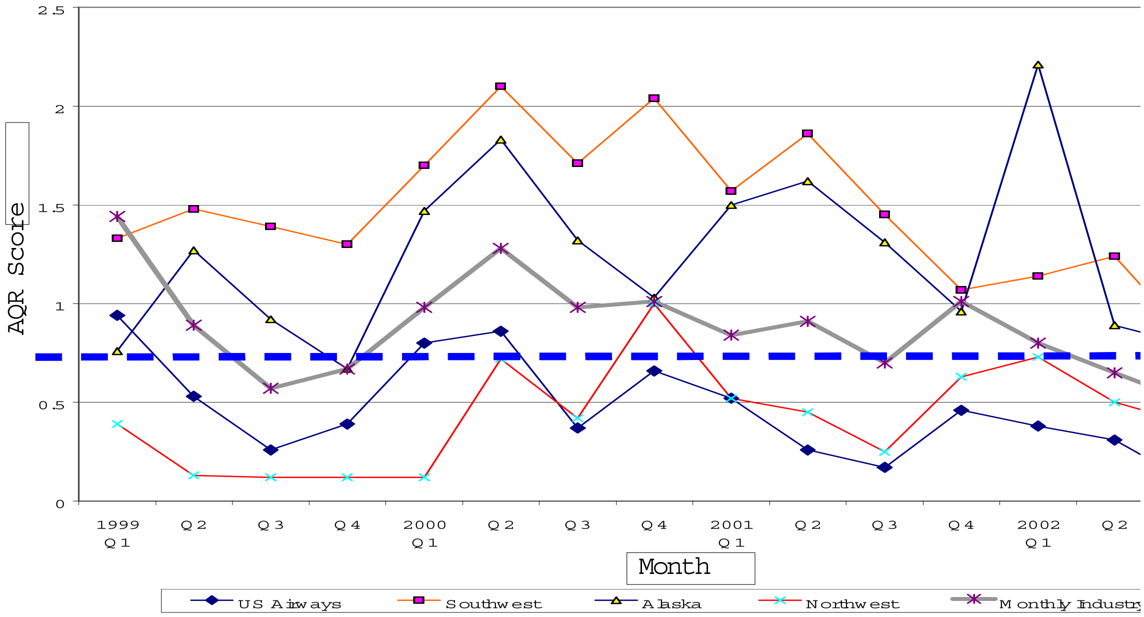
On-Time Performance



Appendix 5

Denied Boardings - Four-Year Benchmarking Analysis (quarterly, per 10,000 passengers)

	1999 Q1	Q2	Q3	Q4	2000 Q1	Q2	Q3	Q4	2001 Q1	Q2	Q3	Q4	2002 Q1
US Airways	0.94	0.53	0.26	0.39	0.80	0.86	0.37	0.66	0.52	0.26	0.17	0.46	0.38
Southwest	1.33	1.48	1.39	1.30	1.70	2.10	1.71	2.04	1.57	1.86	1.45	1.07	1.14
Alaska	0.76	1.27	0.92	0.67	1.47	1.83	1.32	1.03	1.50	1.62	1.31	0.96	2.21
Northwest	0.39	0.13	0.12	0.12	0.12	0.72	0.42	1.00	0.52	0.45	0.25	0.63	0.73
Monthly Industry Average	1.44	0.89	0.57	0.67	0.98	1.28	0.98	1.01	0.84	0.91	0.70	1.01	0.80



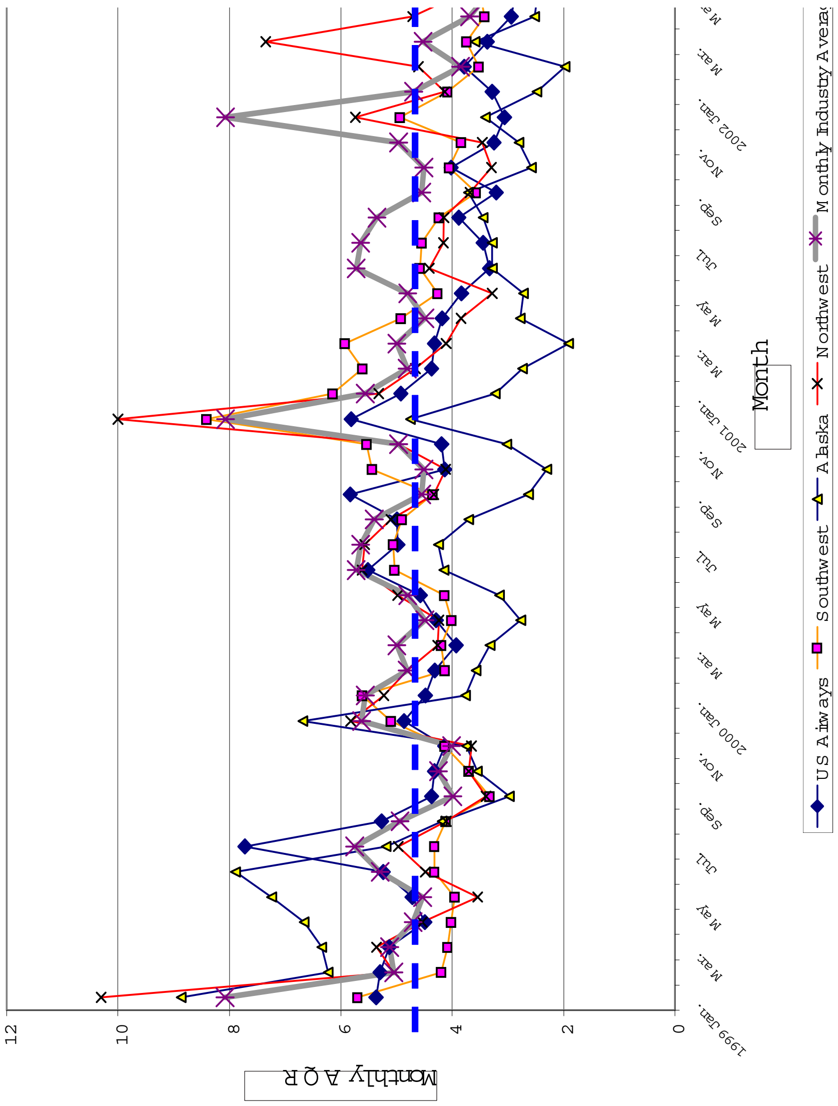
Appendix 6-1

Mishandled Baggage – Four-Year Benchmarking Analysis
(per 1,000 passengers)

	1999 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.
US Airways	5.37	5.29	5.13	4.49	4.72	5.24	7.72	5.27	4.37	4.32
Southwest	5.7	4.2	4.08	4.02	3.95	4.32	4.32	4.12	3.33	3.70
Alaska	8.87	6.23	6.34	6.66	7.24	7.89	5.19	4.18	2.97	3.55
Northwest	10.3	5.04	5.36	4.54	3.54	4.48	4.97	4.11	3.39	3.70
Monthly Industry Average	8.08	5.05	5.12	4.7	4.53	5.29	5.75	4.94	3.99	4.25
	2000 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.
US Airways	4.48	4.31	3.93	4.29	4.57	5.51	4.98	4.99	5.83	4.13
Southwest	5.62	4.13	4.2	4.01	4.14	5.03	5.06	4.9	4.35	5.44
Alaska	3.76	3.57	3.32	2.77	3.15	4.15	4.25	3.7	2.63	2.3
Northwest	5.23	4.81	4.26	4.24	4.98	5.62	5.57	5.1	4.33	4.12
Monthly Industry Average	5.56	4.81	4.99	4.49	4.8	5.72	5.64	5.4	4.55	4.51
	2001 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.
US Airways	4.92	4.37	4.32	4.18	3.83	3.33	3.44	3.88	3.21	4.02
Southwest	6.15	5.61	5.93	4.92	4.26	4.58	4.55	4.24	3.57	4.05
Alaska	3.22	2.74	1.91	2.78	2.72	3.28	3.28	3.44	3.71	2.57
Northwest	5.32	0.66	4.11	3.84	3.28	4.41	4.16	4.15	3.68	3.3
Monthly Industry Average	5.56	4.81	4.99	4.49	4.8	5.72	5.64	5.35	4.55	4.51
	2002 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.
US Airways	3.28	3.78	3.37	2.94	2.86	2.64	2.71	2.81	2.74	2.47
Southwest	4.08	3.52	3.74	3.42	3.50	3.69	3.74	3.40	2.81	3.10
Alaska	2.48	1.97	3.59	2.52	2.46	2.95	3.03	3.04	2.51	2.51
Northwest	4.13	4.60	7.35	4.71	3.74	5.37	5.10	4.36	3.15	3.08
Monthly Industry Average	4.69	3.85	4.52	3.69	3.32	4.03	3.99	3.72	3.04	4.11

Appendix 6-2

Mishandled Baggage



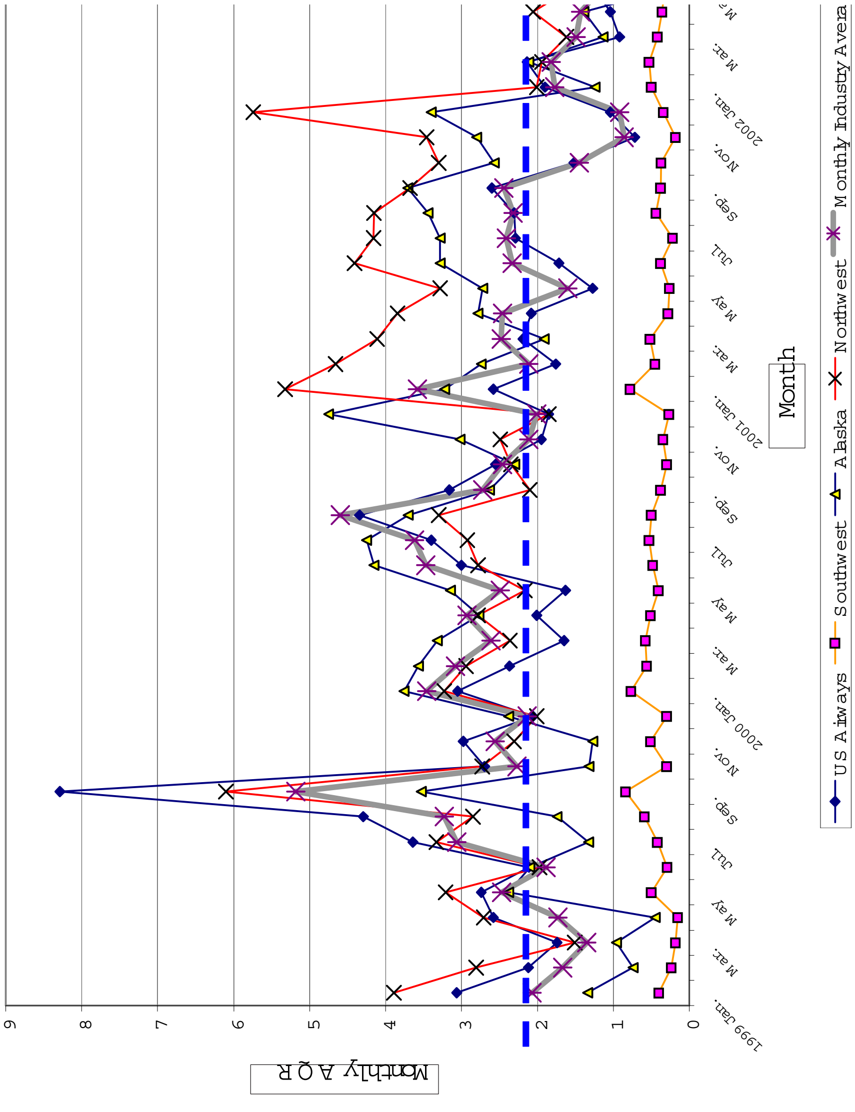
Appendix 7-1

Customer Complaints – Four-Year Benchmarking Analysis
(per 100,000 passengers)

	1999 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oc
US Airways	3.06	2.12	1.74	2.58	2.74	2.11	3.64	4.29	8.29	2.
Southwest	0.4	0.24	0.18	0.15	0.5	0.29	0.42	0.59	0.84	0.
Alaska	1.34	0.74	0.96	0.45	2.38	2.1	1.33	1.74	3.53	1.
Northwest	3.89	2.81	1.51	2.71	3.21	1.97	3.33	2.85	6.1	2.
Monthly Industry Average	2.07	1.67	1.35	1.73	2.47	1.89	3.06	3.23	5.18	2.
	2000 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oc
US Airways	3.05	2.37	1.65	2.01	1.63	3.00	3.4	4.34	3.16	2.
Southwest	0.77	0.56	0.58	0.51	0.41	0.48	0.53	0.5	0.38	0.
Alaska	3.76	3.57	3.32	2.77	3.15	4.15	4.25	3.7	2.63	2.
Northwest	3.23	2.94	2.36	2.78	2.17	2.78	2.92	3.30	2.10	2.
Monthly Industry Average	3.46	3.08	2.61	2.93	2.49	3.47	3.62	4.6	2.72	2.
	2001 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oc
US Airways	2.58	1.76	2.19	2.08	1.27	1.72	2.29	2.31	2.60	1.
Southwest	0.78	0.45	0.52	0.28	0.26	0.38	0.22	0.44	0.38	0.
Alaska	3.22	2.74	1.91	2.78	2.72	3.28	3.28	3.44	3.71	2.
Northwest	5.32	4.66	4.11	3.84	3.28	4.41	4.16	4.15	3.68	3.
Monthly Industry Average	3.58	2.11	2.48	2.46	1.6	2.33	2.41	2.32	2.44	1.
	2002 Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oc
US Airways	1.90	2.14	0.92	1.04	1.26	1.03	1.41	0.73	0.90	0.
Southwest	0.50	0.53	0.42	0.36	0.32	0.47	0.32	0.29	0.13	0.
Alaska	1.24	2.11	1.14	1.39	0.17	1.23	0.93	0.81	0.46	0.
Northwest	2.01	1.94	1.62	2.06	1.52	1.57	1.63	1.26	0.94	1.
Monthly Industry Average	1.77	1.82	1.49	1.43	1.14	1.26	1.41	0.99	0.86	0.

Appendix 7-2

Customer Complaints



International Journal of Applied Aviation Studies, 4, Number 2

Passenger Use and Perception of Personal Electronic Equipment On Board Aircraft

Nora Balfe and Tony Head

College of Aeronautics
Cranfield University
Bedfordshire, MK43 0AL
United Kingdom
Telephone: 01234 750111, Extension 4015
Email: t.head@cranfield.ac.uk

Abstract

Personal electronic devices (PEDs) are becoming increasingly common on board aircraft, as they are in all aspects of life. There are concerns that the electronic noise emitted by such devices may interfere with sensitive aircraft equipment, particularly navigation equipment. Although it has not been possible to prove conclusively that PEDs interfere with avionics, there is strong anecdotal evidence that they have already done so, and experts agree that such interference is theoretically possible. Thus, the safest course of action is to restrict the use of PEDs during aircraft operation. Generally, airlines ban the use of PEDs that transmit a signal, and only allow operation of other devices when the aircraft altitude is greater than 10,000ft.

A survey was conducted at three airports that assessed the PED usage level, knowledge of regulations regarding PED, and participants' perception of the dangers involved in PED usage. The survey revealed that 90% of passengers carry PEDs on board aircraft, with mobile phones being the most common device carried. Participants did not follow the regulations 100% of the time, particularly with regard to mobile phones. Participants were more confused about the regulations and dangers surrounding unintentional transmitters as compared to intentional transmitters. Interestingly, inconsistent airline policies and lack of education regarding PEDs were considered the most likely reasons for this confusion.

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

Introduction

Personal electronic devices (PEDs), such as mobile phones, laptop computers, portable games machines, and portable music players have become increasingly common parts of everyday life and consequently more common on civil transport aircraft. The electromagnetic noise routinely emitted by these devices is suspected of interfering with sensitive aircraft equipment. Although several studies have attempted yet failed to recreate interference reported by pilots, strong anecdotal evidence suggests that PEDs can cause interference, and experts agree that this is possible, particularly with higher power devices (such as mobile phones) that transmit intentional signals.

PEDs can be divided into two separate categories, intentional transmitters (mobile phones, pagers, and remote control toys that must transmit a signal in order to complete their function) and unintentional transmitters (including laptop computers, music players, and cameras that do not need to transmit a signal to operate but may still emit some level of electromagnetic radiation) (Subcommittee on Aviation, 2000). Electromagnetic interference is caused by waves of energy that all electrical devices emit where energy from one device is radiated into another and causes a malfunction (Strauss, 2001).

Within an aircraft the electronic noise emitted by an electronic device rebounds from the inside walls of the fuselage as the metallic fuselage acts as a shield that prevents the signal from escaping (Ladkin, 1997). Complex propagation paths are set up within the confines of the fuselage that lead to signal cancellation or intensification at different locations (CAA, 2000). These signals can escape from the fuselage through window's, hatch seams, and increasingly through composite components (Li, Xie, Ramahi, Pecht, & Donham 2000; Air Safety Week, 1996). Once outside the fuselage, signals may be picked up by externally mounted antennas and interfere with aircraft systems. The navigation and communication systems are particularly vulnerable to this interference, since they are designed to acquire and respond to relatively weak signals (Ladkin, 1997; Perry & Geppert, 1996). Interference is most likely to occur when the interfering signal is at the same frequency as the 'real' signals. However, there are circumstances in which other frequencies can 'break through', particularly if the interfering signal is strong.

At present, most PEDs operate at lower frequencies than aircraft systems; however, there is a clear trend in PED development towards the use of processors operating at higher frequencies (Muller, 2002). Additionally, PEDs may interfere with aircraft systems via harmonics that are multiple emissions of the original frequency of the PED (Ritchie, 1996). Consequently, a laptop computer operating at 55MHz will produce energy at odd harmonics i.e. 165MHz, 275MHz, 385MHz etc. Each successive harmonic becomes weaker than the previous, yet may still retain sufficient power to be recognized by a system whose frequency it overlaps, resulting in interference (this is due to its 'square clock' function) (Helfrick, 1994). Given the number of different PEDs possibly carried on an aircraft, their array of frequencies and associated harmonics, the resulting emitted energy covers almost the entire range of navigation and communication frequencies used by aircraft (Perry & Geppert, 1996).

With the development of new and more powerful PED technologies operating at higher frequencies, the possibility of interference and the range of systems that may be affected are increased (Donham, 2000; Air Safety Week, 2003). While it is unlikely that PED interference alone would ever directly cause a crash, it is not improbable that PED interference could combine with other factors to create a chain of incidents or factors that together result in an incident or accident.

Research

Although there is little hard evidence of PED interference to aircraft systems, previous studies have suggested the possibility.The United Kingdom (UK) Civil Aviation Authority (CAA) has conducted two such studies. The first study examined the potential for interference from mobile telephones (CAA, 2000). It found that transmissions made from mobile telephones in the cabins of two different civil transport aircraft could induce interference to a level greater than the demonstrated susceptibility levels for aircraft equipment approved to pre-1984 standards.

The second CAA study examined the effects of interference from mobile telephones specifically on aircraft avionics equipment (CAA, 2003). It found that mobile telephones were capable of creating a level of interference resulting in:

- Compass freezing or overshooting actual magnetic bearing.
- Instability of indicators.
- Digital VOR navigation bearing display errors of up to 5 degrees.
- VOR navigation To/From indicator reversal.
- VOR and ILS course deviation indicator errors with and without a failure flag.
- Reduced sensitivity of the ILS receiver.
- Background noise on audio outputs.

These studies also determined that interference from a PED could result in false warnings of unsafe conditions, increased crew workload, reduced crew confidence in protection systems, crew distraction, and noise in crew headphones. Any of these could become a contributing factor to an incident or accident.

The Radio Technical Commission for Aeronautics (RTCA) has performed three studies to date between 1963 and 1996 to investigate the potential for interference from personal electronic devices. The most recent research by the RTCA studied unintentional transmitters. Even though the research could not duplicate interference from a PED under controlled conditions, the report stated that the probability of interference, though extremely slight, should be considered as potentially hazardous and clearly represents an unacceptable risk during critical phases of flight (RTCA, 1996). However, not all members of the committee agreed that the potential for interference was slight and some argued that the potential may have been shown as significant had different testing methods been employed (Shawlee, 2000). Finally, this report stated that emissions from new technologies such as Ultrawideband (e.g. as used in wireless computing) could exceed interference limits for aircraft by a factor of 1000 (Ely, Fuller, & Shaver, 2002).

The most compelling evidence supporting the theory that PEDs can cause aircraft interference comes from NASA's Aviation Safety Reporting System (ASRS). A NASA report (Ross, 2001) reviewed 14 years of incident reports and found over 100 reports of possible PED interference, that the pilots believed to be due to a PED on board the aircraft. A wide variety of PEDs was suspected of causing interference; however, mobile telephones and laptop computers were the most frequently cited with 25 references each. Navigation systems were the most commonly affected, confirming that this system is particularly vulnerable to interference. Some reports cited interference at a critical stage of flight, indicating that the possibility of a PED contributing to a crash is very real.

Another review of the Aviation Safety Reporting System (ASRS) data in 2002 (Strauss, 2002) found 125 reported incidents. Again, the type of PED suspected of causing interference was varied. In 57 of these incidents, the interference disappeared as the suspected device was switched off, providing reasonable evidence that the PED was the cause. The most frequently affected aircraft systems were navigation, Instrument Landing, and communications, with mobile phones and laptop computers accounting for 60% of reported incidents. Almost half of the reported incidents occurred within the 'sterile cockpit' phase of flight, creating an unnecessary distraction for pilots at a critical time. The number of reports, together with the strong relationships between the interference and the PEDs in use, suggest that PEDs indeed have the potential to interfere with aircraft systems.

Regulations

The FAA regulations FAR 91.21, 121, and 135 stipulate that no portable electronic devices be operated on board an aircraft registered in the United States (US), excluding portable voice recorders, hearing aids, heart pacemakers and electric shavers, and any other devices that the operator of the aircraft has determined will not cause interference.

In addition to the regulations, the FAA also publishes an Advisory Circular, 91.21-1A that advises operators on compliance with the regulations. It suggests that unintentional transmitters can be used when the aircraft is safely in cruise, since interference occurring at this altitude would allow sufficient time to isolate the source of interference and switch it off. It also states that operation of any intentional transmitters should be prohibited while airborne but permitted when the aircraft is on the ground.

In the U.S., mobile phone use on aircraft is specifically prohibited by the Federal Communications Commission (FCC). This is due to the greater transmitting range that the phone is capable of when airborne, resulting in possible interference to the telephone network (Perry & Geppert, 1996; Ritchie, 1996).

When operating on the ground, a mobile phone can 'see' one or two masts at a time and it transmits to these. When operating in the air, mobile phones can 'see' many different cells, due to both the height of the aircraft and the speed at which it is travelling. If the phone makes contact with many cells at the same time, the result could be serious interference to the mobile telephone network, resulting in interruption to other users' calls (Li et al, 2002).

In Europe, PEDs are regulated by JAR-OPS (Joint Aviation Authority Operating Regulations) 1.110 that states:

> An operator shall not permit any person to use, and take all reasonable measures to ensure that no person does use, on board an aeroplane, a portable electronic device that can adversely affect the performance of the aeroplane's systems and equipment. (JAR-OPS 1.110)

Leaflet 29 published by the Joint Aviation Authority (JAA) discusses the dangers of PED use and recommends a policy for airlines to follow that includes prohibiting use of intentional transmitters from the closing of the aircrafts doors until they are reopened. It further recommends that non-intentional transmitters be switched off during critical flight phases. Furthermore, it recommends that announcements regarding PEDs be made both prior to and during boarding of the aircraft.

Both sets of regulations make the aircraft operator responsible for deciding which devices are safe for use on the aircraft despite the fact that airlines do not have the necessary resources to conduct testing (Strauss, 2001). Although the Federal Aviation Authority (FAA) and JAA regulations are similar, there is no global policy on the use of PEDs (Jensen, 2000) and the result is an issue on which there is no comprehensive and coherent policy.

Airline Regulations

The regulatory bodies, both in the UK and the US, have not produced precise regulations on PED use in aircraft, and because of this, the airlines are required to devise their own regulations on PEDs.

The regulations for PED usage vary across different airlines although many resemble the policy outlined in JAA Leaflet 29. Most airlines divide PEDs into two groups, those that are forbidden at all times and those that are forbidden only during some stages of flight. However, some airlines forbid use of *any* electronic device without the prior permission of the airline. Some airlines add a third category of devices that may be operated throughout the flight; these include pacemakers, electronic watches, and electric shavers (Subcommittee on Aviation, 2000; Ritchie, 1996).

Legal

While there are no specific laws relating to the use of PED on board aircraft in the UK, passengers who refuse to switch off PEDs when asked to do so by the crew have been convicted under the Air Navigation Order (ANO) of "recklessly and negligently endangering" the aircraft (BBC, 2002; BBC, 1999). Article 55 of the ANO states, "A person shall not recklessly or negligently act in a manner likely to endanger an aircraft or any person therein." Article 59 of the ANO may also be relevant; it states the following:

> Every person in an aircraft registered in the UK shall obey all lawful commands which the commander of that aircraft may give for the purpose of securing the safety of the aircraft and of persons or property carried therein, or the safety, efficiency or regularity of air navigation.

The maximum penalty for contravening Article 55 of the ANO is a £2000 fine and/or a two-year prison sentence, although the maximum penalty that has ever been imposed as a result of refusing to switch off a PED in the UK was a one-year prison sentence (BBC, 1999).

Study Rationale and Objectives

As outlined above the regulations governing the use of PED are vague, resulting in different policies implemented in different ways on different airlines, ultimately leading to confusion for passengers (Ritchie, 1996). Some airlines have power outlets for laptop computers, while others do not allow CD players to be operated on board (Shawlee, 1999). Added to this is the fact that some magazines and newspapers challenge the ban on mobile telephone use on aircraft, citing the lack of hard evidence supporting interference (Bicknell, 2000). Many apparently believe that the airlines' ban on mobile telephone usage is based on a desire on the part of the airlines to increase the revenues of the airlines' own seat-back telephones (Bicknell, 2000; Aberbach, 1999; BBC, 2000). A Wall Street Journal article (Aberbach, 1999) quoted a pilot who admits to regularly using his mobile telephone while in flight and is convinced that the ban is bogus.

This confusion and erroneous reporting may explain why there are many reports of passengers being reluctant to switch off mobile telephones and surreptitiously using them and other transmitters on board aircraft (Evans, 1999). One unofficial survey in the US found that 6% of all passengers use their PEDs on domestic flights and this number is undoubtedly growing with the proliferation of PEDs (Rossier, 2000). It may also help to explain why passenger use of PEDs has been reported as the second largest cause of air rage in the US (Subcommittee on Aviation, 2000). Passengers may not understand the potential dangers of PED use and, therefore, consider it unfair to be asked to switch them off. Many passengers find it hard to believe that something as small and innocuous looking as their mobile telephone or portable games machine could affect the operation of the aircraft (Rossier, 2000).

A relatively large amount of research has been conducted into the technical aspect of PED operation on aircraft and has examined the potential for interference, the possible systems affected, and the consequences of interference. However, whilst no study has proven conclusively that PEDs are hazardous to the safety of the aircraft, neither has it been proven that they are not. Expert researchers have agreed that, while it has not been possible to reproduce interference from a PED after a suspected incident, the potential exists for interference, although it may be slight. In light of this, the safest course of action is the one that airlines have been following, namely to restrict the use of PEDs on aircraft and to prohibit the use of some of the higher powered devices and intentional transmitters.

In spite of the technical research, the amount of literature available regarding passenger use and perception of PEDs is minimal. The number of PEDs carried on aircraft is unknown, as is the number operated on board. The level of passenger awareness of the problems posed by PEDs and their desire to use PEDs on aircraft are further unknowns. This study aimed to gather quantitative data regarding passenger use of PEDs on commercial aircraft and further aimed to contribute to the overall understanding of the level of risk that is involved in PED usage. Information gathered regarding passenger knowledge of the dangers and their perception of the risk could be used in future to develop a programme that aims to educate passengers, should that be seen as necessary.

The objectives of this study were to accomplish the following:

1. To establish the number and type of PEDs carried on board aircraft by passengers, and the level of use of these during the flight.
2. To determine passenger knowledge of the regulations governing PED operation on board an aircraft.
3. To determine the level of knowledge of the hazards associated with PED operation on board an aircraft among airline passengers.
4. To determine passenger perceptions of the level of risk posed by PEDs.
5. To determine the level of passenger desire to use PEDs on board aircraft.

Method

The study aims were achieved through an interviewer-administered questionnaire delivered to arriving passengers. The questionnaire was divided into six sections, the first gathered demographics and the following five sections each dealt with a different study objective. The demographics gathered included participants' age group, gender, airline flown, journey type (i.e. long-haul, short-haul, or domestic), and the reason for travel (i.e. business or leisure). These details were taken to determine whether there were any significant relationships between passenger demographics and the use of portable electronic devices.

The first section of the questionnaire investigated the number and type of PEDs that each participant carried on the aircraft. For the purposes of the study, PEDs were defined as any device that is operated under battery power. Devices that were operated on board were noted. Finally, this section investigated when participants carrying mobile telephones had switched them off and switched them back on again. The second section investigated whether participants had ever forgotten to turn off their mobile telephone on a previous flight. It also enquired about the circumstances under which passengers considered it acceptable to use a mobile phone on an aircraft. The next section was related to the regulations governing the use of PEDs. It asked participants about when different types of PEDs should be switched off and about the maximum UK punishment for failing to switch off a PED when instructed to do so. Section 4 asked the participants about the possible adverse effects of PEDs on the aircraft systems. Participants were given a list of possible effects and asked to pick the reasons why PED use is restricted on aircraft. The final section of the questionnaire examined passengers' perceptions of the danger of PED usage along with their desire to use PEDs on board aircraft.

Two hundred questionnaires were administered at three London airports. Two of the airports were relatively small airports, with short-haul flights only, and the third was a large international airport. The two smaller airports provided 50 participants each and the remaining 100 were interviewed at the larger airport. All participants were arriving passengers and the sample for the study was an opportunity sample. The questionnaires were analysed using the Statistic Package for Social Sciences (SPSS). Frequencies were run for each question and graphs plotted of the responses. Independent t-tests were used to determine whether there were any significant differences between demographic groups' perception of the dangers and desire to use PEDs. Independent t-tests were deemed to be the appropriate instrument, as each category of passenger was independent and there were no repeated measures.

Results

Demographics. The majority (77.5%) of participants in this study were short-haul leisure passengers, due to the time of day and season in which the study was carried out. There was an even distribution of gender (49% male, 51% female), and a reasonable spread across the age groups (6.5% under 18, 32% 18-30, 43.5% 31-50, 17.5% 51-70, and 0.5% over 70).

Section One. Ninety percent of participants carried at least one PED on board the aircraft with them. Mobile phones were the most common device to be carried, with 77% of participants carrying one. Cameras were the second most common device. Other devices carried included laptop computers, music players, camcorders, electric shavers, radios, 'Palm Pilots', portable games machines, medical equipment, a Dictaphone, alarm clocks, and a printer. Music players were operated in 42.5% of cases where they were carried, but very few other devices were operated on board the aircraft.

Section Two. The majority of participants turned off their mobile telephones before leaving the airport terminal and did not switch them back on until after entering the destination terminal, but 6.5% of mobile phone carriers neglected to turn them off until just before take-off and 3.2% did not turn off their phones at all. One passenger (0.6%) turned a mobile phone back on during the flight. 13% of participants had forgotten to switch off their mobile telephones on a previous flight and 2% had used their mobile phones during a previous flight.

Participants were asked to consider the circumstances under which they would consider it acceptable to use a mobile telephone during flight. Participants were asked to choose between "always acceptable" and "never acceptable," or they could choose as many of the other three options as they wished. The three options were "if the aircraft was delayed in the air", "under emergency conditions," and "if the aircraft was hijacked." The majority of participants felt that it would be acceptable to use a mobile telephone if the aircraft was hijacked or under emergency conditions. The results are shown in Figure 1.

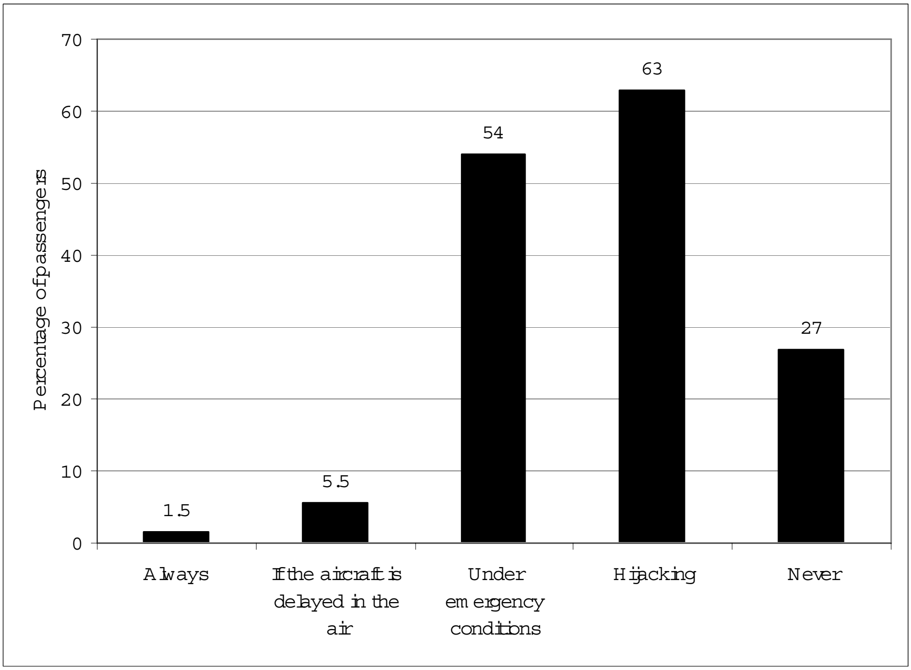


Figure 1. Acceptable circumstances to use mobile telephones.

Section Three. The third section of the questionnaire examined participants' knowledge of the regulations on board aircraft regarding both intentional and unintentional transmitters. 82.5% of participants knew that mobile telephones should be switched off at all times during the flight, but 5.5% thought that they only had to be switched off for take-off and landing and 12.1% did not know what the regulations were. With regard to unintentional transmitters, 42.4% of participants knew that these only had to be switched off during take-off and landing but 32.9% mistakenly believed that they could not be operated at any point during the flight. A much higher percentage (24.7%) did not know what the regulations for unintentional transmitters were, as compared to mobile telephones.

Participants were asked what they thought was the maximum UK legal punishment for failing to switch off a PED in an aircraft, when instructed to do so by the cabin crew. The correct answer was a £2000 fine and/or a 2-year prison sentence. Only 13.6% of participants answered this correctly. The majority thought that it was more likely to be a £500 fine and/or a 3-month prison sentence and 16.6% of participants thought that the punishment would be only a £100 fine.

Section Four. Participants were asked if they knew any of the reasons why mobile phone use is prohibited in aircraft. The majority of participants knew some reasons, the most commonly known was possible interference to the navigation systems (72.4%), but 22.9% did not know of any reason why mobile telephones were prohibited. Participants were more uncertain about why unintentional transmitter use was restricted with 45% stating that they did not know of any reasons for this.

Section Five. The final section of the questionnaire investigated participants' perceptions of PED dangers and participants' desire to use PED on board the aircraft. 77.4% of participants either agreed or strongly agreed with the statement "Mobile phones are dangerous if operated on board aircraft." There was less agreement and more uncertainty with regard to laptop computers. 21.1% of participants did not think that laptop computers were a danger to aircraft and 34.7% did not know. Opinion on whether participants would like to be able to use mobile phones during flight was divided, with 39.2% stating that they would like to be able to use them and 35.7% disagreeing. 25.1% were neutral. Participants were more agreeable to the idea of being able to use other devices like laptop computers and music players at all stages of flight with only 21.1% disagreeing and 20.6% neutral.

Discussion

Although a more methodical study sample incorporating proportional representation of e.g. ages, gender, business, and leisure passengers would have been desirable, due to security restrictions and time limitations only limited access to the airport arrival lounges was available. Therefore, it is recognised that the opportunity study sample of 200 passengers may not be representative of the total airline passenger population.

Section One

The results of the passenger survey carried out show that 90% of participants carried at least one personal electronic device on board the aircraft with them, and 58.5% carried more than one device. Clearly, these are very high numbers but they were tempered by the fact that only 11% of participants operated a device on board the aircraft, and these were all unintentional transmitters operated only during the cruise phase of flight. The number of devices operated did not include mobile phones that were not switched off, as participants did not regard leaving their mobile phone switched on as operating it!

Mobile phones were the most common device carried with 77% of all participants carrying one and 2.5% of participants did not turn off their mobile phones during the flight. Therefore, on an aircraft carrying 100 passengers, it is likely that two or three will have forgotten to switch off their mobile phones. Considering the research carried out by the UK CAA, which found that mobile phones are capable of interfering with aircraft equipment, and the data from the ASRS that lists mobile phones as the most common device to interfere, this is a worrying statistic. Although these mobile phones do not create interference in general, with two to three mobile phones switched on during each flight it must be regarded as inevitable that the circumstances for interference will line up eventually. Similarly, mobile phone use has not yet been implicated in an accident, but it can only be a matter of time before the interference combines with other factors to bring about an incident or accident. This may have happened already but there is no way to determine whether a mobile phone was switched on or operating on board the aircraft so it may never be proven.

Approximately 50% of the participants who did not turn off their phones had simply forgotten, but the remainder stated that they rarely turned off their phones for flight, as they did not regard it as dangerous. One participant stated that since there is no reception at high altitudes it did not matter if mobile phones were switched on once they were in the air. However, this belief is erroneous since a mobile phone that has no reception sends out signals at maximum power to attempt to reach a mast, thus making it more likely to interfere than if it were within signal range (Helfrick, 1994).

Cameras were the second most common device carried by participants, but this may have been influenced by the sample population, as they were predominately leisure passengers. Cameras are extremely unlikely to pose a threat to aircraft, as they are comparatively low powered devices with low radio emissions. Laptop computers were carried by only 9% of participants; but again, had more business passengers been included in the sample this number would almost certainly have been larger. One third of the laptop computers had wireless connection capability that could be operated during cruise without being obvious to cabin crew, although none of the wireless connections were reported to have been used by the participants during flight. Overall, participants carried 164 devices that were capable of intentional transmissions and 191 unintentional transmitters.

The vast majority of PEDs were not operated on board, with music players, such as CD, tape, minidisc and MP3 players, being the only category in which a notable number were operated. Music players were operated in 42.5% of cases where they were carried. Interestingly, music players are the one device over which there are considerable discrepancies across airline policies and this may lead to confusion and irritation for passengers who have become used to operating their music player on one airline but are refused permission to do so on another.

Section Two

Thirteen percent of participants admitted forgetting to switch off their mobile phones on a previous flight and 2% had used them on a previous flight. Many of the participants who had forgotten to switch off their phones on a previous flight claimed that they regularly neglected to switch them off. It is impossible to determine from these statistics how regularly passengers use their mobile phones during flight, but it is clear that some passengers do so. Using a mobile phone may increase the possibility of interference as the signal from the phone is more constant and is sustained over time.

When questioned about the circumstances under which participants felt it would be acceptable to use a mobile phone without permission during flight, the majority (63%) of participants felt that it would be acceptable if the aircraft was hijacked and 54% felt it also would be acceptable to use a mobile phone under emergency conditions. The second statistic is alarming because emergency conditions are exactly the time when interference could potentially increase already high pilot workload and possibly contribute to an accident by exacerbating pilot confusion. Airlines should consider warning passengers not to use mobile phones when preparing passengers for an emergency landing. 5.5% of participants thought that it was acceptable to use a mobile phone if the aircraft was delayed in the air, and although this sounds more trivial than using a mobile phone under emergency conditions, the potential danger involved may be much lower. 1.5% of passengers felt that using a mobile phone in the air was always acceptable.

Section Three

Section three of the questionnaire examined participants' knowledge of the regulations concerning PEDs. 12.1% of participants were not aware of the regulations for mobile phones, although 25% of these did not carry a mobile phone on board and reasoned that they did not need to know the regulations. More worrying is the 5.5% of participants who thought that mobile phones only had to be switched off for takeoff and landing. Fifty percent of these were carrying mobile phones and should have been paying attention to the announcements regarding their use. The remaining 82.5% of participants knew that mobile phones should remain switched off at all times.

With regard to unintentional transmitters such as laptop computers and CD players, the number of participants who were not aware of the regulations rose to 24.7%, while 31.4% mistakenly believed that unintentional transmitters must remain switched off at all times. Only 43.9% of participants knew the correct regulation in this case, indicating a greater confusion and lack of awareness regarding unintentional transmitters as compared with intentional transmitters.

The final question in this section asked participants what they thought was the maximum legal punishment in the UK for failing to switch off a personal electronic device when instructed to do so. The correct answer was the £2000 fine and/or two-year prison sentence, but only 13.6% of participants knew, or guessed this correctly, and 16.6% thought that the maximum punishment was a £100 fine. The most common answer was the third option with 30.2% of participants opting for this. Clearly, the public are not aware of what the actual maximum punishment is and when asked to guess, 86.4% of participants underestimated it. This may give an indication of the lack of seriousness with which they regard the danger of mobile phone use on aircraft. Certainly, the 16.6% who thought the maximum punishment was a £100 fine could not regard the danger as very serious if they thought the punishment was so low.

Section Four

Participants' knowledge of the reasons why PEDs are restricted on aircraft was examined in this section of the questionnaire. Participants who demonstrated knowledge that PEDs could cause interference were asked to choose which systems could be affected. The section was split into two questions, one enquiring about intentional transmitters, represented by mobile phones and the second dealing with unintentional transmitters, represented by laptop computers and CD players. In the case of intentional transmitters, interference to navigation systems was the most commonly cited, followed by communication systems. This is interesting because navigation systems are the most commonly affected system, but it was expected that participants would think communication systems were more vulnerable to interference from mobile phones. Some airlines cite interference to navigation systems as the reason why mobile phones are prohibited during safety announcements and this result showed that passengers have understood this. However, it does not prove that they believe it to be true. Only 11.5% of participants thought that using a mobile phone in an aircraft could interfere with the ground-based mobile network. As this is the only type of interference that has been proven, this was a low percentage. 6% of participants gave all six options available, namely interference to navigation systems, communication systems, cockpit displays, autopilot, engine systems, and the mobile network. 22.9% did not know of any specific reasons why mobile phones were prohibited, although the vast majority of these did know that the prohibition was attributable to some kind of interference.

The percentage of participants who did not know why unintentional transmitters are restricted was much higher at 45% and a relatively low number of these knew that it had something to do with interference. Again, interference to navigation systems was the most frequently cited with 45% of participants, and communications followed with 38.2%. 14.5% of participants thought that the possibility of injury in the event of turbulence or a heavy landing contributed to the restriction on unintentional transmitters. This was the only reason that did not have a higher representation in the question on intentional transmitters. The results seemed to indicate that passengers are more confused about the reasons behind the restrictions on unintentional transmitters and several participants commented, "they could not see how a laptop computer or CD player could interfere with any aircraft equipment as it does not transmit a signal." Some participants also concluded that these devices could not cause interference as if they were suspected of doing so they would not be allowed at any stage during the flight.

Section Five

The final section of the questionnaire investigated participants' perceptions of the danger associated with operating PEDs on board aircraft and participants' desire to use PEDs. Participants were asked to rate their agreement with the statements "mobile phones are dangerous if operated on board aircraft" and "laptops are dangerous if operated on board aircraft." Mobile phones and laptop computers were chosen for these statements because they represented intentional and unintentional transmitters respectively and because they have been reported to be equally likely to interfere with aircraft systems (Ross, 2001).

In general, there was more and stronger agreement with the statement regarding mobile phones as compared with the statement regarding laptop computers. Consequently, more participants felt that mobile phones were dangerous, with only 5.5% disagreeing with the statement regarding the danger of mobile phones. Compared to mobile phones, a larger number of participants did not know if laptop computers were dangerous. These results confirmed that the confusion regarding the potential danger associated with PED is much greater with unintentional transmitters than it is with intentional transmitters.

The number of passengers who strongly agreed with the statement "I would like to be able to use my mobile phone, if it were safe" was not as high as might be expected, with only 5% of participants choosing this option, and a larger number (5.5%) strongly disagreed with the statement. The total number who expressed agreement was 39.2% compared with 35.7% who disagreed and 25.1% were neutral. The reasons for these opinions were not investigated but several passengers commented that they did not feel it would be necessary to use their phones during a flight, while others stated that their decisions were influenced by the possibility of other passengers using their phones throughout the flight. This may indicate a desire on passengers' behalf for the continuation of the prohibition on comfort or annoyance grounds. There was much stronger agreement with the statement "I would like to be able to use laptops and music players at all stages of flight, if it were safe," where only 21.1% of passengers expressed disagreement.

Independent t-tests were carried out on the results of this section to determine whether there were significant differences between any of the demographic groups. It was found that females considered mobile phone and laptop computer use to be more dangerous than males did ($t = 2.839$, $p < 0.05$; $t = 3.174$, $p < 0.05$). Males also were more likely to want to use laptop computers and CD players at all stages of flight ($t = -2.575$, $p < 0.05$), but there was not a significant difference in the desire to use mobile phones on board ($t = -1.385$, $p > 0.05$). Neither were there any significant differences between passenger types for any of the answers to questions.

Study Limitations

Due to the nature of the questionnaire administration, it was not possible to obtain a representative sample of airline passengers; therefore, opportunity sampling had to be used. Furthermore, there were limitations in the time at which it was possible to interview passengers and the ultimate result was a sample dominated by short haul leisure passengers. A different sample that included more business and long haul passengers may have yielded some different results.

As the participants were interviewed at airports, it was necessary to keep the questionnaire as short as possible to maximise the number of participants who would have time to complete it. This resulted in some lines of questioning not being pursued and some data, which may have been useful for the study, could not be obtained. A self-completion questionnaire would have resolved this problem, but the response rate would have been much lower.

Time limitations and a low response rate from airlines meant that the investigation into airline policies suffered from too little data. The lack of data regarding in flight announcements on PEDs meant that it was not possible to analyse the information most commonly available to passengers.

Conclusions

An extremely large percentage of participants (90%) carried personal electronic devices on board the aircraft and a relatively large percentage carried more than one, but only a low percentage of the devices were operated on board the aircraft. Mobile phones were the most common device and 2.5% of participants failed to switch off their mobile phones. Furthermore, 2% of participants admitted to using their mobile phones during a previous flight. Although these figures are low, considering the potential danger, it is in the interest of safety to increase passenger awareness and compliance with the regulations.

The majority of participants were aware of the regulations governing the use of PEDs on the airline with which they flew; however, there were some passengers who thought that intentional transmitters could be used during the cruise phase of flight; many of these were carrying mobile phones and so should have been aware of the regulations. The survey revealed that passengers are less well aware of and more confused about the regulations regarding unintentional transmitters as compared to intentional transmitters. These results pointed to a need for education of passengers.

A high proportion of participants were aware of the possibility of interference to aircraft equipment from PED, but very few knew of the possibility of interference to mobile networks if intentional transmitters are used in the air.Again, there was less awareness regarding unintentional transmitters as compared to intentional transmitters.

Participants overall agreed that mobile phones are dangerous if operated on board aircraft, but 5.5% disagreed with the statement once again indicating the need for education of the public. There was less agreement with the similar statement on laptop computer use that confirmed the confusion on unintentional transmitters noted in the earlier questions. The percentage of participants who would like to use mobile phones on board aircraft was only marginally larger than the number who did not want to use them. Thus, there is a case for prohibition of intentional transmitters not just on safety grounds, but also for the comfort of fellow passengers.

The reasons for passenger confusion and non-compliance with regulations may be at least partially due to the differing regulations across airlines where some devices may be permitted during cruise on one airline and prohibited altogether on another.A coherent policy would ensure that the devices that are regarded as dangerous are regarded consistently across airlines. A small percentage of passengers did not agree that mobile phones are dangerous on board aircraft and a larger percentage did not agree that laptop computers are dangerous. Thus, it would appear that ignorance of the dangers is also part of the problem.

Recommendations

Uniform Airline Policy

A uniform policy across airlines would greatly reduce passenger confusion and would remove the possibility of airlines that allow more PEDs to be operated having a market advantage over those who do not and this in turn might lead to greater reporting of incidents of suspected interference and greater understanding of the problems associated with PEDs. A consistent policy with regard to mobile phones, which required passengers to switch off mobile phones before leaving the terminal, would allow taking more active measures in the airport departure lounges to ensure mobiles are switched off.

PED Detection System

A device for detecting PEDs that are operating on board aircraft above the susceptibility levels for the aircraft equipment, such as the device developed under the FAA (Cross, 2000), should be installed on all aircraft. A PED detection device would allow cabin crew to be aware of any passengers who have forgotten to switch off their devices or who are using them surreptitiously.The device also could record the location and frequencies of all PEDs operating on all flights and in the event of suspected interference, this data could be correlated with the events on the flight deck providing valuable information about the circumstances under which interference can occur.

Further Testing

Testing of PEDs in aircraft and with aircraft equipment should not be neglected. Although it has not proven possible so far to reproduce interference, it may be possible in the future. Testing of new devices is particularly important, as they may be much more likely to interfere with aircraft systems than devices currently available. However, there are limits to the testing it is possible to achieve, particularly considering the vast range of electronic devices available. As the regulatory authorities are in a better position to perform testing it should be their responsibility.This would ensure more comprehensive and uniform testing than individual airlines are able to perform.

Educate Passengers

The results of this study have shown that the majority of passengers carry electronic devices and a small percentage operates them on board. It can only be expected that both of these figures will rise in the future as electronic devices become more and more common. It also was shown that many passengers do not fully understand the potential dangers of PEDs and are not fully aware of the regulations. Therefore, it is recommended that a programme of awareness and education be embarked upon where more information is passed on to passengers regarding PED. This could be done either by making information available to passengers when they receive their tickets or at check in.

Procedures for Interference

Airlines should develop procedures for collecting data when interference is suspected, such as the procedure outlined by the CAA (2002). Data gathered in a systematic way would be more reliable and if the same procedures were used across airlines, it would be possible to compare incidents of interference easily. Ultimately, this would provide sufficient information to develop policies that are synchronised with the potential for interference posed by each device.

References

Aberbach, J.G. (1999). Cell-phone use aloft may not be the danger that airlines claim. *Wall Street Journal, 234* (67), 1.

BBC news. (1999, July 21). *In-flight mobile user jailed.* Retrieved June 2, 2003 from http://news.bbc.co.uk

BBC news. (2002, September 5). *Air passenger jailed for using mobile.* Retrieved May 21, 2003 from http://news.bbc.co.uk

Bicknell, C. (2000, July 25). Are airborne e-devices a danger? *Wired News.* Retrieved May 8, 2003 from http://www.wired.com

Civil Aviation Authority (CAA). (2000, May 2). *Interference levels in aircraft at radio frequencies used by portable telephones.* West Sussex, London, England: CAA.

Civil Aviation Authority (CAA). (2001, October). *Guidance concerning the use of portable electronic devices on board aircraft.* Flight Operations Department Communication (FODCOM) of the Joint Aviation Authorities (JAA), Retrieved June 2, 2003 from http://www.caa.co.uk

Civil Aviation Authority (CAA). (2003, April 3). *Effects of interference from cellular telephones on aircraft avionic equipment.* (ISBN 0 86039 923 0).West Sussex, London, England: CAA. Retrieved June 2, 2003 from http://www.caa.co.uk

Civil Aviation Authority (CAA), Safety Regulation Group. (2002, January 22). *Is electromagnetic interference from aircraft passenger electronics a problem?* Airworthiness Information Leaflet (Ref: AIL/0113, Issue 2). Retrieved June 2, 2003 from http://www.caa.co.uk

Cross, M.W. (2000). Portable electronic device (PED) detection system. Retrieved June 19, 003 from http://www.house.gov/transportation/aviation

Donham, B. (2000). Electromagnetic interference from passenger-carried portable electronic devices. *Boeing Aeromagazine, 10*, 13–18.

Ely, J.J., Fuller, G.L., & Shaver, T.W. (2002). Ultrawideband electromagnetic interference to aircraft radios. In *Proceedings of the 21st Digital Avionics Systems Conference* (pp. 13E4-1 – 13E4-12). Irvine, CA: Institute of Electrical and Electronics Engineers.

Evans, D. (1999). The unknown risks of PEDs. *Avionics Magazine.* Retrieved May 17, 2003 from http://www.aviationtoday.com.

Helfrick, A. (1994). Interference to digital avionics due to electromagnetic radiation from personal electronic devices carried aboard. In *Proceedings of the 13th Digital Avionics Systems Conference* (pp. 227–232). Phoenix, AZ: Institute of Electrical and Electronics Engineers.

Interference from portable electronic devices demonstrates potential for catastrophe. (2003, February 10). *Air Safety Week*, 1–5.

Jenson, D. (2000). The persistent PED problem. *Avionics magazine 24* (8). Retrieved June 25, 2003 from http://aviationtoday.com

Ladkin, P.B. (1997). *Electromagnetic interference with aircraft systems: Why worry?* Retrieved May 8, 2003 from http://rvs.uni-bielefeld.de

Li, L., Xie, J., Ramahi, O.M., Pecht, M., & Donham, B. (2002). Airborne operation of portable electronic devices. *IEEE Antennas and Propagation Magazine, 44* (4), 30–39.

Muller, K. (2002). *Use of electronic devices aboard aircraft*. Retrieved May 18, 2003 from http://www.ila.de/englisch

Perry,T.S., & Geppert, L. (1996). Do portable electronics endanger flight? *IEEE Spectrum, 33* (9), 26–33.

Portable electronic devices still a problem. (1996, *November 4*). *Air Safety Week*, 1–2.

Ritchie, C. (1996). Potential liability from electromagnetic interference with aircraft systems caused by passengers' on board use of portable electronic devices. *Journal of Air Law and Commerce, 61* (3), 683–720.

Ross, E. (2001). *Personal electronic devices and their interference with aircraft systems*. Hampton, Virginia: NASA.

Rossier, R. N. (2000). Problem PEDs: Avoiding the stray CAT II strut. Play it safe with portable electronic devices: Turn them off. *Business and Commercial Aviation, 87* (6), 67-73.

RTCA Inc. (1996). *Portable electronics carried on board aircraft*. Washington DC: RTCA.

Shawlee, W. (1999). Why do the wings dip when I hit play? *Avionics magazine 23* (12). Retrieved July 10, 2003 from http://www.aviationtoday.com

Shawlee, W. (2000). Potential perils of PEDs. *Avionics magazine 24* (1). Retrieved June 25, 2003 from http://www.aviationtoday.com

Strauss, B. (2001). Portable electronics carried on board aircraft: Towards an assessment of the risk. Carnegie Mellon University: Qualifier Paper.

Strauss, B. (2002). Avionics interference from portable electronic devices: Review of the aviation safety reporting system database. In *Proceedings of the 21st Digital Avionics Systems Conference* (pp. 13E3-1 – 13E3-8). Irvine, CA: Institute of Electrical and Electronics Engineers.

Subcommittee on Aviation. (2000). *Portable electronic devices: Do they really pose a safety hazard on aircraft?* Retrieved May 8, 2003 from http://commdocs.house.gov

International Journal of Applied Aviation Studies, 4, Number 2

Flying IFR with GPS: How Much Practice Is Needed?

Stephen M. Casner

National Aeronautics and Space Administration
Ames Research Center
Moffett Field, CA 94035-1000

Abstract

Sixteen instrument-rated pilots with no prior experience with IFR GPS completed a program of ground study and five practice flights in an airplane. Eight pilots completed the ground study using a self-study program, while eight pilots received dual ground instruction. The ground study and flight practice covered knowledge and skills required by the instrument rating practical test standard, which are affected by the use of IFR GPS. A detailed record was kept of errors made by pilots during each practice flight for six selected skills. The data were analyzed to determine: (1) whether or not the ground study and five practice flights were enough to allow pilots to master the skills; (2) how effective was self-study compared to dual instruction; and (3) which skills presented pilots with the most difficulty and accounted for the most errors. The results showed that pilots still had not reached proficiency after five practice flights, regardless of ground study method used. Furthermore, pilots were highly similar in the difficulties they encountered while acquiring these new skills. These results suggested that the learning challenges for proficient IFR GPS use are significant.

Introduction

Since the late 1990s, the installation of IFR-approved GPS units in general aviation aircraft has steadily increased. Initial studies of GPS usage (Heron, Krolak, & Coyle, 1997; Henry, Young, & Dismukes, 1999; Adams, C.,

---

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

Hwoschinsky, & Adams, R., 2001) have prompted concern about what sorts of additional knowledge or experience might be required to use GPS safely as a primary means of navigation under instrument flight rules.

The FAA slowly and conservatively has taken advisory and regulatory steps toward insuring the safe use of IFR GPS. The Aeronautical Information Manual (FAA, 2004) has been expanded to include a section about IFR GPS. The Instrument Rating Practical Test Standard also has been modified to require every pilot applicant to demonstrate proficiency with IFR GPS when an IFR GPS-equipped aircraft is used for a practical test.

The idea that additional training or experience might be required for IFR GPS is not only is a question of safety, but also a question of popular acceptance. Users of IFR GPS might object to new regulations that require additional and expensive pilot training if the need for such training was not carefully documented and made explicitly clear.

We studied a group of sixteen instrument-rated pilots with no prior experience with IFR GPS as they worked toward proficiency with flying under IFR with GPS. Pilots' learning efforts consisted of two parts: (1) ground study; and (2) five practice flights in which pilots practiced the skills they learned on the ground. Eight pilots completed the ground study through dual instruction, while eight pilots studied the same material on their own.

During the practice flights, a detailed record was made of all errors committed by pilots when practicing six selected skills. The skills are:

1. Program IFR flight plan and load GPS approach
2. Program and fly a VNAV descent
3. Demonstrate a straight-in GPS approach
4. Demonstrate a vectored GPS approach
5. Demonstrate a missed approach and hold
6. Demonstrate a GPS approach w/ procedure turn

The ground study and the flight practice covered other knowledge and skills required for safe and proficient use of IFR GPS, but they were not measured as part of the experiment.

Pilots' error data were analyzed to address three questions:

1. Was ground study and five practice flights enough for pilots to master the six skills?
2. How effective was self-study compared to dual instruction?
3. Which skills presented pilots with the most difficulty and ac counted for the most pilot errors?

Method

Participants. Sixteen instrument-rated pilots were recruited from local professional flight training schools. Pilots ranged from 120 to 3,700 hours of flight experience, with a median of 522 hours. Pilots were told they would not be paid for their participation but would receive instrument flight experience using IFR GPS. All pilots met the recent flight experience requirements set forth in 14 CFR 61.57 (c).

Procedure. Eight pilots were randomly assigned to the Self Study group and were told that they would be required to learn the new skills on their own. These pilots were assigned readings in a textbook (Casner & Dupuie, 2002) prior to each session. Pilots were told to master the material as best as they could, and that during the next session, they would have the opportunity to practice and demonstrate their newly learned skills in flight. It was emphasized that pilots' should attempt to master the skills such that they could demonstrate them without the need for intervention by the experimenter, although intervention would be available if needed. Rather than attempt to control the duration of self-study for experimental design purposes, our aim was to make this learning scenario as realistic as possible: as it might occur in everyday practice. Pilots were not asked to report the amount of time they had spent studying.

Eight pilots were assigned randomly to the Dual Instruction group and were told to do nothing to prepare for the flight sessions. These pilots were told that the experimenter would cover all of the concepts and skills needed for each flight during a dual ground instruction session immediately prior to the flight. The material presented during the dual instruction sessions was the same as that presented in the textbook read by pilots in the Self Study condition. Pilots were told that they should attempt to master the skills such that they could demonstrate them without the need for intervention by the experimenter, although intervention would be available if needed. Dual instruction sessions continued until the pilots felt they were ready to demonstrate successfully the skills they had learned.

Both groups of pilots had access to a desktop IFR GPS unit that could be used to learn and practice GPS skills prior to each practice flight. The desktop IFR GPS unit was the same make and model installed in the airplane that was used for the practice flights.

For both groups, prior to each practice flight, the experimenter briefly reviewed the skills that would be needed during the flight, provided the pilot with charts covering the routes and approaches to be flown, and answered any questions the pilot had about the material.

The six skills were introduced before the practice flights as shown in Table 1. Table 1 also lists the number of times that each skill was practiced during each flight. It is important to note that not every skill was practiced on every flight. No new skills were introduced during the fifth practice flight.

Table 1

Six IFR GPS skills practiced during the five practice flights.

	Flt1	Flt2	Flt3	Flt4	Flt5
Skill					
Program IFR flightplan and load GPS approach	1	3	3	3	4
Program and fly a VNAV descent	1				1
Demonstrate a straight-in GPS approach	1			1	
Demonstrate a vectored GPS approach		3	3	1	3
Demonstrate a missed approach and hold			1		1
Demonstrate a GPS approach w/ procedure turn				1	1

Practice flights were conducted using the same protocol and evaluative technique used in similar studies of pilot proficiency (Talleur et al, 2003). During the practice flights, participants acted as sole manipulator of the controls under simulated instrument conditions (a standard view-limiting device was used). The experimenter rode in the right seat and acted as flight instructor and observer. A script for each flight was prepared in advance and used by the experimenter to ensure that each flight proceeded in accordance to a set plan, and that each pilot was asked to practice and perform the same skills in the same order. The scripts used for each flight are given in Appendix A.

A palmtop computer was used to record errors made by the pilot on any skill, or assistance requested by the pilot for any skill. A scorecard was kept for each pilot and flight. For each skill, if the pilot was able to demonstrate the skill without error or help from the experimenter, the pilot received a score of one. If an error was made or help was provided, regardless of how subtle (e.g., words, gestures, sounds), a score of zero was recorded for that skill. Help was provided by the experimenter only when the flight could no longer continue to the next step in the planned flight.

Results and Discussion

Figures 1(a) and 1(b) show the error rates for each of the skills during each of the five flights, across all pilots in the Self Study and the Dual Instruction groups. Error rate means the proportion of failed attempts to demonstrate each skill among the total number of attempts to demonstrate each skill, for all pilots in each condition. For example, in Figure 1(a), the error rate for the Program Route skill during the first flight is roughly 0.38. This means that pilots collectively succeeded in demonstrating the Program Route skill 62 percent of the time, and failed to demonstrate the skill 38 percent of the time, during the first flight. Recall that not every skill was demonstrated during every flight; hence, some skills appear fewer than five times in the graphs.

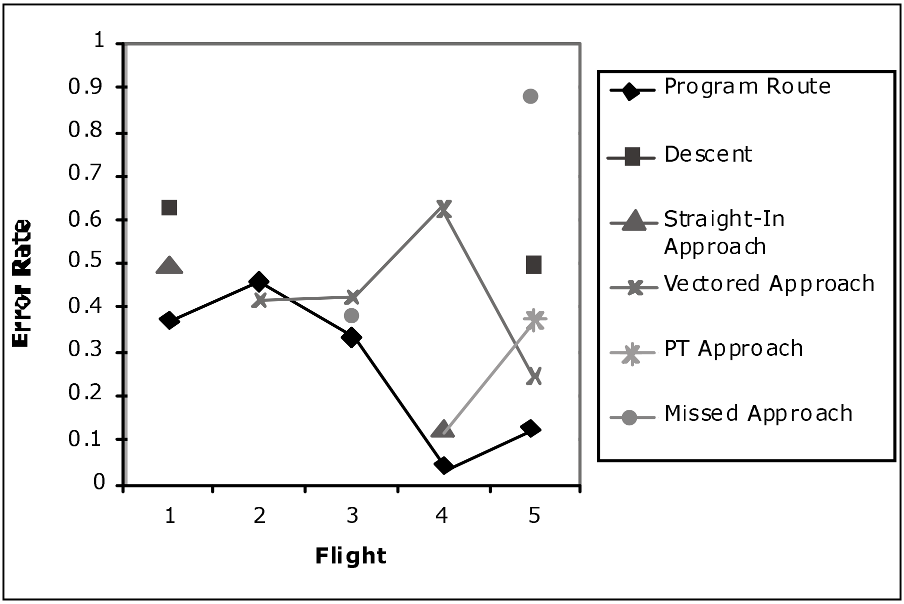


Figure 1(a). Error rates for the six skills (dual instruction)

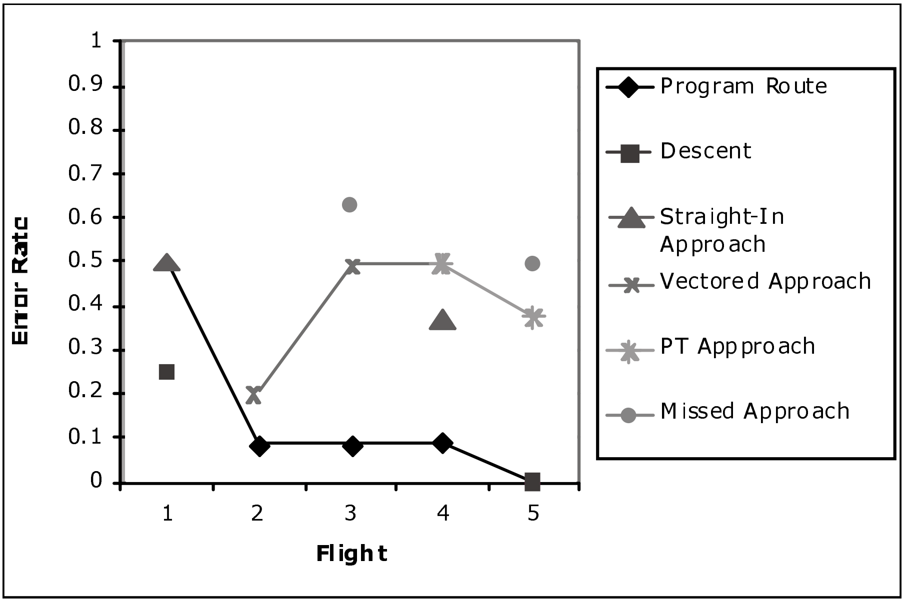


Figure 1(b). Error rates for the six skills (self-study)

1. Were five practice flights enough? The first question to address is whether ground study and five practice flights were enough to allow pilots to reach proficiency with the six skills. Looking at the data points for the later practice flights in both graphs in Figure 1 we can see that errors persist for most skills.

Although pilots seemed to have mastered the knobs-and-dials procedures needed to program a route, skills associated with flying instrument approaches seemed to require more practice.

The data in Figure 1(a-b) make it abundantly clear that our initial expectations of reaching proficiency after two, three, four, or five practice trials were off the mark. Our expectations were based on the understanding that all pilots had already demonstrated proficiency for these same skills using different navigation systems (i.e., all pilots held an instrument rating and were instrument current). However, the data in Figure 1 does show that the likely number of needed practice trials for each skill is more than what was provided here.

To verify that the adage "practice makes perfect" is at work here, the number of practice trials for each skill was correlated with the error rates for the last trial in which each skill was practiced. For example, looking at Figure 1 (a), the Vectored Approach skill was practiced four times and the error rate was 0.25 on the final trial. Comparing these two numbers for each skill, the correlation coefficients for the Dual Instruction and Self-Study groups were –0.53 and –0.42, respectively.That is, for both groups, high numbers of practice trials were associated with lower error rates. This result suggests that learning was occurring, only at a slower rate than what we had originally hypothesized.

2. Did having more total flight experience make a difference? It is interesting to look at the relationship between the total flight experience of the participants in the study, and their error rates for the six skills. The correlation coefficient for total flight hours and error rates for the six skills for all sixteen pilots was $r = 0.01$. This lack of correlation suggests that proficiency with IFR GPS is a separate set of skills to be acquired. Having extensive flight experience in airplanes not equipped with IFR GPS does not appear to help. Flying proficiently with IFR GPS seems to be the result of training and experience flying with IFR GPS.

3. Was there a difference between dual instruction and self-study? A third question to consider is whether there are any observed differences between the two methods of ground study: dual instruction and self-study. Dual instruction seems to offer the advantage of two-way interaction between student and instructor. Self-study offers the advantage of a persistent record of the instructional material that can be later reviewed. It is interesting to note that none of the pilots in the Dual Instruction group made use of notes.

A simple comparison of mean error rates between the two groups for all six skills (i.e., the data shown in Figure 1) yielded no significant difference. Next, six t-tests were performed comparing the error rates for each of the six skills individually. Only the Build and Fly Descent task yielded a significant difference ($t = 2.65$, $p < 0.05$). Since this task was only practiced twice during the course of the five practice flights, no strong conclusions are warranted.

Overall, it seems that the two ground learning methods yielded similar results.

4. A more detailed breakdown of errors made on each of the six skills.

Figures 2 – 7, show the specific criteria that were used to determine success or failure for each of the six tested skills. The data in Figures 2 – 7 breakdown overall performance for each skill into performance on component sub-skills. Figures 2 - 7, list the sub-skills associated with each skill, and show the proportion of cases for which each sub-skill was a contributing factor in pilots' failure to perform each of the six skills.

Since no significant differences were found between the two learning methods, Figures 2 - 7 combine the results for the two ground learning methods.

Program Route and Install GPS Approach

The Route Programming skill consisted of two sub-skills shown in Figure 2.

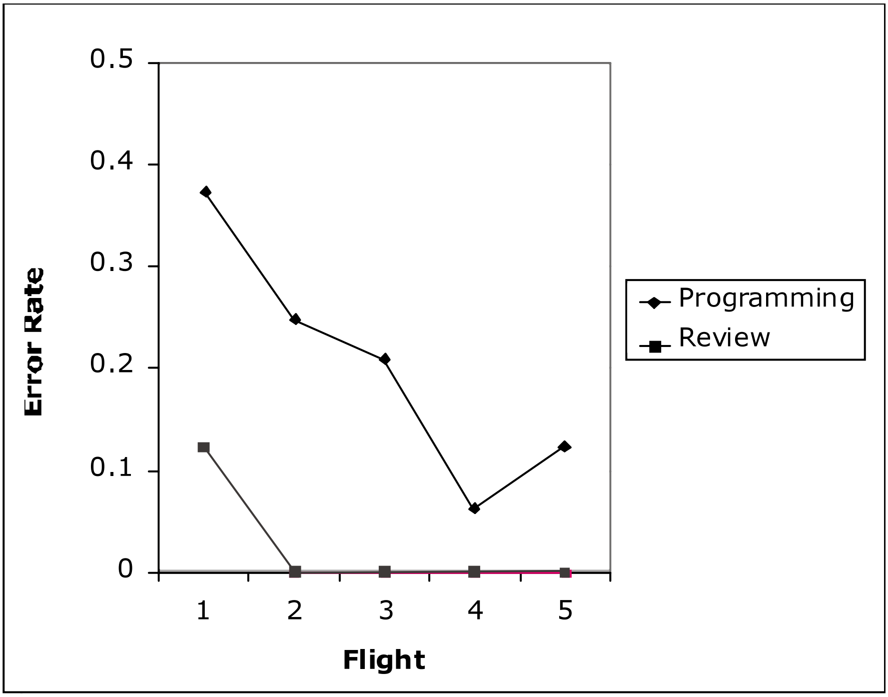


Figure 2. Sub-skills that comprise the Program Route skill.

The Programming sub-skill required pilots to recall and perform the knobs-and-dials procedures needed to install the route. This sub-skill is essentially a memory task aided by any cues provided by the GPS unit interface. For example, a button marked FPLN might allow pilots to successfully reach the flight-planning page when the procedure has not been memorized. The Programming sub-skill was the principle cause for the occasional unsatisfactory performance for the Route Programming task.

The Review sub-skill requires pilots to remember to review the accuracy of a flight route once it is installed. Pilots seemed to have well grasped the importance of checking their work.

Build and Fly a Descent

The Descent skill consisted of two sub-skills shown in Figure 3. Recall that the Build and Fly a Descent skill was only demonstrated during the first and fifth flights.

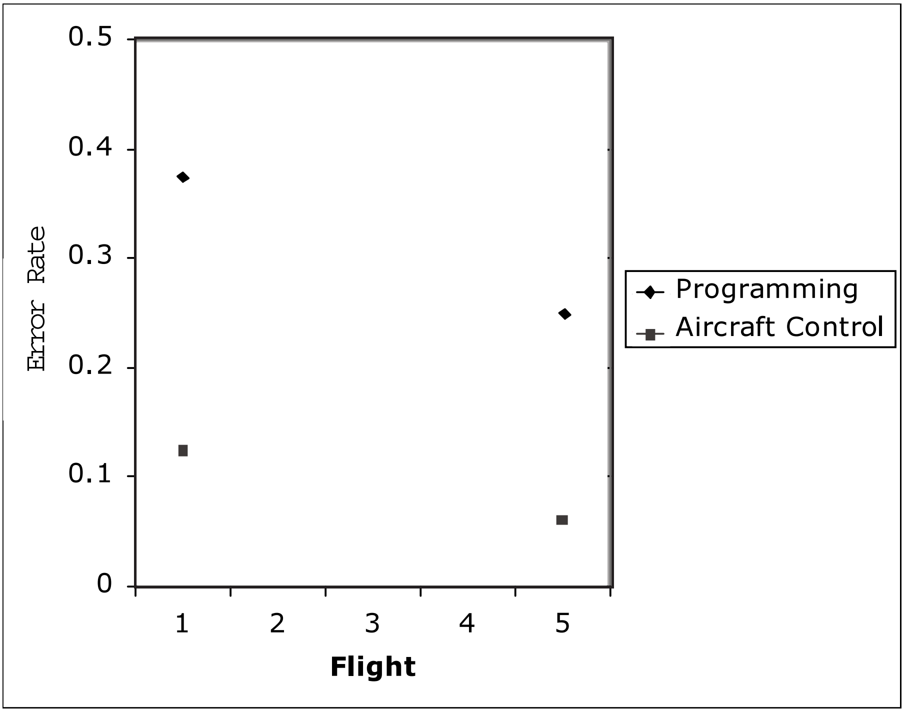


Figure 3. Sub-skills that comprise the Descent skill.

The Programming sub-skill requires pilots to recall the knobs-and-dials procedure required to build a VNAV descent path. Again, the Programming sub-skill was the primary cause of errors. Since the Descent skill was practiced only twice in flight, the high error rate observed for this skill cannot be regarded with any certainty. In fact, the two data points in Figure 3 exactly match the improvement trajectory observed for the programming sub-skill for the Program Route skill (see Figure 2).

The Aircraft Control sub-skill required pilots to meet the crossing restriction they had programmed. Errors on this sub-skill were related to inattention: failure to start the descent at the top-of-descent point computed by the GPS unit, or failure to maintain the target rate of descent.

Straight-In GPS Approach

The most basic type of GPS approach was scored with three sub-skills shown in Figure 4. This skill was demonstrated only during the first and fourth flights.

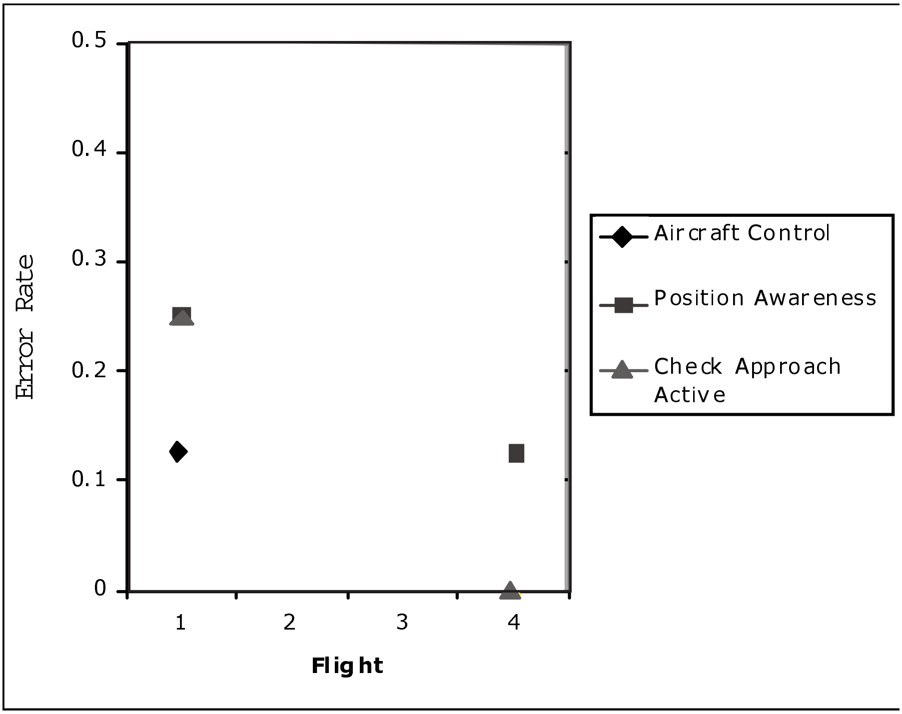


Figure 4. Sub-skills that comprise the Straight-In Approach skill.

The Check Approach Active sub-skill is particularly important. Every IFR-approved GPS unit features an annunciation that informs the pilot that all necessary conditions are met to continue an approach beyond the final approach fix and descend to the minimum descent altitude (MDA). Continuing the approach without an approach active indication could result in disastrous consequences since the integrity of the course guidance is not guaranteed. Pilots initially struggled with this important skill but seem to have resolved the problem by the end of the practice flights.

The Aircraft Control sub-skill was a simple measure of how frequently pilots deviated more than 100 feet from a required altitude, or allowed a full-scale deflection of the CDI needle. It is widely known by instructors and pilots alike that aircraft control performance varies when workload is increased and distractions are introduced.

The Position Awareness sub-skill resulted in an error when pilots failed to announce their position at an important approach waypoint, or took a required action at an inappropriate place. Several pilots began a descent to the MDA prior to reaching the final approach fix.

Vectored GPS Approach

Five sub-skills shown in Figure 5 comprised this more sophisticated type of GPS approach. This skill was not demonstrated by pilots during the first flight.

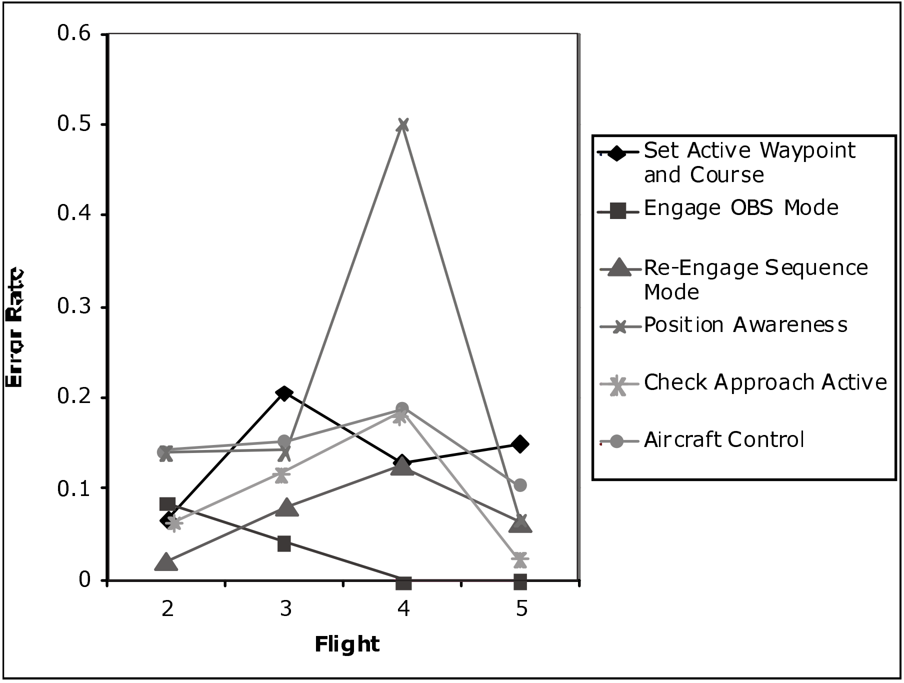


Figure 5. Sub-skills that comprise the Vectored Approach skill.

The Set Active Waypoint and Course sub-skill required pilots to change the active waypoint in the GPS computer to a different waypoint that was farther ahead in the approach procedure. This is required when ATC vectors the pilot inside of the initial approach fix in order to shorten an approach. Error rates for this sub-skill never significantly improved over the course of five flights. The consequences of making an error on this sub-skill are severe. Entering the wrong waypoint or course means that the pilot is following a course other than the published approach course.

The Engage OBS Mode sub-skill requires pilots to engage the GPS unit's non-sequencing mode, which allows the pilot to use the OBS knob to dial arbitrary courses to any waypoint. For this type of approach, the pilot dials in the final approach course. Pilots seem to have mastered this sub-skill quickly.

The Re-Engage Sequence Mode sub-skill is somewhat challenging in that it requires pilots to remember to take a future action, a cockpit memory task known to be difficult (Nowinski, Holbrook, & Dismukes, 2003). Pilots were still forgetting roughly seven percent of the time even after five practice flights.

Errors on the Position Awareness, Check Approach Active, and Aircraft Control sub-skills continued to be somewhat problematic for vectored approaches.

Missed Approach and Hold

The Missed Approach and Hold skill was scored using the five sub-skills shown in Figure 6. This skill was only demonstrated during the third and fifth flights.

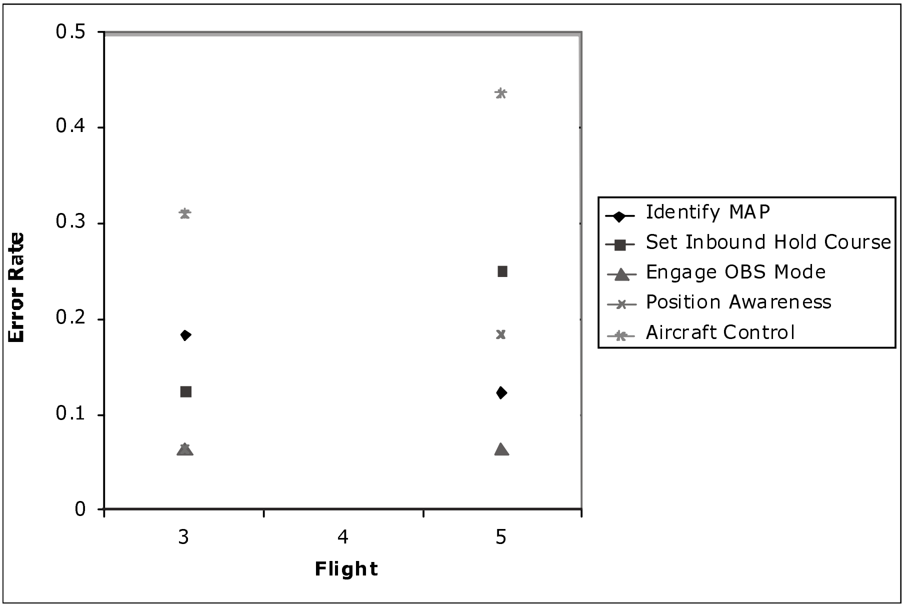


Figure 6. Sub-skills that comprise the Missed Approach and Hold skill.

Identify Missed Approach Point is another sub-skill with potentially serious consequences. Pilots failed to recognize the missed approach point roughly 12% of the time on their second missed approach procedure. One pilot overran the missed approach point by 2.4 NM. This sub-skill is particularly worrisome because the sixteen pilots have already demonstrated their ability to recognize missed approach points using other navigation systems.

The two practice trials proved insufficient for most other sub-skills. Pilots consistently had trouble dialing the correct inbound hold course and in controlling the aircraft.

GPS Approach with Procedure Turn

The GPS Approach with Procedure Turn sub-skill was scored using the five sub-skills shown in Figure 7. This skill was demonstrated during the fourth and fifth flights.

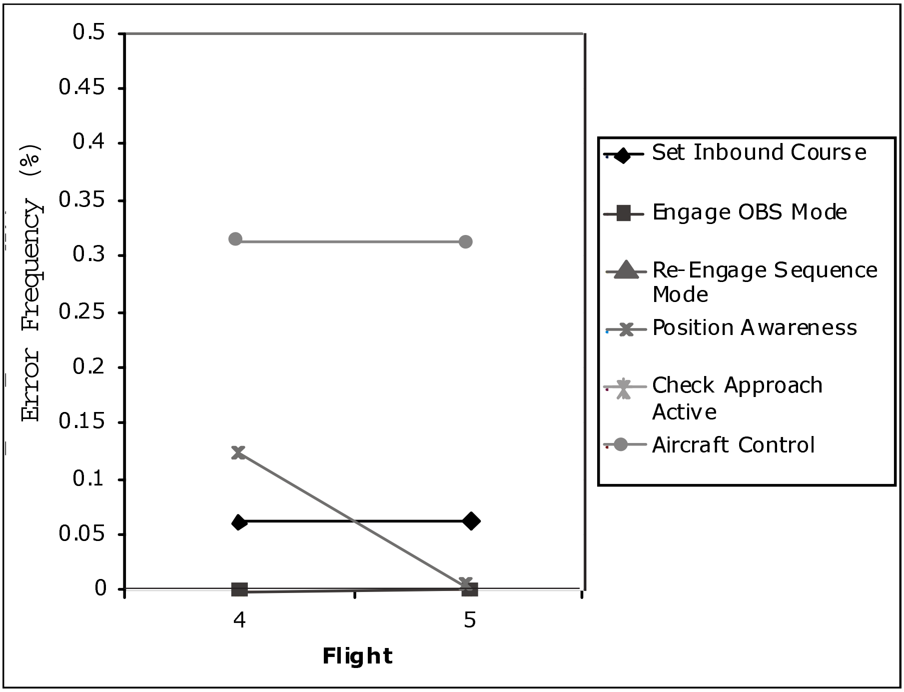


Figure 7. Sub-skills that comprise the Procedure Turn Approach skill.

After two practice trials, pilots were still sometimes failing to dial the inbound hold course, and were still experiencing problems with aircraft control.

5. Did similar skills result in similar performance? Looking at the sub-skills listed in Figures 2 - 7, we notice that some skills require the pilot to perform similar sub-skills. For example, Position Awareness and Aircraft Control are both required sub-skills for all four approach-related skills. Similarly, Engage OBS Mode, Re-Engage Sequence Mode, and Check Approach Active are common to the Vectored and Procedure Turn Approach skills. It is interesting to note whether performance on sub-skills were similar across different skills that used them. It may be that sub-skills that are learned and practiced on one skill might help expedite learning and improve performance on later skills that use them.

Development of sub-skills related to engaging or monitoring modes appeared to follow a similar trajectory across skills. The Check Approach Active sub-skill seemed to have been soundly learned by Flight 5 for all three GPS approach skills (Straight-in, Vectored, and Procedure Turn). It is not clear whether or not practicing this sub-skill in three different contexts helped to develop a more general skill. The Re-Engage Sequence Mode sub-skill is another memory-related skill that presented most pilots with initial difficulty. This sub-skill also was well mastered across the Vectored and Procedure Turn Approach skills.

There appeared to be little similarity among performance on the Aircraft Control sub-skill across the different skills. Pilots who flew within tolerances for some skills were sometimes quite out of control during performance of other skills. This cast doubt on theories that claim that aircraft control is an indicator of a more general division of attention skill that, once mastered, applies to pilot performance in the large.

Conclusion

After ground study and five practice flights in an airplane, the data showed that the pilots we studied had not yet reached proficiency for our six instrument flying skills when GPS was used at the primary navigation system. It important to recall that these six skills were drawn from the Instrument Rating – Airplane practical test standards (FAA, 1998) and are ones for which each pilot already had been tested formally and certified to perform while acting as pilot-in-command in instrument flight conditions. At the present time, no additional training or qualifications are required for pilots to exercise the privileges of an instrument rating using GPS as the primary means of navigation.

How much practice is needed then? In short, we have failed to answer our original question in much detail. The only answer that can be provided by the data and analysis given here is that it is likely that more than five ground-learning sessions and five flights are required for the average pilot. It is clear that another study must be done in which pilots are permitted to continue practicing until reaching a point of asymptotic performance. It must be noted for the present study that not all skills were practiced on every flight. Indeed, at the end of the training, pilots performed most poorly on skills that they had practiced only a few times. Future studies might be designed to carefully control the number of practice trials for each skill, not just the number of flights.

These results suggested that IFR GPS is not a "walk-up-and-use" system for pilots at any experience level. Considerable learning and practice are required to achieve proficiency with flying IFR with GPS.

References

Adams, C. A., Hwoschinsky, P. V., & Adams, R. A. (2001). Analysis of adverse events in identifying GPS human factors issues. *11th International Symposium on Aviation Psychology*. Columbus, OH.

Casner, S. M. & Dupuie, D. A. (Illustrator) (2002). *Cockpit Automation for General Aviators and Future Airline Pilots*. Iowa State Press.

Federal Aviation Administration (2004). *Aeronautical Information Manual*. Washington, DC: FAA.

Federal Aviation Administration (1998). Instrument rating practical test standards for airplane, helicopter, powered lift. *FAA-S-8081-4C. Flight Standards Service*. Washington, DC: FAA.

Henry, W.L, Young, G. E., & Dismukes, R. K. (1999). The influence of global positioning system technology on general aviation pilots' perception of risk during in-flight decision making. In R. Jensen (Ed.), *Proceedings of the 10th International Symposium on Aviation Psychology* (pp.147-151). Columbus, OH: Ohio State University.

Heron, R. M., Krolak, W., & Coyle, S. (1997). *A human factors approach to use of GPS receivers*. Transport Canada Aviation.

Nowinski, J. L., Holbrook, J. B., & Dismukes, R. K. (2003). Human memory and cockpit operations: An ASRS study. In R. Jensen (Ed.), *Proceedings of the 12th International Symposium on Aviation Psychology* (pp. 888-893). Dayton, OH: The Wright State University.

Talleur, D. A., Taylor, H. L., Emanuel, Jr., T. W., Rantanen, E., & Bradshaw, G. L. (2003). Personal computer aviation training devices: Their effectiveness for maintaining instrument currency. *International Journal of Aviation Psychology 13* (4), 387-399.

Appendix 1

Script of Events Used for the IFR GPS Practice Flights

Flight 1: SQL-O27-SQL

- SQL-O27
  - Program SQL-Sunol-Tracy-ECA-O27 on ground
  - Announce Sunol
  - Program VNAV ECA @ 3,000
  - Announce Tracy
  - Announce ECA
  - Announce Moter
  - Announce approach active mode
  - Announce Eltro
  - Aircraft control
- O27-SQL
  - Program O27 to SQL on ground
  - Insert Tracy and Sunol
  - Program diversion
  - Look up rwy length and frequency
  - Program Sunol to SQL
  - Aircraft control

Flight 2: SQL-MOD-SCK-LVK-SQL

- SQL-MOD
  - Program SQL-Sunol-Tracy-Cazli-MOD on ground
  - Set OBS 009 to Sunol
  - Set GPS to sequencing mode

- Announce Sunol
- Announce Tracy
- Set OBS 018 to Awoni
- Announce Awoni
- Set GPS to sequencing mode
- Announce approach active mode
- Announce Wowar
- Aircraft control

MOD-SCK

- Program MOD-SCK on ground
- Set OBS 291 to Oxjef
- Set GPS to sequencing mode once established
- Announce Oxjef
- Announce approach active mode
- Announce Ipdew
- Aircraft control

SCK-LVK

- Program SCK-LVK on ground
- Set OBS 246 to Uhhut
- Set GPS to sequencing mode
- Announce Uhhut
- Announce approach active mode
- Announce Oyahi
- Aircraft control

Flight 3: SQL-STS-KDVO-O69-SQL

SQL-STS

- Program SQL-STS
- Set OBS 321 to Zijbe
- Set GPS to sequencing mode
- Announce Zijbe
- Announce approach active mode
- Announce Gokuw
- Aircraft control

STS-DVO

- Program STS-DVO on ground
- Set OBS course to Oriby
- Announce Oriby
- Announce approach active mode
- Announce Eyeji
- Program direct to SGD
- Set OBS 180 to SGD for hold
- Program SGD-O69
- Aircraft control

DVO-O69

- Set OBS 268 to Ipary

Set GPS to sequencing mode when established
Announce approach active mode
Announce Ipary
Aircraft control

Flight 4: SQL-MRY-WVI-HAF-SQL

SQL-MRY

Program SQL-OSI-Sapid-Santy-Mover-SNS-Llynn-MRY on ground
Engage Heading Select
Engage VS and arm Altitude Hold
Set OBS 141 to Sapid
Arm Nav to capture course
Set GPS to sequencing mode
Announce Sapid
Engage VS and arm Altitude Hold
Announce Santy
Engage Heading Select
Set OBS 286 to Raine
Arm Approach to capture course
Set GPS to sequencing mode when established
Announce approach active mode
Announce Raine
Announce 7.2NM waypoint

MRY-WVI

Program MRY-WVI on ground
Engage VS and arm Altitude Hold
Set OBS 314 to Dyner
Arm Approach to capture course
Set GPS to sequencing mode when established
Announce approach active mode
Announce Dyner

WVI-HAF

Program WVI-HAF on ground
Announce Giruc
Set GPS to OBS mode for hold
Set GPS to sequencing mode
Engage Approach to capture course
Announce approach active mode
Announce Wohli

Flight 5: SQL-O27-SCK-1O3-LVK-SQL

SQL-O27

Program SQL-Sunol-Tracy-ECA-O27 on ground
Announce Sunol
Engage VS and arm Altitude Hold
Program VNAV ECA @ 3,000
Engage VS and arm Altitude Hold

Announce Tracy
Set OBS 090 to Moter
Engage Heading Select and arm Approach
Set GPS to sequencing mode
Announce Moter
Announce approach active mode
Announce Eltro
Program direct Wraps
Use autopilot to accomplish missed approach
Set OBS 180 Wraps for hold
Announce Wraps

Wraps-SCK
Program Wraps-SCK
Set OBS 234 to Oxjef
Engage VS and arm Altitude Hold
Engage Heading Select and arm Approach
Set GPS to sequencing mode when established
Announce approach active mode
Announce Ipdew

SCK-1O3
Program SCK-1O3
Set OBS 285 to Quads for PT
Use autopilot to accomplish PT
Announce Quads
Set GPS to sequencing mode inbound to Quads
Engage approach function
Announce approach active mode
Announce Quads

1O3-LVK
Program 1O3-LVK
Engage VS and arm Altitude Hold
Set OBS 246 to Uhhut
Engage Heading Select and arm Approach
Set GPS to sequencing mode when established
Announce Uhhut
Announce approach active mode
Announce Oyahi

International Journal of Applied Aviation Studies, 4, Number 2

Using Binary Logistic Regression to Screen Candidates for Instructor Pilot Upgrade

Raymond A. Hamilton II and Robert Ripley

Auburn University
Aviation Management
Lowder Business Building
415 W. Magnolia Avenue
Auburn, AL 36849

Abstract

Researchers who focus on aircrew performance and its relation to aviation safety have endeavored to understand personality traits and attributes and their effect on team coherence and performance. Moreover, corporate aviation departments and commercial air carriers have vested interests in hiring those applicants who have the propensity to perform effectively in a team environment. As a continuation of ongoing research in pilot resource selection, the authors have examined the efficacy of binary logistic regression in screening candidates for instructor pilot upgrade by analyzing psychological surveys of 77 commercial pilots based on the Big Five Personality Inventory, the Leadership Intelligence Quotient (LIQ) Self-Assessment and three select attribute scales (Mach V, locus-of-control, and self-monitoring scales). Results of this research indicated a significant relationship between leader decision-making and personality attributes complementing team formation. In this particular study, 36 experienced instructor pilots and 41 line pilots were categorized with an overall accuracy of 93 percent employing a two-variable model based on Machiavellian view and self-monitoring as independent or predictor variables. These results underscored the efficacy of logistic regression as a statistical methodology in human resource selection and these psychological dimensions as a means of screening pilot candidates for program assignment.The demand among commercial air carriers and corporate aviation for greater consistency and validity of aircrew re-

---

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

> *source selection methods grows continuously. In the past, the emphasis has been on evaluating technical qualifications and operational experience. There has been no shortage of applicants meeting the basic standard in these areas. The challenge today and in the future is to go beyond that stage of evaluation to develop more effective means of measuring the social aspects of the human variable in the aviation safety equation. Selection protocols are under close, continuous scrutiny and now include personality and temperament characteristics related to crew coordination and team assimilation (Chidester, 1991). Research suggested that in the past selection protocols have suffered in predictive validity with content remaining basically unchanged (Damos, 1995). These protocols in the past have indicated higher predictive validity in the training program, not in the operational environment (Damos, 1996).*
>
> *Effective screening of pilot candidates, whether the purpose is to identify instructor upgrades or to select hires, will improve the return on investment of sunk costs related to training (Hackman, 1993).*

Research Questions

This research seeks to continue the development of a comprehensive pilot screening model based on the personality and leadership temperament of the optimum pilot cadre. To this end, we raised the following research questions:

1. Are there common personality traits and attributes shared among professional instructor pilots who have demonstrated effective flight crew leadership during Line Oriented Flight Training (LOFT) training?
2. Are there instruments or elements of instruments that have the discriminatory power to accurately categorize candidates who possess effective aircrew leadership qualities?

Methodology

We elected to employ binary logistic regression analysis since our dependent variable, pilot status is dichotomous (e.g., instructor pilot – line pilot) and hence nonmetric. As with multiple regression, the independent variables are assumed to be metric. Logistic regression analysis applies in those cases when the total sample can be divided into groups based on a nonmetric dependent variable characterizing several known classes or categories. Using this methodology, our intent was to understand categorical differences and to predict the likelihood that an individual will belong to a particular category based on several metric independent variables. In our analysis, binary logistic regression analysis was used to distinguish instructor pilots from line-assigned pilots according to their psychographic profiles as measured by the following select survey instruments, which are explained in the next section:

1. Mach V Machiavellianism Scale
    a. Tactics subscale
    b. Views of people subscale

2. Leadership Intelligence Quotient (LIQ) Self-evaluation Questionnaire
3. Self-monitoring Scale
4. Locus-of-Control Scale
5. Big Five-factor model of personality
   a. Extroversion subscale
   b. Agreeableness subscale
   c. Conscientiousness subscale
   d. Emotional stability subscale
   e. Openness-to-experience subscale

Instrumentation

We selected one instrument to evaluate personality traits, three instruments to evaluate attributes, and one to evaluate pragmatic leadership knowledge.

Personality Traits

We selected the five-factor model of personality—referred to as the "Big Five" (Smith, 2001). Recent research strongly underscored the five basic dimensions of the Big Five (described below) as constituting the greater part of the scope of human personality.

1. Extroversion. This factor is concerned with the individual's comfort level with relationships. Extroverts are oriented toward gregariousness, assertiveness, and sociability. In contrast, introverts are more reserved, quiet, and timid.
2. Agreeableness. This factor refers to a propensity to yield or defer to others. Those who are highly agreeable are trusting, cooperative, and warm in their demeanor toward others. Those who possess a lower score on this factor are relatively more disagreeable, antagonistic, and cold or withdrawn from others.
3. Conscientiousness. Reliability is the main focus of this factor; those scoring high tend to be dependable, persistent, organized, and responsible. In contrast, low scorers tend to be disorganized, unreliable, and become distracted easily.
4. Emotional stability factor centers on people's ability to deal with stressful situations. They tend to be self-confident, calm, and secure in their situation. Lower scores suggest nervousness, insecurity, and anxiety
5. Openness to experience refers to a person's scope of interests and preoccupation with innovation and change. Those scoring lower tend toward the conventional approach to problem solving and are more comfortable with a routine to which they are adapted.

Research conducted using this instrument has examined a variety of occupations and work environments (Judge, 2001). Results indicated that conscientiousness relates positively to job performance across the spectrum of employment from professional to semi-skilled jobs (Mount, 1994). Consistent with these findings, evidence also indicated a pronounced relationship between conscientiousness and "organizational citizenship behavior." Other dimensions related to job performance depending on the nature of work. For example, extroversion correlated positively to job performance in managerial positions; while, openness-to-experience correlated with performance in the training environment (Podsakoff, 2000).

We would expect an effective instructor to reflect above-the-norm agreeableness, conscientiousness, and emotional stability—key elements to performing as a strong team leader exercising referent power. It is important, however, to realize that to try categorize pilots according to personality type is difficult if not impossible (King, 1999, p. 21).

Leadership Acumen

As a surrogate for leadership skill, we chose the Leadership Intelligence Quotient Self-evaluation Questionnaire (Murphy, 1996). The instrument draws on research into the beliefs and practices of more than 18,000 organizational leaders in 562 public and private organizations. From this database, 1,029 individuals were identified by Murphy (1996) as highly effective leaders representing the full scope of organizational levels. Murphy termed them "workleaders," denoting the combination of work and leadership, which he viewed as comprising the essence of effective leadership. In Murphy's model, workleaders know how to:

- Select the right people,
- Connect them to the right cause,
- Solve problems that arise,
- Evaluate progress towards objectives,
- Negotiate resolutions to conflicts,
- Heal the wounds inflicted by change,
- Protect their cultures from the perils of crisis, and
- Synergize all stakeholders in a way that enables them to achieve improvement together (Murphy, 1996).

Consisting of 36 items, the Leadership Intelligence Quotient (LIQ) Self-Assessment instrument measures the degree to which a person conforms to Murphy's workleader model in making decisions affecting interpersonal relations and their consequence for organizational benefit. The instrument presents scenarios that leaders face daily with decision options that reflect both conventional wisdom and non-traditional approaches. Based on experience and leadership skill associated with a pilot selected for instructor upgrade, one would expect instructors to make decisions that positively affect interpersonal relations complementary to mission goals.

Personality Attributes Influencing Team Member Interaction

We chose the following attributes because of their reported strength in predicting human behavior in organizations.

Locus of Control. Locus of Control concerns the degree to which individuals view themselves in control of their own destiny and circumstance. Those who believe thus are termed "Internals." Individuals who are of the view that what happens to them is controlled by random outside factors are referred to as

"Externals." An individual's perception of the controlling element determining personal fate is referred to as "locus of control" (Rotter, 1966). This instrument consists of 23 items measuring two dimensions— external and internal locus of control. Research suggested that individuals scoring as externals experience more job dissatisfaction, absenteeism, alienation in the work environment, and have a weaker commitment to their job (Blau, 1987). Internals show higher motivation and initiative during the employment interview process (Cook, 2000). These results implied that externals might experience job dissatisfaction because of perceived lack of control over decisions affecting their well-being and interests. In contrast, internals might view events and outcomes in the organization as attributable to their personal actions and performance.

How turnover relates to locus-of-control is complex. Internals are prone to take action and to terminate more often; but, at the same time, they tend to be more successful on the job and hence more satisfied then their external counterparts (Spector, 1982). Given the strong cultural emphasis in accepting personal responsibility as a professional pilot, we would anticipate instructor pilots in particular to possess a pronounced internal locus-of-control.

Machiavellianism. Machiavellianism refers to a personality characteristic named after Niccolo Machiavelli, a sixteenth century writer who wrote on the subject of power—how to gain it and keep it. Christie and Geis (1970) developed the Mach IV and Mach V instruments to measure three dimensions: tactics in dealing with people, views of the nature of people, and conventional morality. Persons who score high on these instruments emphasize pragmatism, maintain emotional distance, and view the ends justifying the means. Extensive research has focused on the relationship of high- and low-Mach personalities to organizational and group behavior (Vleeming, 1979). We employed the Mach V, a triadic, forced-response scale, to measure the two primary dimensions of Machiavellianism: tactics in dealing with people and views of the nature of people. Being a forced-response instrument, the Mach V is preferable since it is less susceptible to strategic responses. The Mach V is a triadic, forced-response scale consisting of twenty items—nine measuring the tactics dimension, nine measuring the views dimension, and two measuring conventional morality. We elected to dispense with this third dimension and focused on the means by which the respondent endorses deceptive means in controlling others and the overall view the respondent has of the general nature of people. Those scoring high on the Mach V scale tend to be more successful when competing for control of the group, less susceptible to persuasion, and are able to persuade others more successfully than those scoring low on the scale (Christie, 1970). Nevertheless, the situation has some influence on the relative success of the high Mach. Face-to-face interaction, not being encumbered by rules and regulations and little effect of interpersonal relationship favor the high Mach's success. Low Machs are more successful in structured situations where human interaction is indirect. At the same time, research indicated low Machs are more susceptible to emotional involvement with details that do not relate to success in achieving goals (Christie & Geis, 1970).

We would expect effective instructors to exhibit skills in influencing others through the exercise of referent power versus power emanating from authority or coercion. This would suggest that they possess a relatively high Machiavellian orientation regarding tactics in dealing with others.

Self-monitoring. Self-monitoring is a trait related to a person's ability to adjust personal behavior to the situation at hand (Day, 2002). Those who score high reflect adaptability in adjusting their behavior to meet the challenge of acceptance in social situations. Low self-monitors are less capable of this social adjustment and tend to show their internal dispositions and attitudes regardless of the situation. The self-monitoring scale we employed in this study consisted of 29 items designed to measure the respondent's self-perception of personal ability to adjust behavior to situational factors. Those who score as high self-monitors display sensitivity to external cues and can present a public persona in marked contrast to their private self. Low self-monitors lack this ability and proclivity to alter their public persona. Research in this area is in the embryonic stage, but it does suggest that there might be a strong correlation with the tactics dimension of the Mach V scale. That being the case, the expectation would be that high self-monitors are successful in supervisory positions where individuals are expected to present various personas depending on with whom they are interacting and the setting (Mehra, 2001). Effective instructors would emerge as high self-monitors with the ability to adapt to varying social conditions and personalities exhibited in the cockpit.

Procedure

For the first group (the instructors), the population consisted of flight officers in examiner or instructor status for their respective crew positions who had excelled in LOFT training. All were employed by certified commercial air carriers and all but six were actively flying in the Air National Guard or the Air Force Reserve. The second group (line pilots) was comprised as well of flight officers employed by a commercial air carrier. Data on these two groups was collected over a two-year period 1998-2000 during one of the author's service as a flight officer with a major air carrier.The respondents completed the instruments in confidence and received guarantees of anonymity regarding the results.

Estimation of the Logistic Regression Model and Assessing Overall Fit

Statistical Assumptions

The research problem was to determine if differences in personality traits and attributes can distinguish between commercial flight officers serving as line pilots versus instructor pilots. The analysis sample consisted of 77 respondents—each employed by a Part 121 air carrier at the time of survey completion. Logistic regression served well as a statistical method for this study, given the robustness of logistic regression to violation of the assumption of equality of the variance/covariance matrices across groups.

Model Estimation

Logistic regression is estimated similar to multiple regression since a base model is estimated initially to produce a standard for comparison. In multiple regression, the mean is used to set the base model and calculate the total sum of squares. A like procedure is employed in logistic regression with the mean used to generate a log likelihood value in lieu of setting the sum of squares as is done in multiple regression. This model allows partial correlations for each variable to be established and the most discriminating variable identified honoring the criteria for selection.

Table 1 presents the base model results for the logistic regression analysis. The log likelihood value (-2LL) here is 106.745. The score statistics are measures of association used to identify the variable for selected in the stepwise procedure. We employed the reduction in the log likelihood value (-2LL) value as our choice of criteria. As seen in the first step of the estimation procedure (Table 2), X5[SELFMN] (self-monitoring score) and X1[TACT] (tactics score derived from the Mach instrument) were selected for entry.These two variables had the highest score statistics—37.948 and 37.622, respectively.

Table 1

Logistic Regression Base Model Results

OVERALL MODEL FIT -2 log likelihood (-2LL): 106.475			
VARIABLES NOT IN THE EQUATION			
Variable		Score [df=1]	Significance
X1[TACT]	Machiavellian Tactics	37.622	.000
X2[VIEW]	Machiavellian Views	7.957	.005
X3[LIQ]	Leadership IQ	29.344	.000
X4[LOC]	Locus of Control	31.377	.000
X5[SELFMN]	Self Monitoring	37.948	.000
X6[EXPER]	Openness to Experience	.002	.963
X7[CONSCI]	Conscientiousness	33.904	.000
X8[EXTRAV]	Extraversion	.341	.559
X9[AGREEA]	Agreeableness	34.600	.000
X10[EMOSTA]	Emotional Stability	4.661	.031

Although the entry of X5 and X1 obtained a reasonable model fit, a look at the remaining independent variables not in the equation indicated that five meet the threshold of .05 significance thus suggesting inclusion. Table 3 shows the entry variables and the remaining four steps to include the corresponding change in the -2 log likelihood (-2LL). Note that the improvement in all measures of model fit ranging from a decrease in the -2LL value to the various R^2 measures.

Table 2

Step 1: Entry of X5 and X1 (Self Monitoring and Machiavellian Tactics) Stepwise Logistic Regression Model					
OVERALL MODEL FIT					
Goodness of Fit Measures	*Value*	*Change in -2LL*			
-2 log likelihood (-2LL)	54.652			*Value*	*Significance*
Cox and Snell R Square	.492	From base model		52.093	.000
Nagelkerke R Square	.655	From Prior step		52.093	.000
	Chi-square		*df*	*Significance*	
Hosmer and Lemeshow	1.940		8	.983	
VARIABLES IN THE EQUATION					
Variable	*B*	*S.E.*	*Wald*	*Sig.*	*Exp(B)*
X5	-.273	.495	16.273	.000	.761
X1	.164	.041	15.987	.000	1.178
CORRELATION MATRIX					
		X5		X1	
X5		1.000		-.997	
X1		-.997		1.000	
CLASSIFICATION TABLE			Y0[INST]	Y1[PLT]	Percentage Correct
Y0[INST]			29	7	80.6
Y1[PLT]			7	34	82.9
Overall Percentage					81.8

Again, although extremely high levels of model fit were obtained in the first step with the two entry variables, X5 and X1, examination of the variables not in the equation indicated that other variables could enter the stepwise solution possibly increasing the discriminatory power of the model.

Table 3

Model Summary			
Entry Variables	-2LL	Cox & Snell R^2	Nagelkerke R^2
Initial State	106.745	.492	.655
X5 & X1	54.652	.492	.655
X9 & X7	49.878	.522	.697
X4	48.781	.529	.705
X2	40.254	.578	.771
X10	39.839	.581	.774

Statistical Significance

There are two statistical tests for the significance of the final model (Table 4). First, a chi square test for a change in the -2LL value from the base model is comparable to the overall F test in multiple regression. In the seven-variable model (Table 4), this reduction is statistically significant at the .000 level. In addition, the Hosmer and Lemeshow (1989) measure of overall fit indicated statistically no significance difference in the observed and predicted classifications (Table 5). These two measures in combination provide support for acceptance of the seven-variable model (Table 4) as a viable binary logistic regression model and suitable for further examination.

Table 4

Step 5: Entry of X10 (Emotional Stability of The Big Five) Stepwise Logistic Regression Model				
OVERALL MODEL FIT				
Goodness of Fit Measures	*Value*	*Change in -2LL*		
-2 log likelihood (-2LL)	39.839		*Value*	*Significance*
Cox and Snell R Square	.581	From base model	66.906	.000
Nagelkerke R Square	.774	From Prior step	.415	.519
	Chi-square	*df*	*Significance*	
Hosmer and Lemeshow	6.273	8	.617	

VARIABLES IN THE EQUATION					
] *Variable*	*B*	*S.E.*	*Wald*	*Sig.*	*Exp(B)*
X5	-.477	.235	4.130	.042	.621
X1	.000	.154	.000	1.000	1.000
X9	-.110	.158	.492	.483	.895
X7	.321	.189	2.884	.089	1.379
X4	.291	.443	.432	.511	1.338
X2	.112	.048	5.487	.019	1.118
X10	.098	.153	.408	.523	1.103

CORRELATION MATRIX							
	X5	X1	X9	X7	X4	X2	X10
X5	1.000						
X1	-.232	1.000					
X9	.157	-.755	1.000				
X7	-.915	.095	-.196	1.000			
X4	-.316	-.439	.153	.224	1.000		
X2	.225	-.250	-.244	-.016	-.102	1.000	
X10	-.199	.164	-.190	.203	-.009	-.196	1.000

CLASSIFICATION TABLE	INST	PLT	Percentage Correct
INST	34	2	94.4
PLT	3	38	92.7
Overall Percentage			93.5

Assessing Overall Model Fit

In assessing model fit, several measures are available. First, the -2LL value is given. Note that the -2LL value is reduced from the base model from 106.745 to 39.839 with the final entry of the seven variables (Table 3). Reduced values of the -2LL value indicate an improved model fit. The goodness-of-fit measure compares the predicted probabilities to the observed probabilities with higher values indicating a better model fit. There is no upper or lower limit for this measure and the value for the seven-variable model is 39.839 (Tables 3 and 4).

Next, three measures comparable to the R squared measure in multiple regression are available. The Cox and Snell measure operates in the same manner with higher values indicating greater model fit. However, this measure is limited since it cannot reach the maximum value of 1. Nagelkerke proposed a modification that has a range from 0 to 1 (Hosmer, 1989). In our model, the Cox and Snell and Nagelkerke R Squares had a value of .581 and .774, respectively (Table 4). The third measure of model fit is the Hosmer and Lemeshow value, which measures the correspondence of the actual and predicted values of the dependent variable. In this case, better model fit is indicated by a smaller difference in the observed and predicted classification. A good model fit also is demonstrated by an insignificant chi square value (Table 5).

Table 5

Hosmer and Lemeshow Test			
Entry Variables	Chi Square	df	Significance
X5 & X1	1.940	8	.983
X9 & X7	6.545	8	.586
X4	8.532	8	.383
X2	27.164	8	.001
X10	6.273	8	.617

In the seven-variable model (Table 4), all the measures of model fit showed improvement over the two-variable model (Table 2). The -2LL value decreased to 39.839, the Cox and Snell R Square increased from .492 to .581 and the Nagelkerke R Square increased from .655 to .774 (Table 3). All the improvements from the double-variable model (Table 2) reflected in the seven-variable model have indicated a good model fit. The Hosmer and Lemeshow measure still showed no statistical significance in differences in the distribution of the actual and predicted values of the categorical (dependent) variable.

Finally, the classification matrices identical in nature to those used in discriminant analysis showed extremely high hit ratios of correctly classified cases for the seven-variable model (Tables 4 and 6). The overall hit ratio was 93.5 percent; likewise, the individual group hit ratios were consistently high and did not indicate a problem in predicting the two groups (Instructor pilots versus line pilots). The seven-variable model demonstrated excellent model fit and established statistical significance as well as the variables included in the model.

Table 6

Classification Table*				
Entry #	Variable(s)	Instructor Pilots	Line Pilots	Overall Correct Pct.
Initial	None	0 (0.0%) 36 Pilots	0 Instructors 41 (53.2%)	53.2%
1	X5 & X1	29 (80.6%) 7 Pilots	7 Instructors 34 (82.9%)	81.8%
2	X9 & X7	30 (83.3%) 8 Pilots	6 Instructors 33 (80.) 5%	81.8%
3	X4	29 (80.6%) 7 Pilots	7 Instructors 34 (82.9%)	81.8%
4	X2	32 (88.9%) 4 Pilots	3 Instructors 38 (92.7%)	90.9%
5	X10	34 (94.4%) 2 Pilots	3 Instructors 38 (92.7%)	93.5%

*Values in parentheses are the correct percentages for each category.

Case-wise Diagnostics

The analysis of the misclassification of individual observations provided further insight into the possible improvements of the model; but in our study, there were only two misclassified cases providing an inadequate basis for making any further analysis. Case-wise diagnostics such as residuals and measures of influence were available, but were of little use in these results. Given the low level of misclassification, no further analysis of misclassifications was performed.

Interpretation of the Results

The logistic regression model produced a variate that reflected a positive association of self-monitoring, Machiavellian tactics, three of the Big Five measures (agreeableness, conscientiousness, and emotional stability), and locus of control with the dependent variable, while Machiavellian views reflected a negative association.

Conclusions and Recommendations

This study sought the specific objective of predicting instructor status of commercial flight officers using binary logistic analysis. The resulting model demonstrated excellent discriminant power in sorting instructor flight officers from those line-assigned. The study's broader objective of assessing the viability of binary logistic regression in screening pilot candidates for upgrade or initial hire was achieved. To establish a broadly based, consistent, and reliable pilot-screening model, further study in researching a full range of instruments is warranted. The optimum selection/screening model will more than likely be based on a composite instrument, which draws on select elements of various instruments that add to the model's overall discriminatory power.The success of this methodology employed for this purpose rests on the selection of the referent group—the members of which should ideally represent the standard in flight crew performance. Such a model should be employed to select-in, not select-out candidates. The select-out process refers to screening of candidates based on medical criteria regarding presence of psychiatric disorders disqualifying individuals from consideration (Santy, 1994). In contrast, select-in criteria are psychological selection criteria crucial to personnel selection and are not based on psychopathology but on social characteristics relevant to commercial flight crew performance. The ultimate goal is to pick the best person for the job. A selection model using binary logistic regression and based on a quality, high-performance reference group would add greatly to commercial aviation safety.

References

Blau, G. J. (1987). Locus of control as a potential moderator of the turnover process. *Journal of Occupational Psychology 60*, 21-29.

Chidester,T. R., Helmrich, R. L., Gregorich, S. E., & Geis, C. E. (1991). Pilot personality and crew coordination: implications for training and selection. *The International Journal of Aviation Psychology 1* (1).

Christie, R. & Geis, F. L. (Eds.) (1970). *Studies in Machiavellianism*. New York: Academic Press.

Cook, K. W.,Vance, C. A., & Spector,P. E. (2000). The relation of candidate personality with selection-interview outcomes. *Journal of Applied Social Psychology 30*, 867-85.

Damos, D. (1995). *Pilot selection batteries: a critical examination*. Aldershot, United Kingdom: Avebury Technical.

Damos, D. (1996). Pilot selection batteries: shortcomings and perspectives. *The International Journal of Aviation Psychology, 6* (2), 199-209

Day, D. V., Scheicher, D. J., Unckless, A. L., & Hiller, N. J. (2002). Self-monitoring personality at work: A meta-analytic investigation of construct validity. *Journal of Applied Psychology, 87* (2) 390-401.

Hackman, J. R. (1993). *Teams, leaders, and organizations: new directions for crew-oriented flight training*. San Diego, CA: Academic Press.

Hosmer, D. W. & S. Lemeshow (1989). *Applied Logistic Regression*. New York: Wiley.

Judge, T.A. & Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology 86* (1), 80-92.

King, R. E. (1999). *Aerospace clinical psychology*. Brookfield, VT: Ashgate.

Mehra, A., Kilduff, M., & Brass, D. J. (2001). The social networks of high and low self-monitors: Implications for workplace performance. *Administrative Science Quarterly 46*, 390-401.

Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the big five personality factors. *Journal of Applied Psychology 79* (2), 272-280.

Murphy, E. C. (1996). *Leadership IQ: a personal development process based on a scientific study of a new generation of leaders*. New York: John Wiley and Sons, Inc.

Podsakoff, P. M., MacKenzie, S. B., Paine, J. B., & Bachrach, D. G. (2000). Organizational citizenship behaviors: A critical review of the theoretical and empirical literature and suggestions for future research. *Journal of Management 6* (3), 513-63.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs 80* (609).

Santy, P.A. (1994). *Choosing the right stuff: the psychological selection of astronauts and cosmonauts*. Westport, CT: Praeger Publishers.

Smith, D. B., Hanges, P. J., & Dickson, M. W. (2001). Personnel selection and the five-factor model: Reexamining the effects of applicant's frame of reference. *Journal of Applied Psychology 86* (2), 304-15.

Spector, P. E. (1982). Behavior in organizations as a function of locus of control. *Psychological Bulletin 91*, 482-97.

Vleeming, R. G. (1979). Michiavellianism: A preliminary review. *Psychological Reports 44*, 295-310.

International Journal of Applied Aviation Studies, 4, Number 2

Visual Detectibility of Dents on a Composite Aircraft Inspection Specimen: An Initial Study

David Erhart

Fairfield Services Group
Idaho Falls ID 83402, USA
(208) 523-7021
email: erhart@gobigwest.com

Lee T. Ostrom

University of Idaho
1776 Science Center Drive
Idaho Falls, ID 83402, USA
(208) 282-7903
facsimile: (208) 282-7950
email: ostrom@if.uidaho.edu

and

Cheryl A. Wilhelmsen

University of Idaho
1776 Science Center Drive
Idaho Falls, ID 83402, USA
(208) 282-7718
facsimile: (208) 282-7950
email: cherylw@if.uidaho.edu

Abstract

During the ground inspection of an aircraft, ground crews inspect the wings, inspection specimens, and fuselage of the aircraft for damage. This damage can take the form of cracks, holes, corrosion, dents, or other anomalies. It has been shown that a dent with a depth of 0.05 inches can indicate the potential delamination of composite materials. Without any special tools and in all kinds of weather conditions, these inspections are conducted and are critical for safe flight. This research focused on determining the level of detectiblity of dents and determining the probability of detection of 0.05 inch deep dents.

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

Introduction

The current and next generation of aircraft will contain a much greater percentage of structural components that are made of composite materials. The Boeing Commercial Aircraft Company's (Boeing) 7E7 aircraft, for instance, will have much more composite material than current generation Boeing commercial airplanes. Composite materials can delaminate due to impact and the visual indication of a delamination can be very subtle. This type of damage is called "Barely Visible Impact Damage" (BVID) (Poon, Benak, & Gould, 1990). Numerous nondestructive testing (NDT) techniques have been developed to find these types of delaminations. These techniques include using ultrasonics, X-ray, infrared, and laser technologies. However, the visual detection of anomalies by airline inspectors and other airline personnel has been and will continue to be a major component of the aircraft inspection process.

Visual inspection of aircraft for structural damage has been the main defense against in-flight aircraft breakup since the beginning of flight. Goranson and Rogers (1983) reported that over 80% of all aircraft inspections are conducted visually.A great amount of research has been conducted on the visual inspection of metal, structural components. The Visual Inspection Research Project is a good example of a very thorough and well-developed research project, designed to determine the effect of certain variables on the probability of detecting cracks on aircraft structure (Spencer, 1996). Research on this topic was performed before and after this project and there continues to be a great amount we wish to understand about how human inspectors perform inspections. Research on visual inspection of composite materials is in its infancy.There is little literature on this topic. Though, there is a great amount of literature on the equipment used to perform non-destructive testing/inspection of composites (Hsu, et al., 2002), there is little research that correlates BVID surface anomalies with interior defects. This paper is an initial examination on the topic of visual inspection of composite materials for surface anomalies, mainly dents. We performed this initial, limited study to begin to understand the variables associated with this topic and as a test bed as to how to perform future research.

The commercial aircraft industry was interested in determining the visual probability of detection (POD) of 0.05-inch dents. A 0.05-inch dent is considered a reasonable lower bound cutoff value for visual indication of potential delaminations. These and larger dents would then be considered for further NDT to determine whether delamination is occurring. Aircraft manufacturers specify the types of NDT to use.

In prior studies, we developed POD curves for detecting cracks in metal structural components of commercial aircraft and developed an automated tool for developing these data from inspection data (Ostrom, Wilhelmsen, Valenti, & Kanki, 2002). In this current work, we begin a similar process for dents in composite material. As a part of this research, Boeing asked the University of Idaho to conduct a pilot study to determine what size dent the average person would detect. Northwest Airlines provided two previously rejected inspection specimens off one of their Boeing aircraft for use in this study as research specimens.

Methods

This section will discuss the inspection specimens used in the study, how the dents were created, the design of the study, and the process with which the study was conducted. It is important to note that this was an initial study into the topic of detecting dents in composites and as such did not study even a fraction of the possible experimental variables.

Two inspection specimens were used to conduct this test. Inspection specimen A was 28 inches wide and inspection specimen B was 32 inches wide. One-half of each inspection specimen was cleaned from top to bottom and the other half left dirty.The Inspection specimens were displayed on either side of a separating wall (see Figure 1). Inspection specimen B was displayed in ambient light and Inspection specimen A was displayed under a lamp to simulate a bright flashlight.

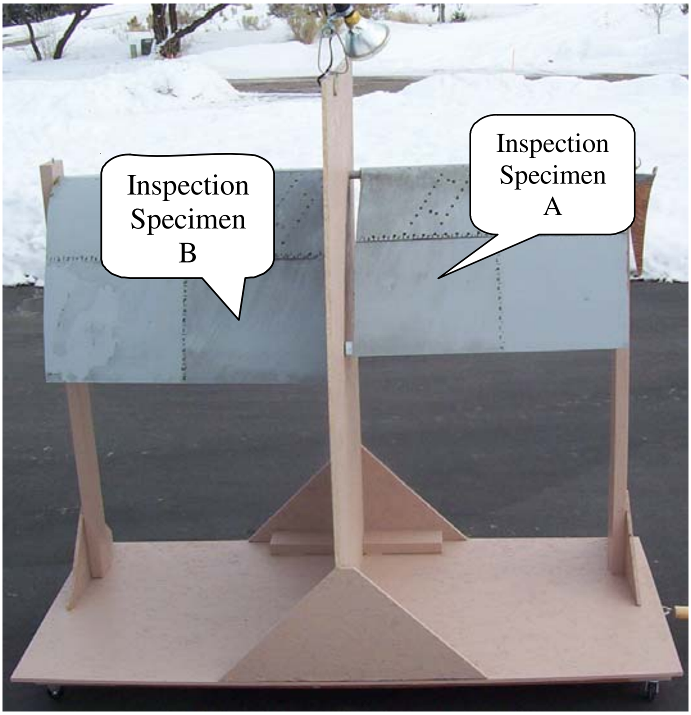


Figure 1. Inspection specimen Test Display

The inspection specimens were displayed at a 30 to 45 degree angle. This was to simulate the position they would be found on a plane after landing. Both the top and bottom surfaces of the inspection specimen were accessible for inspection. A poster was included atop the display to show the participants where these inspection specimens were located on the aircraft.

Defects (dents) in the inspection specimen were produced by dropping a 0.5 oz. lead ball or an 8 oz. lead weight from a variety of heights. The data for creating the dents are contained in Table 1. While trying to create a consistent dent, two problems were noted. The difference between a dent and failure of the composite surface was not always achievable. The weight dropped from the same height in different locations on the inspection specimen would produce different results. If the weight was dropped near the trailing edge of the inspection specimen, a small dent was created. This situation was also repeated when the weight was dropped at the upper end of the carbon fiber section. The same weight dropped in the center of the carbon fiber section from the same height would cause complete failure of the carbon fiber surface. There was no effort to control the volume of the dents.

The dents were measured using a depth probe on a digital caliper.A depth micrometer was found to be unsuitable, as the face of the depth micrometer did not always reach the bottom of the small defects. To provide a flat measuring surface, a piece of shim stock with a hole in the center was used to bridge the defect.

Table 1
Inspection Specimen Dents.

Inspection specimen A										
Top						Bottom				
X Grid	Y Grid	Dent Depth	Drop Distance	Weight		X Grid	Y Grid	Dent Depth	Drop Distance	Weight
E	19	0.004	existing			J	12	0.007	15	8 oz
F	20	0.004	existing			P	26	0.007	24	8 oz
G	16	0.006	50	8 oz		M	24	0.009	20	8 oz
S	12	0.010	24	8 oz		M	24	0.016	200	.5 oz
L	17	0.013	24	8 oz		O	17	0.016	35	8 oz
G	3	0.016	75	8 oz		O	10	0.021	200	.5 oz
E	23	0.019	70	8 oz		J	17	0.023	30	8 oz
M	10	0.050	30	8 oz		B	27	0.024	75	8 oz
B	24	0.061	70	8 oz		L	5	0.038	30	8 oz
P	24	0.076	100	8 oz		D	5	0.061	100	8 oz
O	4	0.085	75	8 oz						

Table 1 *(Continued)*
Inspection Specimen Dents

Inspection specimen B										
Top						Bottom				
X Grid	Y Grid	Dent Depth	Drop Distance	Weight		X Grid	Y Grid	Dent Depth	Drop Distance	Weight
U	26	0.008	15	8 oz		F	3	0.004	80	8 oz
E	21	0.013	70	8 oz		K	9	0.007	200	.5 oz
M	4	0.013	30	8 oz		N	16	0.009	200	.5 oz
M	10	0.019	30	8 oz		R	5	0.010	24	8 oz
W	4	0.020	30	8 oz		K	24	0.016	20	8 oz
O	18	0.030	20	8 oz		S	25	0.025	24	8 oz
D	14	0.031	70	8 oz		C	29	0.037	80	8 oz
G	12	0.033	70	8 oz		K	3	0.041	30	8 oz
N	25	0.035	24	8 oz		C	19	0.047	existing	
R	12	0.043	24	8 oz		E	19	0.063	80	8 oz
E	7	0.139	80	8 oz						

A grid pattern was laid out on the top and bottom of each inspection specimen to aid in locating the defects. This grid pattern was repeated on the inspection forms. The participants were asked to mark the inspection form where they found the defect on the wing. The participants were not told what depth of dent they were looking for. In the evaluation of the dents, we would give credit if the mark on the data sheet was within one unit of the actual location. In the example shown in Figure 2, the defect is located at Q-24.

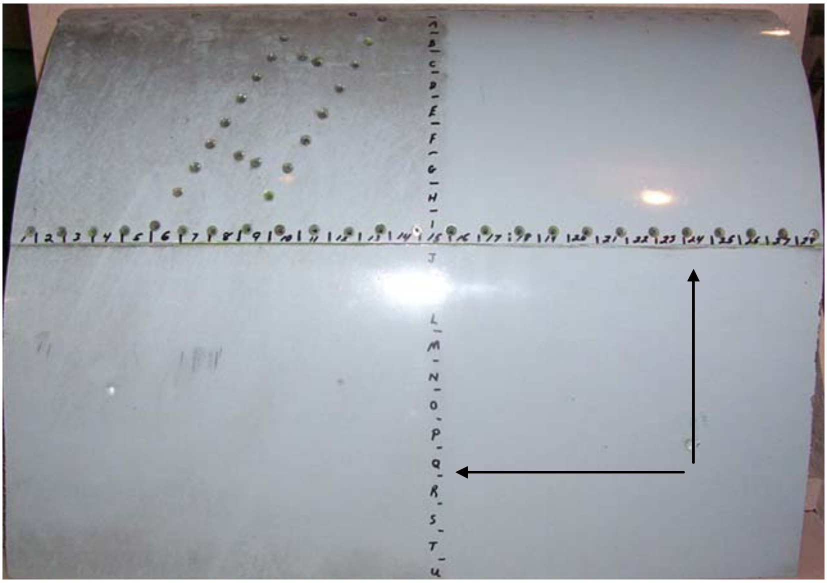


Figure 2. Grid pattern on inspection specimen A.

Inspection Variables

The preflight inspection of an airplane can occur at any time of day and in any weather/lighting conditions. It was our intention to examine two of the numerous possible variables. These were:

1. Clean vs. Dirty Surface
2. Bright vs. Ambient Light

To do so, one-half of each inspection specimen was cleaned of all oxidation, grease, and dirt. By having a clean and dirty side half of each wing, we wanted to know how well one could see the defects on different wing conditions.

In addition, a directed light was provided over inspection specimen A to simulate a bright flashlight. Inspection specimen B was displayed in existing ambient light conditions. However, due to the problem of not being able to make dents of a consistent size in the composite material we were not able to statistically examine these variables. We will discuss this further in the Discussion Section of this paper.

There were 73 total participants in the study. The sex and age of the participants were captured and compared in the study.

The age of the participants was broken into five groups:

- <20 years old
- 21–30 years old
- 31–40 years old
- 41–50 years old

Data Gathering Process

As this is an initial study with future detailed studies anticipated, no attempt was made at selecting a "target" audience. For this initial study, a wide range of participants was selected with a broad range of training and work experiences. Three facilities were selected for the gathering of data. The Test Reactor Area (TRA) at the Idaho National Engineering and Environmental (INEEL) site, the airframe mechanics class located at Idaho State University (ISU), and the Technical Support Building (TSB) located in Idaho Falls. There were 24 participants from the TRA area, 17 from ISU, and 32 from TSB.

The TRA personnel represented a mixed group of craft personnel, reactor operator personnel, engineers, and office workers. The craft personnel were made up of electricians, mechanics, machinists, and laborers.

The ISU personnel were all students in a two-year airframe maintenance program. Most of the students were in the <20 and 21-30 age group. The class had experience in aircraft engines and metal airframes/skin. The class had not been exposed to carbon fiber surfaces.

The TSB personnel represented a mixed group of engineers and office workers. This group had the largest representation of females. Figure 3 shows the ages of the participants.

To provide an interesting comparison, we gathered data from one other group. This group was composed of 6th grade science students from Sandcreek Middle School in Idaho Falls. These students examined one inspection specimen only, (limited by class time), and their observations have been included at the end of this paper.

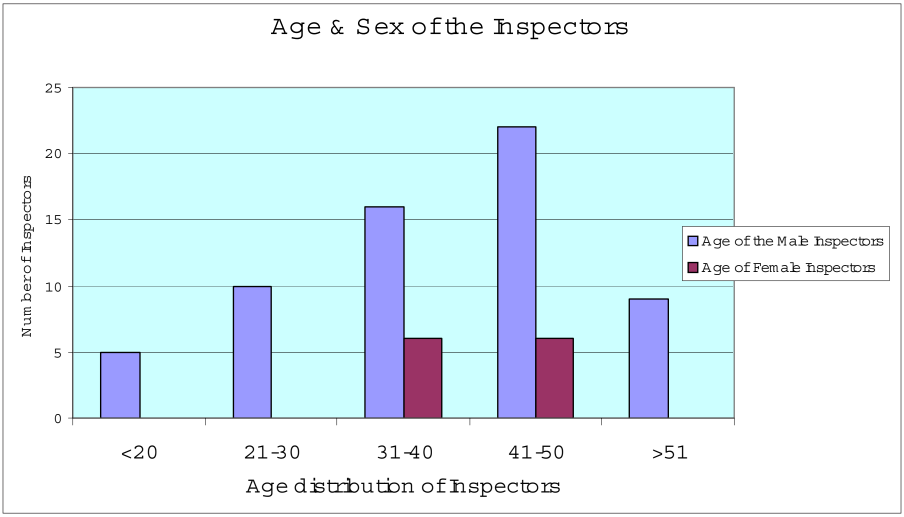
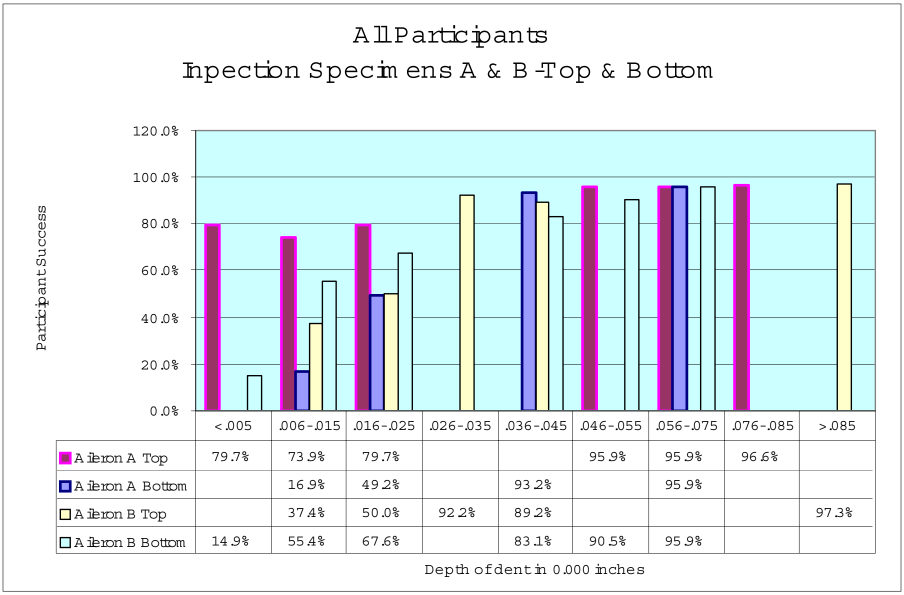


Figure 3. Age of the participants.

Results

The overall results of the data gathering are shown in Figure 4. This is the summation of all the adults surveyed, (regardless of sex or lighting conditions), at the three sites. In almost all cases involving the smaller dents, the same size dent is missed when it is located on the underside of the inspection specimen. In the case of the 0.05 dents, 95.9 % found it on the upper surface and 90.5% found it on the lower surface.



	<.005	.006-.015	.016-.025	.026-.035	.036-.045	.046-.055	.056-.075	.076-.085	>.085
Aileron A Top	79.7%	73.9%	79.7%			95.9%	95.9%	96.6%	
Aileron A Bottom		16.9%	49.2%		93.2%		95.9%		
Aileron B Top		37.4%	50.0%	92.2%	89.2%				97.3%
Aileron B Bottom	14.9%	55.4%	67.6%		83.1%	90.5%	95.9%		

Figure 4. Summary Graph, all participants, Inspection specimens A & B top and bottom.

When a comparison is made between men and women, it yielded inconclusive results. On some of the dents, the women were more observant, and on others, the men were. There was no clear differentiation in the attention to detail between men or women.

Age yielded more conclusive results. For locating dents of 0.05, the most accurate group was the 31 to 40 year olds when the top of the inspection specimen was observed. However, when the bottom of the inspection specimen was examined, the most accurate group was the 21 to 30 year olds.

When we compared the data gathered from the three locations, we were very surprised at the results. As a whole, the ISU students (airframe mechanics) found the fewest dents. The conditions in the ISU classroom were no different from the other display locations and the students were encouraged by their instructor to locate the dents.

False Positives

The number of false positive dents was determined as well. A false positive was defined as a falsely identified dent. This value was calculated by dividing the total number of falsely identified dents by the number of actual dents on the test specimen. We felt this was the best definition of false positives for this study.These data are summarized in the Table 2.

Table 2
False Positives

Inspection specimen	Mean	Median	Standard Deviation
A - Top	23.5%	18.2%	26.5
A - Bottom	29.0%	22.2%	28.9
B - Top	16.5%	9.1%	18.7
B - Bottom	31.0%	25.0%	26.7
Overall	25.5%	18.2%	25.0

The number of false positives for the tops of the inspection specimen was less than for the bottom surfaces. This was not statistically significant for Part A, but was for Part B. In all cases, the median value was less than the mean indicating that most subjects noted fewer false positive dents, but a few subjects found a great number of false positive dents.

Discussion

We noted that the participants each had their own personal way of examining the inspection specimens. Some people touched the surfaces more than others did. When looking at the underside, some bent over, some knelt, and some pulled up a chair. Others asked to have the floodlight moved. Some of the participants asked for a flashlight when examining the undersurfaces and others commented that the underside was much harder to examine. It can be noted that for locating a defect of 0.05 inch on the bottom side, the most accurate age group was 21–30. For the topside, the most accurate age group was 31–40. This comparison poses the question of whether or not age verses dexterity was a factor in the detection of defects between the top and bottom of the surfaces.

The inspection specimen display attempted to simulate bright sunlight and a clean and dirty surface. In the tabulation of the data, no clear conclusion can be drawn in the relationship to these conditions. It did not appear to make any difference if the surface was clean or dirty. Some of the participants commented that the bright light made it easier to see the dents, while others thought the bright light interfered with their observations.

No restrictions were placed on the amount of time that the participants could have while examining the inspection specimens. Although not recorded, we noticed that the average amount of time spent in examining the four surfaces took about 15-20 minutes

This study had some shortcomings. The most significant of these was is the inability to consistently make dents of exactly 0.05 inch on the different surfaces. We had to adjust the height from where the weight was dropped to attempt to create a 0.05 inch depth. This problem created some gaps in "depths of dents" and consequently the data. These gaps in data are seen in both Table 1 and in all Figures in this report.

Summary

A summary of the data collected in this research project showed that an average of 93.2% of untrained adults were able to find dents of 0.05 inch. This confirmed that there is potentially a problem with people being able to see a small dent, and that much more detailed and more specific studies need to be conducted. We have made recommendations for some future research in relationship to our results.

.

Future Work

Many variables were not examined during this initial study on visual detection of dents in composites. We are obtaining a much larger section of an inspection specimen that will be used in future studies. We will also be determining better methods for inflicting damage in the composite materials with more consistency.

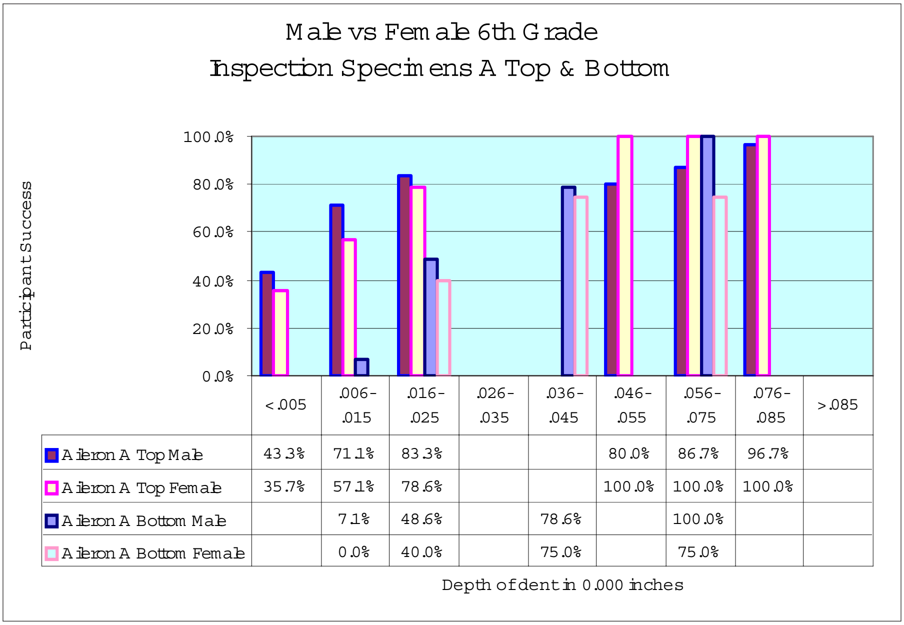
The variable of most interest to the aircraft industry are:

- Clean vs. dirty surfaces
- Dull vs. glossy finishes
- Lighting
- Viewing distance
- Visual angle
- Correlation of identified interior damage with a surface defect
- Experience and training
- Combination of visual and tactile inspections

We will begin this work in the fall of 2004 and hope other researchers consider performing similar studies.

Fun Data: The Sixth Graders

The sixth grade students at the Sandcreek Middle School in Idaho Falls were very excited to be part of a NASA research project. Because of the class sizes and limited length of a class period, the students were asked to only examine one surface. No attempt was made to distribute equally the boys/girls on the different surfaces. The results of the students' examinations are shown in Figure 5 below.



	<.005	.006-.015	.016-.025	.026-.035	.036-.045	.046-.055	.056-.075	.076-.085	>.085
■ Aileron A Top Male	43.3%	71.1%	83.3%			80.0%	86.7%	96.7%	
□ Aileron A Top Female	35.7%	57.1%	78.6%			100.0%	100.0%	100.0%	
■ Aileron A Bottom Male		7.1%	48.6%		78.6%		100.0%		
□ Aileron A Bottom Female		0.0%	40.0%		75.0%		75.0%		

Figure 5. Sixth Grade Students

It is interesting to note that the sixth graders did a much better job of locating the defects on the bottom surface. This may be because they were very excited to participate or simply that they had an easier time looking at the under side being smaller in stature.

Acknowledgements

We wish to acknowledge Dr. Barbara Kanki of NASA Ames Research Center for sponsoring this work. We also wish to thank Jordan Argyle for performing a data review for this work.

References

Goranson, U. F. & Rogers, J.T. (1983, May). Elements of damage tolerance verification, *12th Symposium of International Commercial Aeronautical Fatigue*, Toulouse, France.

Hsu, D. K., Barnard, D. J., Peters, J. J., Dayal, V., & Kommareddy, V. (2002, September). Nondestructive testing of composites and their repair, *6th Joint FAA/DoD/NASA Aging Aircraft Conference*. San Francisco, CA.

Ostrom, L., Wilhelmsen, C., Valenti, L., & Kanki, B. (2002). Determining airline inspectors' probability of crack detection, *International Journal of Applied Aviation Studies*, *2* (2), 71-86.

Poon, C., Benak, T., & Gould, R. (1990, May). Assessment of impact damage in toughened resin composites, *Theoretical and Applied Fracture Mechanics*, *13* (2), 81-97.

Spencer, F. W. (1996). *Visual inspection research project report on benchmark inspections*, (Technical Report DOT/FAA/AR-96/65). U.S. Department of Transportation, Federal Aviation Administration, Washington, D.C.

International Journal of Applied Aviation Studies, 4, Number 2

Coherence and Correspondence Decision Making in Aviation: A Study of Pilot Incident Reports

Catherine N. Jacobson

U.S. Army Research Laboratory
Fort Knox, KY

and

Kathleen L. Mosier

San Francisco State University
San Francisco, CA

Abstract

Technological advances in aircraft automation have revolutionized the operating environment of pilots (Billings, 1996). As more tasks originally delegated to the pilot are governed by automated systems, cognitive activity changes considerably, such that it moves away from "perception and response" toward "thinking and deciding" (Mosier, 2002). In light of this change in cognitive activity, research on judgment and decision-making has become increasingly relevant in its application to aviation. Research in the field of judgment and decision-making, however, has long been divided into two ostensibly opposing theoretical perspectives. On one side, the work of researchers is guided by coherence theories, which focus on the rationality of human judgment. On the other side, research is guided by correspondence theories, in which the value of judgments is based on the degree to which they reflect empirical accuracy. In this study, we examined pilot decision making in light of work by Hammond (1996a, 1996b, 2000), who proposed that judgment and decision-making research can, and should, incorporate both of these contrasting theories. Using data from voluntary incident reports, we examined pilots' use of coherence- and correspondence-based decision-making processes as a function of variables such as phase of flight, weather conditions, type of event, and level of aircraft automation. We hypothesized that, in

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

accordance with the tenets of each theory, characteristics of the various conditions would systematically elicit either coherence- or correspondence-based decision. The findings suggested that pilots' use of decision strategy is dependent on the situational conditions and constraints present. Data suggested that weather conditions and the type of event, but not level of automation, affected the frequency of pilots employing either coherence- or correspondence-based decision strategy. Phase of flight in combination with weather conditions impacted the use of a particular decision strategy; in particular, during cruise and approach, pilot decision strategies varied as a function of weather conditions.

Introduction

Over the last 25 years, the technological advances in aircraft automation have revolutionized the operating environment of pilots (Billings, 1996). As more and more tasks originally delegated to the pilot are put into the control of complex, automated systems, the work of the pilot and the accompanying cognitive processes change considerably (Mosier, 2002). In particular, the presence of increasingly sophisticated equipment designed to fly the plane, monitor its multiple systems, manage flight and navigation data, and a multitude of other operations, has affected the cognitive activity of pilots such that it moves away from "perception and response" toward "thinking and deciding" (Mosier, 2002, p. 93). In light of this change in cognitive activity, research on judgment and decision-making has become increasingly relevant in its application to aviation. Current understanding of the flying task in the context of the automated cockpit, however, has been limited by the fact that the evolution in aviation technology has not been accompanied by a parallel evolution of cognitive theoretical models and frameworks within which to study it.

Setting the Theoretical Context

Research in the field of judgment and decision-making has long been divided into two ostensibly opposing theoretical perspectives within which the goals of decision-makers as well as the decision evaluation criteria are different. On one side, the work of researchers has been guided by coherence theories, which focus on the rationality of human judgment, and the ability to form and maintain a logical and coherent understanding of a given situation; judgments are evaluated solely on their intellectual and analytical merits. Research guided by coherence theories suggested that humans are very poor decision makers whose judgments are prone to frequent errors arising from various situational factors. The work of Tversky and Kahneman (1974; Kahneman, Slovic, & Tversky, 1982), for example, was devoted to unearthing the cognitive illusions and biases present in human decision-making. The concept of coherence-based decisions within the context of aviation is best illustrated with an example. For pilots, the medium in which coherence-based decision strategies are made resides in the information that is provided to pilots through the data displays. Pilots use empirical information to assemble a logical and coherent picture of a situation in which they find themselves and then base their decisions on that picture. Their abilities to combine all relevant information rationally and consistently to form an assessment of a situation will determine the goodness of any decisions made from that assessment. A pilot's decision to maneuver to avoid a potential mid-air traffic conflict based solely on information provided by data from the instrument panel or a pilot's decision to shut down an engine based on an array of data displays that suggest an engine malfunction are both examples of pilots engaging in coherence-based decisions. The situations that may immerse pilots in coherence-based decision-making are not always so straightforward however, as many times conflicting information is presented to pilots through different cockpit displays. The pilot must use all available information to develop a logical and coherent picture of the situation from which decisions are then made.

Research that falls under the other theoretical perspective is guided by correspondence theories in which human judgment and decision making is seen as an adaptive trait which thrives in natural situations. The value of judgments as viewed through correspondence theories is derived from the accuracy of perception, and the resulting decision behavior based on that perception. Signal Detection Theory is one example of a correspondence theory (Green & Swets, 1974) in which decisions are evaluated on the basis of whether or not the presence or absence of a stimulus is correctly perceived. The decision to report its presence or absence is rated according to the decision maker's accuracy in perception of that stimulus. Brunswik's Lens Model, another correspondence theory, measures accuracy in the use of probabilistic cues to judge or predict a criterion (Brunswik, 1943; 1956; see Cooksey, 1996, for a review). Correspondence-based decisions within the realm of aviation involve the use of real time, environmental, probabilistic cues such as references to air traffic, landmarks, or runways as seen outside the cockpit window, fumes that can be smelled from within the airplane cabin, clouds, vibrations and any other physical cues that a pilot perceives. The goodness of a pilot's correspondence-based decision depends on its empirical accuracy and is a function of how the pilot perceives and interprets physical cues, the assessment of a situation based on those cues, and how well that assessment corresponds to the actual state of the world.

In aviation, an effective approach to understanding the modern pilot's task requires a metatheoretical framework that captures all of its elements. Thinking of aviation diagnosis and decision-making processes in terms of coherence and correspondence metatheories would allow identification relevant independent and dependent variables according to the theoretical research context and goals. For example, a correspondence-based examination might focus on clarity or 'goodness' of cues, appropriate cue weighting, or expert recognition of cue patterns, while a coherence approach would focus on consistency, rationality, and comprehensiveness in the use of cues and information. By defining variables and cognitive tactics associated with each approach to judgment, we can examine factors that elicit or subvert particular strategies or processes, and look at the match between a given situation and effective cognitive processes.

Why Do We Need Both Perspectives?

Arguably, the 'best' decision-making processes would be those that are both coherent (rationally sound) and correspondent (objectively accurate). Aviation is a *correspondence-driven work domain* in that it exists within an objective reality that imposes real constraints on interactions, and that must be perceived correctly (Vicente, 1990). Clearly, in this environment, pilots must make accurate decisions based on probabilistic cues, such as altering a flight plan based on weather predictions or performing a go-around in potentially hazardous landing conditions, and strive for precision in their assessment of traffic threats or in the accomplishment of a visual approach. Naturalistic decision-making frameworks, such as the Recognition-Primed Decision Making model (e.g., Klein, 1993; 2000), which focus on experts' achievement of accurate decisions in dynamic environments such as aviation, embody the correspondence perspective of judgment. Much of the correspondence process, as defined by these and other frameworks, is rooted in non-analytical, intuitive cognition. Several researchers have documented usage of these processes in aviation under conditions in which cues are uncertain or ambiguous predictors of reality (e.g., Klein, 1993; Fischer, Orasanu, & Montalvo, 1993; Mosier, 1991; Orasanu & Fischer, 1997).

Examining decisions solely in terms of accurate use of probabilistic cues, however, does not capture the judgment process in the automated environment. In high-technology electronic cockpits, automated systems bring many probabilistic cues from the outside environment into the cockpit, and display them as highly reliable and accurate data and information. Pilots in this environment must monitor these systems and cockpit displays, make judgments about whether or not their indications are appropriate and consistent with the flight context, and, in the event of discrepancies or anomalies, decide what should be done to bring elements back into sync with each other.Advances in aircraft automation have greatly increased the need for this type of cognitive activity and decision-making behavior, which exemplifies the coherence perspective of judgment.

Neither the correspondence nor the coherence perspective, then, is sufficient to describe decision-making processes in the electronic cockpit. Within this domain, pilots must exercise competence in both coherence and correspondence. In many situations, they may alternate between strategies (Hammond, 2000). While landing an aircraft in good visibility conditions, for example, a pilot may switch back and forth rapidly from correspondence to coherence – checking between cues outside of the window to estimate the aircraft's position, and inside at cockpit instruments to ensure that glide slope, localizer, altitude, and speed indications are as they should be. In some cases, one crewmember may be responsible for coherence and the other for correspondence. A standard landing routine may call for one pilot to keep his or her head "out the window" while the other monitors instruments to makes sure that glide slope capture takes place and altitude readings are appropriate. In other situations, the two processes may represent sequential stages of the judgment process. Hammond (1996b) suggested that correspondence processes may represent the inference phase and coherence processes the justification or rationalization phase of diagnostic decision-making. So, for example, the pilot may see or smell smoke in the aircraft, infer that there is an engine fire, and verify that diagnosis by checking the engine instruments and electronic displays that would be present during an engine fire.

It is quite possible that the use of either coherence- or correspondence-based approach to decision-making will hinge, to a great extent, on the task being accomplished and the circumstances under which diagnoses and decisions are made. An essential first step in the integration of these theoretical perspectives, then, would be to determine what factors are associated with each type of diagnostic and decision-making strategy. In this study, we explored the relationships between certain variables and situations inherent in the operational environment and the accompanying approaches to decision making by pilots. Specifically, we examined how weather conditions (outside visibility, VMC; no outside visibility, IMC), phase of flight (initial climb, cruise, approach), sophistication of aircraft automation (no automation, highly automated), and type of incident event (equipment malfunction, traffic incursion) affect the pilot's use of a coherence-based or correspondence-based approach to decision making. The pilot incident reports were collected using the above parameters and a coding system was developed to ascertain the particular decision strategies used by pilots during these incidents. While many of the narratives involved decisions that were purely coherence- or correspondence-driven, several other reported incidents involved both decision strategies. To understand fully how the context in which decisions are made affects decision strategies, it is useful to study the spectrum of contexts ranging from purely coherence- to purely correspondence- inducing and all of those in between. While the parameters used in this study were selected in order to investigate the impact that certain variables had on decision strategies, a useful byproduct was the selection of contexts that represent several points along such a spectrum.

Voluntary pilot incident reports submitted to the Aviation Safety Reporting System (ASRS) were gathered and reviewed systematically on the premise that pilot decision-making can be characterized according to the stipulations of either coherence or correspondence approach by looking for decision behavior that exemplifies a particular approach. The use of archival data such as pilot incident reports represents the necessary "causal texture of the environment" (Hammond, 1993, p. 210), as reported by the pilots. ASRS reports contain valuable information that has been reported by pilots in real life in-flight situations, from their perspectives. In many cases, such information is not even accessible in high fidelity simulations (Degani, Chappell & Hayes, 1991). The ability to draw data from the recollection of pilots during actual incidents is an important strength when looking at approaches to decision-making, since the context in which a decision is made may be vital to the use of either coherence- or correspondence-based approach.

Method

Data were collected from pilot incident reports submitted to the Aviation Safety Reporting System between 1990 and 1998. Because this study was conducted in conjunction with other research involving regional air carrier operations, we searched the Part 135 database only and did not include reports from Part 121 (major carrier) operations. In that there were many (24) combinations of variables under consideration, the data sets for each combination were limited to 10. To obtain data sets of 10 reports for each of the 24 combinations of variables (2 event types x 2 weather conditions x 2 levels of automation x 3 phases of flight), a total of 240 reports were collected using search parameters that, combined, brought up only those that contained the specified parameters. Queried reports were screened in the order of most recent to the oldest report until 10 viable reports were selected for each variable combination. For example, a query for reports that contained a combination of the variables equipment malfunction, and IMC weather conditions was created. The query generated a list of 377 reports that matched those search parameters. Of those 377 reports, the 10 most recent reports for each combination of parameters were examined for inclusion in the data set. Exclusion criteria were:

(a) no indication of information or cues used by pilot(s),
(b) the incident was reported by someone other than a member of the flight crew, such as a flight attendant or an air traffic controller,
(c) insufficient narrative to describe adequately the incident,
(d) the aircraft was a helicopter or non-engine plane, and
(e) the narrative did not actually report a traffic or equipment problem.

This process continued until we had the 10 most recent usable reports for each combination.

Once all reports were collected, they were coded for decision strategy. The reports to be read and coded contained only the accession number and the pilot narrative; information regarding the variable combinations was listed in a separate database, and linked to the "narrative only" reports by the unique accession numbers. Due to the nature of the coding system, the variability of how behaviors and information were presented in the pilot narratives, and the technical nature of the subject matter, all 240 reports were coded two times. All 240 narratives were first read by the primary coder. The narratives were then divided into two groups of 120 reports and each group was assigned to a secondary coder. The secondary coders worked independently of each other, and discrepancies that arose between the primary and secondary coders were reconciled. A preliminary check revealed an inter-rater reliability of 75% when coders worked independently from each other and the primary coder. The relatively low initial inter-rater reliability was largely a result of differences in interpreting narrative details that describe the decision strategy used by the pilots and, to a lesser extent, misinterpretations of the narratives due to unfamiliarity with acronyms used within the text. This agreement was increased to 100% when all reports were reconciled through consensus with the primary coder. Consensus was reached by breaking down each narrative to point out what details within the text support a particular decision type and to define unfamiliar acronyms.

Coding System

A coding system was developed in order to reliably determine from reading the narratives whether or not pilots engaged in coherence-based, correspondence-based, or both types of decision-making strategies. A compilation of 20 yes/no questions was created, the answers to which could determine the strategy used. The coding system was based on the characterizations of coherence and correspondence in decision-making, and assumptions of what behaviors each type of decision strategy would elicit. The underlying questions used to develop the yes/no items in the questionnaire were, "What does coherence-based decision-making look like during the task of flying an aircraft? What does correspondence-based decision-making look like?"

The first 18 items of the coding sheet were designed to identify the pilot behaviors that could determine which decision strategy was used. For example, item number one stated, "The pilot uses information only," with examples such as gauges and engine parameters, publications, and visual or aural warnings/indicators. A *yes* answer to this question indicated that the pilot engaged in coherence-based decision-making. Another item, "The pilot uses one or more direct sensory cues only," included examples such as outside views (looks outside the cockpit window), traffic, terrain, runway conditions/features, odors and/or sounds *other* than automated warning sounds (e.g., "thud," "bang"), or anything kinesthetic (e.g., "felt" the plane yaw, roll, pitch, etc). Other items clarified strategy by coding whether the pilot mentioned the fallibility of any of the cues used, where the initial indication of a problem came from, whether the pilot crosschecked other sources during diagnosis, etc.

As mentioned earlier, one of the fundamental distinctions between coherence and correspondence-based decision-making is the overall goal of the decision maker. The final two questions of the coding sheet were designed to query outright which decision strategy was used based on the overall objective of the pilot (i.e., to create a logical picture of the current situation, or to verify the empirical accuracy of information and cues). Criteria for answering the last two questions were based on the overall focus of the pilot's attention (gathering information, or cues), and what he or she seemed to consider most when resolving (or attempting to resolve) the incident. Some incidents seemed to involve both coherence- and correspondence-based decision strategies, where both strategies played an equally important role in the resolution (or attempt at resolution) of the incident. A "yes" answer to the last two questions indicated in this study that both decision strategies were used. Few incidents were categorized as both, but it is noteworthy to mention that identifying the use of coherence or correspondence was not always straightforward and, in fact, a combination of both strategies may have been employed for a single decision. It may also be likely that in a given incident, several decisions were made in order to resolve, or attempt to resolve a situation, and each decision required a different approach. For example, in one incident a co-pilot used both strategies in a decision to alert the captain of a need to deviate from a current flight plan. The co-pilot used both physical cues (visible icing on the leading edge of an aircraft wing) as well as data (an assimilation of information from instrument readouts signifying control surface problems) to arrive at a decision to alert the captain of a need to alter the flight plan to accommodate the possible aircraft malfunction.

Results

This study employed the inferential statistic X^2 (Chi-square), an appropriate statistic for analyzing qualitative data such as pilot incident reports. Chi-square is an inferential statistical test that is based on nominal data (e.g., data that describe categories of events). A Chi-square test determines whether or not the frequency of an event is dependent on the context in which it occurs. The null hypothesis of a Chi-square test is that the frequency of an event (the dependent variable) is not dependent on the context in which it occurs (independent variable); rather, the frequency is a result of chance alone. A rejection of the null hypothesis indicates that the frequency of an event reveals a systematic pattern and that the frequency of an event is dependent on the context in which it occurs. The proportion of frequencies of coherence- and correspondence-based decisions by pilots was coupled with the variables outlined above: (a) weather conditions, (b) type of incident event, (c) level of automation, and (d) phase of flight. Analysis of these frequencies revealed whether or not the use of coherence- or correspondence-based decisions was dependent on these variables.

Weather Conditions

The findings indicated that, overall, the use of coherence or correspondence does depend on the weather conditions X^2 (2, N = 240) = 37.80, p < .001. For all incidents that occurred during IMC, pilots made coherence-based decisions (N = 90) more often than they made correspondence-based decisions (N = 18), and although the frequencies of decision strategies were quite similar during VMC, pilots still made correspondence decisions (N = 61) more often than they made coherence decisions (N = 46). These findings support the notion that in an effort to conduct operations from an accurate perception of reality, pilots will likely attempt to make correspondence-based decisions when outside visibility permits. Conversely, when outside visibility does not exist, pilots will forgo attempts to make decisions that correspond to objective reality and instead, develop a coherent picture of the information supplied to them through cockpit displays, and then base their decisions on the coherence of that information.

Event Type

As proposed in this study, when pilots encounter an aircraft equipment problem or malfunction, it seems likely that they would employ coherence in their judgment making because their goal would be to detect and resolve the problem within the closed-circuit environment of the aircraft's systems and subsystems by building a logical and coherent picture of the available information regarding the aircraft's operating systems. When pilots encounter incidents that involve a problem with surrounding traffic, however, it is suggested here that they would attempt to establish correspondence in their decisions if outside visibility permits, and would attempt to establish coherence if outside visibility does not permit. The findings suggested that the use of coherence or correspondence does depend on the type of event X^2 (2, N= 240) = 35.33, p < .001. For all incidents involving equipment problems, pilots made coherence-based decisions (N = 88) much more often than they made correspondence-based decisions (N=18), and during traffic problems, pilots made correspondence decisions (N =61) more often than they made coherence decisions (N = 48). These findings suggested that the type of activity in which pilots are engaged does impact their decision-making strategies, at least when the events either involve equipment malfunctions or traffic conflicts.

Phase of Flight

The likelihood that a pilot's decisions will be based on either correspondence or coherence during a particular phase of flight may be a function of the types of incidents that occur most frequently at any one phase, as well as the visibility conditions at the time of the incident. Incidents involving traffic may be more likely to occur during times when planes are flying in the traffic pattern, yet incidents involving aircraft equipment malfunctions may happen at any time. Overall, without considering incident type, phase of flight bore no significance in terms of whether or not coherence or correspondence was used more often X^2 (4, N = 240) = 5.31 p = .26. Across all phases, coherence was used more often, yet it was used relatively less often during initial climb than during cruise and approach. The difference in frequencies of each decision strategy was smallest during initial climb.

Phase of flight, in combination with traffic problems and weather conditions, did produce some interesting findings. During traffic problems while in cruise the use of coherence or correspondence depended on weather conditions X^2 (2, N = 40) = 21.50, p < .001. This also was true for approach X^2 (2, N = 40) = 19.89, p < .001. This was *not* true, however, for initial climb X^2 (2, N = 40) = 5.46, p = .065. A possible explanation for this would be the fact that, during initial climb, the airplane was beneath the cloud deck such that the weather conditions (outside visibility) were VMC regardless of the reported weather conditions. While the differences in strategy during initial climb traffic situations were not significant, the pattern of use of coherence and correspondence was similar to the other comparisons of weather conditions - the use of coherence was more frequent during IMC conditions while the use of correspondence was more frequent in VMC conditions.

Level of Automation

It was suggested in this study that the level of automation would affect pilots' approach to decision making. A cockpit environment with little or no automation would draw pilots away from the inside, closed-circuit environment of aircraft systems and direct them toward the outside, real-time environment where it is more likely that decisions would be based on correspondence with the outside environment. In contrast, pilots flying a heavily automated aircraft are working in an information-rich environment in which the status of the airplane's systems and flight performance are presented to the pilot in the form of information displays. In this heavily automated environment, the relative frequency of coherence-based decisions was hypothesized to be high because the tasks of the pilot would require maintaining a coherent picture of the information available inside the cockpit. The level of aircraft automation, overall or in combination with any of the other three variables, however, did not significantly affect the use of coherence or correspondence-based decisions.

Discussion

While the evolution of increased automation in aviation has greatly enhanced airline safety, flexibility, and efficiency, the impact that it has on the human operator is extreme. As the role of the pilot gradually moves from active controller to passive monitor, the accompanying decision processes must evolve accordingly. In light of this progression, the need for research in judgment and decision making in aviation is transcended only by the need to build a solid and unified theoretical framework upon which such research can be conducted. This study has applied the coherence/correspondence metatheorical approach advocated by Hammond to the aviation domain. The findings suggested that in aviation operations the use of either coherence- or correspondence-based approach to decision-making does hinge, to some extent, on the circumstances under which decisions are made. In particular, weather/visibility conditions and type of event, as well as phase of flight in combination with other variables will systematically and predictably influence the type of decision strategy employed by pilots.

The findings from this study warrant further investigation on the decision processes of pilots within a theoretical framework that combines both coherence and correspondence theories. It is clear that research that attempts to define, describe, or evaluate pilots' decision-making with respect to only one decision theory is likely to be misleading and unrepresentative of how and why pilots actually make decisions. A more structured approach to understanding how pilots make decisions and a refinement of what behaviors constitute the use of one decision strategy versus another would be a valuable extension from this study.The methodology and analysis used here was limited by the nature of the data; pilot narratives provide rich qualitative data, yet cannot offer complete and trustworthy information on every aspect involved in pilots' decisions (Chappell, 1994). Careful manipulation of the circumstances under which pilots are observed as well as a thorough documentation of pilot actions may supplement with more reliability the findings of this study.

The failure to find differences in decision strategies as a function of level of automation was surprising and may have been an artifact of the report selection process. The search criteria for level of automation used in this study were such that a low level of automation included only those reports where level of automation was "None," and a high level of automation included only those reports categorized as "Automated/Integrated Navigation and Flight Control Systems." These categories were used with the intention of creating a substantial difference between reports involving aircraft low and high levels of automation and we did not use reports from other categories of automation. It may be that if this study had included incidents from all levels of automation then significant findings with regard to aircraft automation may have been found. Additionally, the ASRS CD-Rom used for the study did not include reports after 1998, which limited the number of reports involving aircraft with sophisticated flight management computer systems in our sample population.

A closer look at aircraft automation and the use of coherence and correspondence, either by comparing incidents categorized in the ASRS reporting system or through a laboratory study, may reveal significant findings. It is important to consider that even the lowest level of aircraft automation has displays and that some decision strategies may be used across all levels of automation. In such a case, decision strategy would not vary across the levels of automation but would, instead, reside within all conditions of aircraft automation sophistication, thus revealing no significant differences. Furthermore, it is possible that decision strategies are dependent on factors other than the level of automation, factors that have not been explored in this study. It is likely that a more rigorous approach to learning the effects of cockpit automation on the use of coherence- or correspondence-based decisions will yield significant results.

The efforts taken in this study provided an important first step in understanding the decision processes within a unifying framework of two historically contradictory decision theories. While there are notable limitations to this study, the patterns of use for coherence- and correspondence-based approaches to decision-making under varying situations are evident, and further research efforts are encouraged.

Acknowledgments

The authors would like to thank Dr. Kenneth Hammond for providing direction and perspective in the development of this study. Without Dr. Hammond's extensive contributions in the field of judgment and decision-making, the foundations upon which this study relies would not exist. Thanks also to Nikita Sethi and Laurel Bailey for their much-valued contributions, and to anonymous reviewers for their comments.

References

Billings, C. E. (1996). *Human-centered aviation automation: Principles and guidelines*. (NASA Technical Memorandum No. 110381). NASA Ames Research Center, Moffett Field, CA.

Brunswik, E. (1943). Organismic achievement and environmental probability. *Psychological Review, 50*, 255-272.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.

Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications*. San Diego, CA: Academic Press.

Chappell, S. L. (1994). Using voluntary incident reports for human factors evaluations. In N. Johnston, N. McDonald & R. Fuller, (Eds.), *Aviation Psychology in Practice*. Aldershot, England: Avebury Technical.

Degani, A., Chappell, S. L., & Hayes, M. S. (1991, May). *Who or what saved the day? A comparison of traditional and glass cockpits*. Paper presented at the 6th International Symposium on Aviation Psychology, Columbus, Ohio.

Green, D. M. & Swets, J. A. (1974). *Signal Detection Theory and Psychophysics*. New York: Krieger, Inc.

Hammond, K. R. (1993). Naturalistic decision making from a Brunswikian viewpoint: Its past, present, and future. In G.A. Klein, J. Orasanu, R. Calderwood & C.E. Zsambok (Eds.), *Decision making in action: Models and methods.* Norwood, NJ: Ablex Publishing Corp.

Hammond, K. R. (1996a). *Human Judgment and Social Policy*. New York: Oxford Press, Inc.

Hammond, K. R. (1996b). How convergence of research paradigms can improve research on diagnostic judgment. *Medical Decision Making, 16(3)*, 281-287.

Hammond, K. R. (2000). *Judgments under stress*. New York: Oxford Press, Inc.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press.

Mosier, K. L. (2002). Automation and cognition: maintaining coherence in the electronic cockpit. In E. Salas (Ed.), *Advances in Human Performance and Cognitive Engineering Research, Volume 2*. Elsevier Science Ltd. (pp. 93-121).

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.

International Journal of Applied Aviation Studies, 4, Number 2

Importance of International Oversight in Aviation Safety and Security: The Need for Political Commitment and Regional Cooperation

Vahid Motevalli[1], Farooq Ansari, Alfredo Lagos, and George Novak

School of Engineering and Applied Science
The George Washington University
Ashburn, Virginia, 20147

Abstract

The air transport industry is currently facing many challenges due to worldwide economic pressures and increasing costs of security and safety. It also has tremendous growth opportunities because of increased travel, growth in global manufacturing, and product development, as well as an increasing dependence of national economies on air transport as a facilitator for economic activity. Based on International Civil Aviation Organization (ICAO) audits, compliance with ICAO requirements by many civil aviation authorities around the world is not adequate and often not sustained. Lack of political commitment and needed resources are often deficient in sustaining the minimum regulatory activities and oversight needed to match the pace of growth in aviation activities.

This paper discusses the role of oversight in the aviation industry worldwide by examining the critical elements of an effective civil

---

[1] This research was supported in part by a grant from the Federal Aviation Administration Flight Standards Service, which in part supported this work, under a Cooperative Agreement with the George Washington University, along with assistance from the Office of International Aviation and support from the Contracting Officer Technical Representative Mr. William Nash is greatly appreciated.

Correspondence concerning this article should be addressed to Vahid Motevalli, School of Engineering and Applied Science, The George Washington University,Ashburn, Virginia, 20147. E-mail: vahidm@gwu.edu.

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

aviation oversight authority. Analysis of data on civil aviation authority oversight organizations showed that it is difficult to identify clearly quantifiable metrics for measuring the strength or degree of compliance of a civil aviation authority. Furthermore, the resource limitations of many civil aviation authorities make it clear that regional organizations are the best approach to strengthen oversight capabilities.

Introduction

The air transport industry has played an increasingly important role during the last quarter of the 21st Century as a facilitator of overall economic activity and as a critical element in certain economic sectors. Air transport has become a necessity to ensure the efficient and cost-effective movement of goods and services. However, many civil aviation authorities that are mandated by national governments to ensure safety and security are not able to sustain effective regulatory activities needed to match the pace of growth. This paper discusses the role of oversight in the aviation industry worldwide by examining the critical elements of an effective civil aviation oversight authority.These critical elements have been used as the basis for an evaluation program by aviation safety and security authorities in the United States to categorize Civil Aviation Authorities (CAA) whose airlines fly to the United States. According to the United Nation's technical agency for aviation, the International Civil Aviation Organization (ICAO), each Member State is responsible for ensuring that its airlines are in compliance with the country's safety regulations.

The George Washington University Consortium for Aviation Safety and Security, through a grant from the Federal Aviation Administration (FAA), has developed a program to advocate the improvement of oversight by CAAs around the world. This program, developed and conducted over a three-year period, addresses key challenges and issues facing oversight officials from around the world. The George Washington University Consortium has hosted a series of on-going meetings among high-level foreign government officials with a focus on resources and structure to achieve full compliance with the Standards and Recommended Practices (SARPs) adopted by ICAO.

International Aviation Oversight

Background

The number of worldwide airline passengers in 2003 has reached an all time high of 1.8 billion, an increase of 400 million over a nine-year period from 1994 to 2003 (see Figure 1 - ICAO, 2003). While a dip occurred in 2001 and 2002 due to a world economic downturn, followed by the September 11, terrorist attacks and the Severe Acute Respiratory Syndrome (SARS) outbreak, growth continued in 2003. Not withstanding the importance of passengers carried, air transport has become a necessity to ensure the efficient and cost-effective movement of goods and services. Today, air transport has become crucial in maintaining contemporary

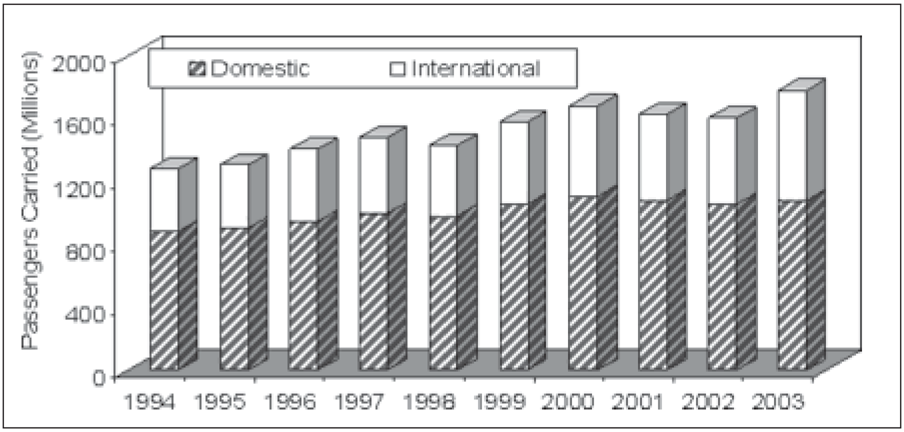


Figure 1. Increasing Number of Passengers Carried by Air Transport (*Note*. From "Air Traffic Remains Flat, Improvements Expected in 2004," by ICAO, 2003).

just-in-time business processes such as fresh food and flower transport. Thegrowth in the number of passengers and cargo being carried annually by commercial air transport, the ever decreasing tolerance of the public for aviation accidents (due to improved safety records and media attention to such accidents), and the stated goals to reduce worldwide accident rates (for example the FAA's goal to reduce the fatal accident rate by a factor of five in 10 years) makes international oversight activities more critical than ever.

Table 1 provides select data on the national economy and indicators of the size of governmental authorities mandated to oversee aviation safety.Table 1 shows countries that have been de-identified from eight different regions around the world. Next to each country are their annual gross domestic product (GDP) and four columns with data on their aviation oversight. The size of the national aviation industry is represented by both commercial air transport operators (CATO) and aircraft registered (AR). These figures are then compared with the number of aircraft operations inspectors (AOI) as well as aircraft inspectors (AI) to demonstrate the relative size of the national oversight authority. Country E for example has almost $500 billion GDP, with seven AOIs overseeing almost 80 operators and eight AIs to inspect almost 500 aircraft. Table 1 shows that most countries have a low ratio of inspectors to operators, and one can reasonably state that the lower the ratio, the better. While a quantitative assessment is not possible without specific details about the inspection system for a country and their quality control approach, some qualitative observations are easy to make. For example, general aviation aircraft are included in AR, but even ignoring general aviation the number of operators (CATO) is a key factor in determining the inspector force needed. The inspectors are responsible for overseeing the compliance with safety regulations of both civil and general aviation aircraft, and thus the workload and quality of inspection depends on many factors.

Table 1

Selected Data Relevant to Oversight Activities in Select Countries

Region	Country	GDP ($)	CATO	AOI	AR	AI
Asia & Pacific Region	A	429.0 B	14	17	227	17
Middle East	B	53.0 B	13	4	127	4
East and Southern Africa Region	C	50.6 B	3	4	69	4
Western & Central African Region	D	42.5 B	3	4	43	10
European & North Atlantic Region	E	468.0 B	76	7	469	8
Northern American Region	F	900.0 B	500	267	6631	98
Central American & Caribbean Region	G	53.0 B	12	28	225	26
Southern American Region	H	391.0 B	106	45	4580	90

Note. From "The World Fact Book," by Central Intelligence Agency, 2003 and from "Select Audit Findings," by ICAO, 2003. GDP = gross domestic product, CATO = commercial air transport operators, AOI = aircraft operations inspectors, AR = aircraft registered, AI = aircraft inspectors.

Throughout the past few decades, both actual evidence and economic studies have clearly demonstrated the benefit of having a robust, safe, and secure air transport system (George Washington University Consortium, 2004). It is easy to conclude that there are no direct correlations between the readily available economic data, the oversight data, and the effectiveness of a country's oversight.

The Need for International Aviation Oversight

Effective oversight, while difficult to quantify, is both desirable for economic development and tourism, necessary for international air operations, and mandatory under ICAO treaty obligations for each Member State. Today, there are 190 countries that have agreed to abide by mandatory safety and security regulations as Member States to the Convention on International Civil Aviation (ICAO, 2003). One-hundred of these countries have close to 600 air carriers flying into the United States (FAA, 2003). Due to the high number of air carriers, there has been an ongoing public concern about non-United States airlines' safety and compliance with mandatory safety regulations. According to the Convention, each Member State is responsible for ensuring that its airlines are in compliance with its own safety regulations.

On January 25, 1990, the Colombian airline Avianca Flight AV052 crashed on the north shore of Long Island, New York. This accident brought the public concern for compliance of foreign airlines to the forefront of United States public policy. Ultimately, the AV052 crash brought about significant changes in how the FAA oversees foreign flight operations into the United States. The National Transportation Safety Board (NTSB) (1990) concluded that the probable cause of the crash was the failure of the flight crew to adequately manage the airplane's fuel load and their failure to communicate an emergency fuel situation to air traffic control before fuel exhaustion occurred. The failure to communicate the emergency fuel situation was due in part to the flight crew's lack of knowledge of the Standard English phraseology used for emergency situations. The probable cause of this crash was particularly significant to aviation oversight policy because the NTSB strongly recommended (1990) that the FAA should take immediate action to make certain that foreign flight crews operating within the United States were thoroughly knowledgeable of the flight operating and air traffic control rules and procedures, including standard phraseology, for operating in the United States. The United States Congress pressured the FAA to address the deficiency in the training and qualifications of Flight AV052's flight crew. This resulted in further questions on the responsibility for ensuring compliance with safety rules by foreign airlines flying into the United States. This set in motion the development of the International Aviation Safety Assessment (IASA) program by the FAA. The IASA program was designed to measure the degree of compliance with the ICAO standards and recommended practices by countries that maintain flight operations into the United States.

International Aviation Safety Assessment Program (IASA)

There are nearly 600 foreign air carriers that currently operate into the United States from 100 countries or regional country alliances that have oversight responsibilities for those air carriers. The IASA program assesses whether another country's oversight of its air carriers that operate, or seek to operate, into the United States comply with minimum international standards for aviation safety. In 1992, the FAA began to assess whether each foreign Civil Aviation Authority (CAA) with air carriers operating into the United States met the safety standards for International Commercial Air Transport contained in the Operations of Aircraft Annex 6 to the Convention on International Civil Aviation. To determine whether a CAA complies with ICAO standards regarding its oversight of air carriers under its authority, the FAA analyzes collected information from several sources. These sources include the foreign CAA responsible for providing safety oversight to its air carrier, pertinent records, and officials of the subject foreign air carriers. Based on the IASA assessment of a CAA, the United States Department of Transportation recommends whether or not to permit initiation, continuation, or expansion of air services of all carriers overseen by that CAA.

The FAA originally established three categories to rate the level of compliance by a CAA (2003). In 2003, the FAA revised the rating system to two categories to remove some of the stigma from the program. Focusing only on the CAA themselves, and not the air carriers, the IASA program classifies a country as Category I if it is in full compliance or Category II if it is not in compliance with ICAO Standards (FAA, 2003). The FAA places a country in Category II status if it determines that there is deficiency in one or more of the critical areas of safety oversight as established by ICAO. Areas critical to compliance include: laws and regulations, technical expertise, resources and organization to license and oversee air carrier operations, adequately trained and qualified technical personnel, adequate inspector guidance to ensure enforcement of and compliance with minimum international standards, documentation and records of certification, and adequate continuing oversight and surveillance of air carrier operations. These critical elements are discussed later and depicted in Figure 3. Category II is divided between two groups of countries, as one group includes countries that have carriers with existing operations at the time of the assessment, which are permitted to operate at their current level without an increase or change, and also under heightened FAA surveillance (represented as "Category II"). Countries in Category II with existing operations cannot add new air services or aircraft to their existing operations. Furthermore, if an air carrier reduces its air services or aircraft fleet while the country is in Category II status, it cannot replace those lost services or aircraft until it achieves Category I status. The second group includes countries that have been completely prohibited from operations due to a particularly high degree of non-compliance. This group also includes countries that do not have carriers with existing operations into the United States at the time of the assessment. In both cases the carriers from these countries will not be permitted to commence service while in Category II* status (represented with the addition of an asterisk, thus "Category II*"). The FAA concluded (2003) that, in 1994, two-thirds of the countries with existing operations to the United States were not fully complying with minimum standards and practices. According to the 2003 assessment, the number not fully complying has dropped from two-thirds to one-third (FAA). Figure 2 shows the number of countries categorized by IASA arranged by geographic regions designated by ICAO and modified by the authors. On the x-axis are eight regions and on the y-axis is the number of countries audited per region. Seventy-two of the total 100 countries are in full compliance.

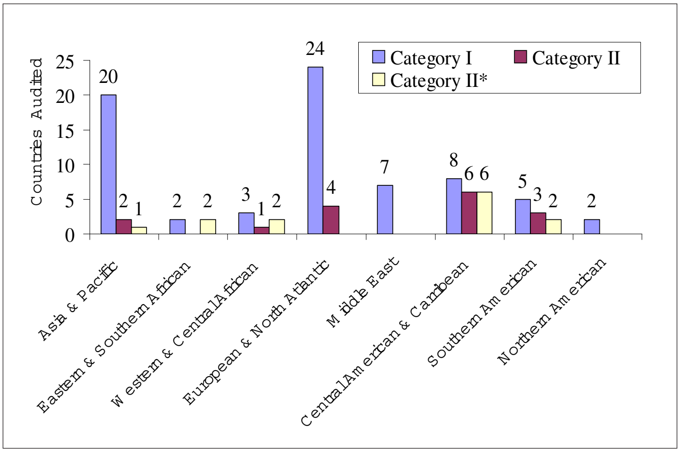


Figure 2. IASA Categorization of Countries Shown According to Region (*Note*. From "Overview of the Federal Administration Flight Standards Service International Aviation Safety Assessment (IASA) Program," by FAA, 2003).

The Standards of the Convention on International Civil Aviation

The Convention on International Civil Aviation was signed in 1944 and resulted in the establishment of the ICAO. The Convention stipulates several conditions, including that all countries maintain sovereignty of the air space over their territory, that signatories have a national responsibility for fulfilling conditions with respect to licensing, certification, and inspection, and that they are responsible for navigation services, accident investigation, and sharing of pertinent information. The Convention also established ICAO as the United Nation's technical agency for civil aviation and its mandate was defined as the safe and orderly growth of international civil aviation. To achieve this end, ICAO approves and adopts

Standards and Recommended Practices (SARPs) that are included in the Annexes to the Convention. Each country that signs the Convention must implement an effective safety oversight system that conforms to the SARPs. Each ICAO Member State must strive for conformity with the SARPs and must notify ICAO if it intends to differ from any standard. Non-conformity must always be noted on aircraft and flight crew licenses and certificates, and non-conforming aircraft and/or flight crews may operate in other countries only with its expressed approval. Notification of non-conformity as required by ICAO can be a valuable asset for Member States in order to be aware of discrepancies that may raise safety or security concerns. Ultimately, individual Member States provide the enforcement authority for compliance.

Member State Responsibilities for Aviation Safety Oversight

The obligations of Member States under the Convention on International Civil Aviation include compliance with all the SARPs under the eight critical elements of aviation safety oversight (ICAO, 2003). Figure 3 describes these critical elements as interdependent, fundamental building blocks of a safety oversight system. When examining these elements noted in Figure 3, it is clear that primary aviation legislation is the basis for effective implementation of the other elements. Primary aviation legislation should provide the basis for operating regulations and a range of enforcement mechanisms, and establish the structure of a CAA. It should be consistent with the environment and complexity of a state's aviation and legal system. The operating regulations should be properly implemented and promulgated in accordance with the laws of the Member State. Once the regulations are implemented, developing a CAA structure comes next, which includes that a director should head the CAA with legal authority over civil aviation. The CAA structure should provide the necessary technical and administrative personnel, and other resources. Furthermore, both the regulations and CAA structure must provide adequate procedures and systems to check and balance the separation between regulatory and operating functions.

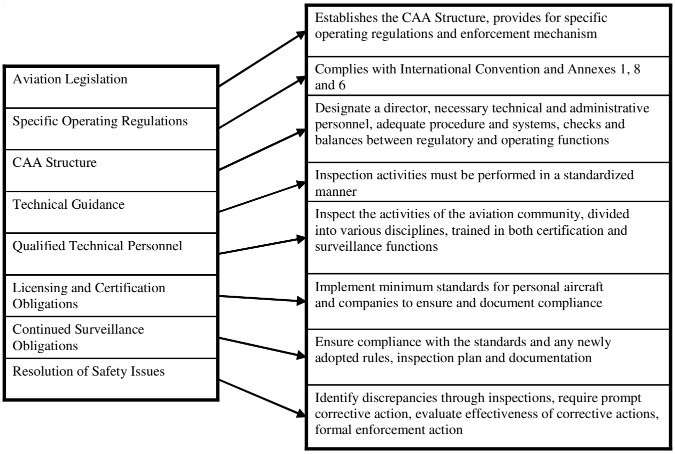


Figure 3. Eight Critical Elements of Safety Oversight (Note. From "Overview of the Critical Elements of Security," by George Washington Consortium [GW Consortium], 2003).

In addition to safety, signatories to the Convention are responsible for implementing and conforming to aviation security standards and recommended practices. The terrorist attacks of September 11, 2001 raised the question of government's responsibility versus that of airlines for the security of the aviation system. The new level and type of terrorism involving suicide attacks and use of airplanes as weapons has forced oversight authorities worldwide to alter their approach to aviation security. Member States have appealed for both revisions and a greater adherence to existing security SARPs. Following the ICAO Ministerial meeting in February of 2002, a number of additional security measures were adopted. Perhaps most importantly, ICAO established the Universal Security Audit Program, which was launched in 2003. This program consists of trained and certified audit teams identifying minor and serious areas of improvement regarding compliance with the ICAO security annex. Audits are conducted to determine compliance and to notify the deficient countries that they must undertake immediate corrective action.

New security requirements approved by a Ministerial meeting in February of 2002 have been organized into four critical elements for aviation security oversight by the George Washington University Consortium (2003). Although these four elements are not identified in ICAO guidance materials, they are clearly integral components of any effective security oversight system. Figure 4 shows the four elements in detail, on the left are the elements and on the right their description.

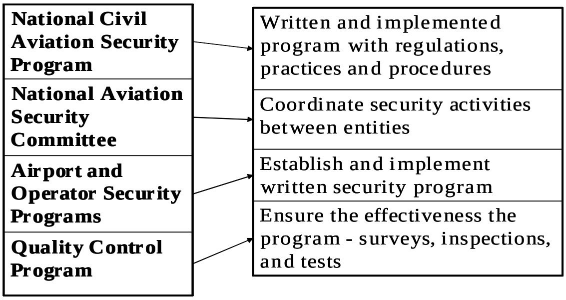


Figure 4. Four Central Elements of Security (Note. From "Critical Elements of Security," by George Washington University Consortium [GW Consortium], 2003).

Prior to establishing any kind of program, a country must first designate and clearly identify an authority to be responsible for aviation security.This not only ensures accountability, but also alleviates unproductive competition between different agencies struggling for partial or total control of the aviation security apparatus. Once an authority is designated, it has the responsibility to establish a National Civil Aviation Security Program (NCASP). The NCASP is a written and implemented program with corresponding regulations, practices, and procedures. The NCASP sets all standards for security oversight while taking into account the safety, regularity, and efficiency of flights. With the security program in place, the National Aviation Security Committee is established for the purpose of coordinating security activities between the departments, agencies, and other organizations of the State; airport and aircraft operators; and other entities concerned with or responsible for the implementation of various aspects of the NCASP.

The third element requests that airports and operators must also establish and implement a written security program. Member States must require each airport serving international civil aviation and the operators providing service from that country, to abide by a security program appropriate to meet the requirements of the NCASP. Security activities for airports and operators include operational procedures such as contingency plans for acts of unlawful interference and the reporting of any such acts to ICAO. Security activities also include the development of emergency plans that act as the bridge between airport security plans and contingency plans. Finally, a quality control program must be implemented to ensure effectiveness of the NCASP.The quality control program should be conducted using surveys to identify security needs, inspections of the implementation of security controls, and the conducting of tests of security controls to assess their effectiveness.

Common Compliance Deficiencies

Deficiencies in both safety and security have been found in CAA's worldwide. ICAO and FAA IASA audits found a pronounced lack of legislation, regulations, and guidance procedures (Kotaite, A., 1998; FAA, 2003). Both, a lack of adequate number of qualified technical personnel and inadequate certification and licensing systems are common deficiencies between ICAO and FAA audit findings. If deficiencies are found by either ICAO or FAA, it indicates problems with the oversight organization, which could lead to adverse consequences. Table 2 describes six consequences of which four relate directly to the lack of personnel.

Table 2

Consequences of Common Oversight Deficiencies

Inability to attract, recruit, and retain qualified technical personnel
Frequent turnover of technical personnel
Job insecurity
Low morale and active searching for other employment
Frequent changes of leadership
Program funding, adequate support staff, computer, record keeping facilities

Note. From "An Update on the ICAO Universal Safety Oversight Audit Program," by ICAO, 2004.

Considering the current state of oversight (see Table 1), the increasing workload for inspectors due to the steady growth of the air transport industry could further exacerbate an already urgent problem. Frequent changes of leadership and lack of program funding, inadequate support staff, and computerized record keeping facilities point to what ICAO has determined as the number one deficiency found in audits worldwide, the lack of an established and adequately funded CAA.

Solutions for improving oversight effectiveness include political commitment at the highest-level, regional cooperation, training, and sharing resources. Training includes the ICAO TRAINAIR program, which has a goal to establish and maintain high standards of training and competency for aviation personnel. In addition, the FAA provides extensive technical assistance including the TRAINAIR program. The Netherlands CAA and the United Kingdom CAA are also very active in assisting other countries worldwide. The key to improving oversight mechanisms in all eight critical elements lies with having political commitment at the highest level of government. To this end, the George Washington University Consortium developed a program for executives and leadership in civil aviation of other countries under a grant from the FAA.

International Summit on Aviation Safety and Security

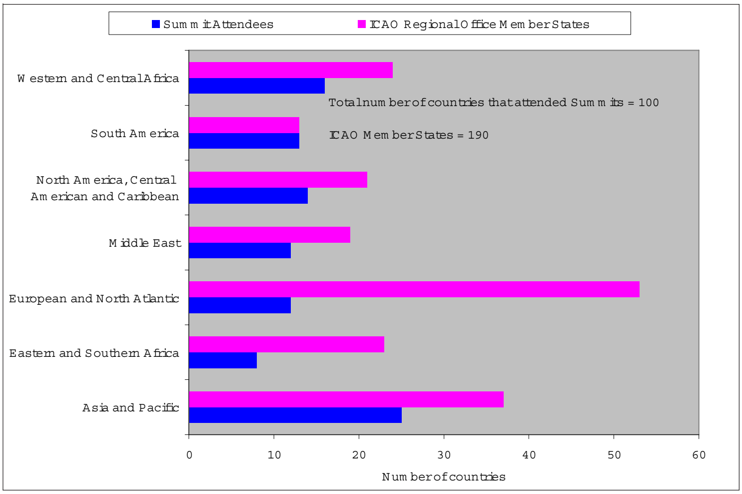
The George Washington University Consortium for Aviation Safety and Security was established in 2001 with a three-year grant from the FAA, by the George Washington University Aviation Institute in partnership with George Mason University Center for Transportation Policy.The George Washington University Consortium (GW Consortium) developed and currently administers an executive level program to address obligations of ICAO Member States in aviation safety and security oversight. The Summit program officially started in September 2001. It is sponsored by the FAA Flight Standards Service, with support from the FAA Office of International Aviation, along with active assistance from the Office of Secretary of Transportation, the Department of State, and the Transportation Security Administration (established in February 2002). ICAO, the U.S. Trade and Development Agency, the World Bank, and the International Development Bank have also been active participants in the program. The International Summit on Aviation Safety and Security emphasizes political commitment and identifies the human and financial resources needed for full compliance with the standards of the International Convention on Civil Aviation. To this end, the Summit program promotes regional cooperation to pool oversight resources among countries. Ministerial-level discussions are an integral part of the training program in order to build consensus concerning the economic benefits of a safe and secure aviation system as well as how to secure available funding and technical resources.

The choice of the word "Summit" is to symbolize the intent of this program to facilitate an open exchange of views and information at the highest levels of government. Using a discussion format, the goals of the Summit program include:

- Raise the level of global aviation safety and security through commitment at the policymaking level
- Strengthen the advocacy by Ministers and Director Generals for aviation safety and security
- Identify the human and financial resources needed to meet ICAO obligations for aviation safety and security
- Exchange views on development of regional cooperation plans, and maintaining conformance with ICAO obligations

The Summits are not intended to provide in-depth technical knowledge; they are designed to offer the GW Consortium as a resource, provide an opportunity to exchange views, and to focus on the oversight roles under ICAO obligations. Recognizing the fact that many countries have limited resources, the Summit program identifies the human and financial resources needed to meet ICAO obligations for aviation safety and security oversight. To overcome the challenge, the Summit program promotes regional cooperation among countries to pool human and financial oversight resources. The main elements of the Summit program include economic benefits of a safe and secure aviation system, critical elements of safety and security, and available funding and technical resources. The program has been designed to provide a forum for high-level discussions and review of the ICAO Convention and obligations. The first day and a half of the Summit program is deliberately designed for ministerial level discussion, providing an opportunity for these leaders to share their views with all participants and to discuss the essential elements of civil aviation safety and security from a policy perspective.

The Summit program includes both international and regional summits. The International Summits are held in Washington DC in order to ensure as much dialogue as possible with top U.S. government officials, including meetings of delegations with appropriate level U.S. officials. Regional Summits are held in specific regions of the world, with participants coming to one host country in the region. The regional basis allows for a greater number of participants from civil aviation authorities across the region. In addition, other training and outreach programs have been held with a focus on inspector training, development of laws and regulations and managing a CAA. At the end of the third year, 14 International Summits, 3 Regional Summits and 3 training and outreach programs had been held. Figure 5A shows the number of countries that have attended the International and Regional Summits organized by the seven ICAO Regional Offices. One can see that the Summit program has covered most parts of the world from a regional perspective.



Total number of countries that attended Summits = 100

ICAO Member States = 190

Figure 5A. Countries attending the Summit Program (*Note*. From "Compilation of Minutes and Participant Evaluations," by GW Consortium, 2004. Total number of countries that attended Summits does not include repeat attendees).

The participants to the Summit program include representatives from a pool of Category I and Category II countries as well as countries that do not maintain flight operations into the United States. Figure 5B shows that 118 countries have attended the International and Regional Summits, including several countries, which attended the program twice.

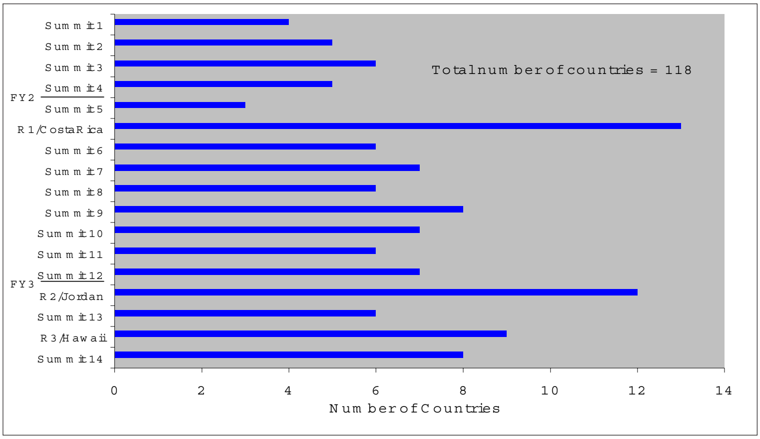


Figure 5B. Number of Countries attending the Summit Program (Note. From "Compilation of Minutes and Participant Evaluations," by GW Consortium, 2004. FY = Fiscal Year, R1 = Regional Summit 1, etc; excludes United States and organizational representatives, total number of countries include repeat attendees).

Finally, the Summit program data shows that the participating countries were serious about addressing the critical aviation issues. Figure 5C shows that nearly 44 Ministerial Attendees (which includes Ministers, as well as Deputy and Assistant Ministers), 70 Director Generals attended the Summits.

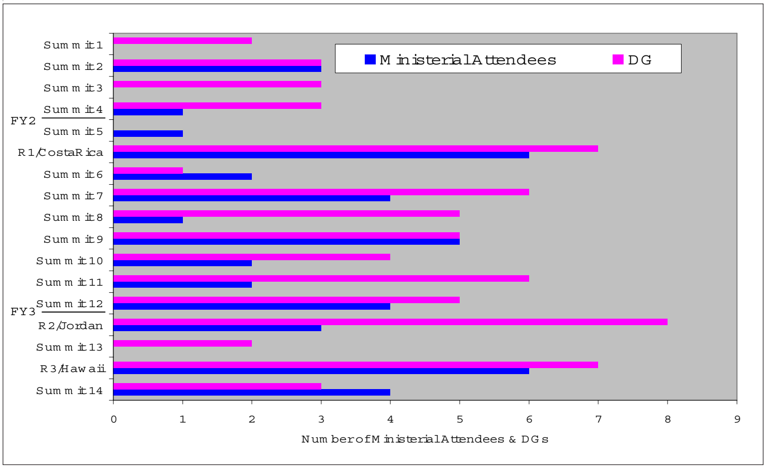


Figure 5C. Number of Ministerial Attendees and Director Generals per Summit (Note. From "Compilation of Minutes and Participant Evaluations," by GW Consortium, 2004. FY = Fiscal Year, R1 = Regional Summit 1, etc.).

Selected Summary of Issues and Challenges Facing Oversight Authorities

According to virtually all the respondents (81/84) of a post-summit survey, the Summit provided delegates with knowledge that they could use to assist them in their improvement of aviation safety and security oversight activities (GW Consortium, 2004). Figure 6 shows twelve areas of knowledge that the delegates indicated were beneficial.

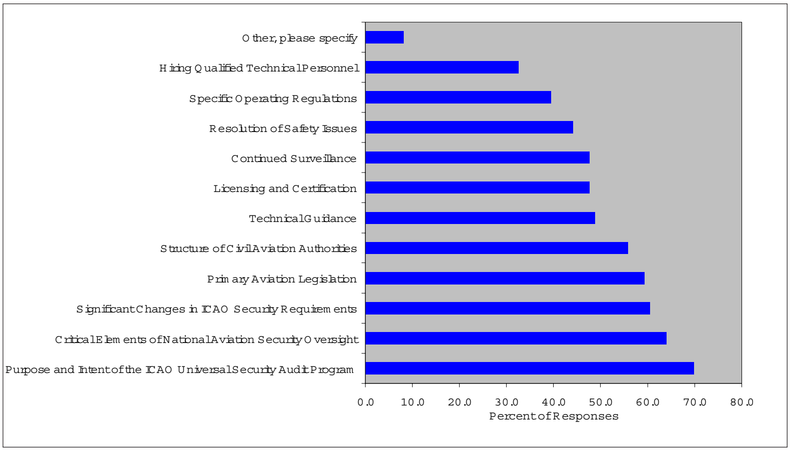


Figure 6. Areas of Knowledge considered Beneficial to the Delegates (*Note.* From "Report on the Past Delegates Questionnaires for the International Summit on Aviation Safety and Security," by GW Consortium, 2004).

According to the discussion minutes and feedback in the form of evaluations from participants, there are key issues that can be organized into five categories described in Figure 7. Not considering the "Other Category," which is made up of several single issues, training of CAA personnel is the primary concern, followed by funding and technical guidance. A less urgent issue involves establishing regional organizations to share resources and/or to improve IASA standings in order to increase flight operations into the United States.

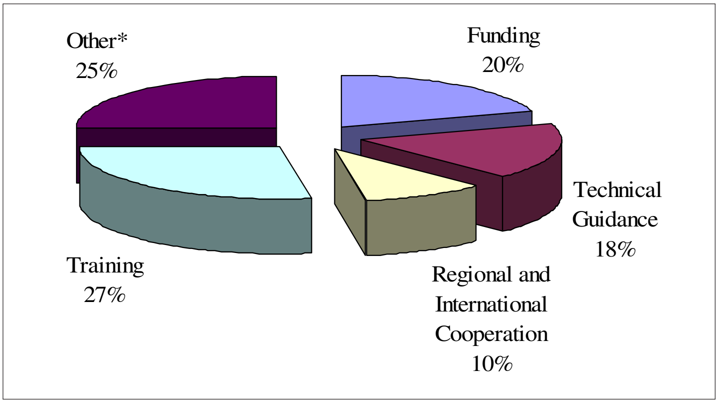


Figure 7. Key Issues for Aviation Oversight as identified by Officials (*Note.* From "Compilation of Minutes and Participant Evaluations," by GW Consortium, 2003. Other* = challenges noted by officials).

The "Other Category" includes ten recurring challenges to oversight that have been identified by oversight officials from the summit proceedings. Several of these challenges are identical to the audit common deficiencies, while the two challenges of aging aircraft and too many aviation safety audits are unique. Greater regional and international cooperation may help to immediately resolve the resource challenges and share effective methods to address the retention of trained personnel, the resolutions of safety issues, as well as aging aircraft.

Need for Political Commitment and Regional Cooperation

Cooperation among countries may take many forms given the distinctiveness of national interests and culture. For instance, countries may cooperate to form regional organizations concerned with sharing safety information and expertise. Alternatively, countries may form alliances to leverage complementing air carrier routes. Ten regional organizations created through an ICAO initiative to promote regional cooperation are of particular interest to the authors. The Cooperative Operational Safety and Continuous Airworthiness Activities Program (COSCAP) organizations are shown in Table 3. Each one is described by several characteristics; including, the name of the participating countries, the number of donors funding the project, the duration in years and the cost in millions of dollars for administering the project.

Table 3

COSCAP Regional Organizations and Their Characteristics

Project	Participating States	Donors	Duration	Cost
Former CIS States	Armenia, Belarus, Georgia, Kazakhstan, Kyrgyzstan, Azerbaijan, Tajikistan	Boeing, GE, Airbus, Ilyushin, IAC	2002-06	$3M
West Africa (UEMOA)	Benin, Burkina Faso, Cote d'Ivoire, Guinee Bissau, Mali, Niger, Senegal, Togo, Mauritania	TBD	2003-05	*
South Africa (CEMAC)	Central African Republic, Cameroon, Congo, Equatorial Guinea, Gabon, Chad, Sao Tome & Principe	TBD	2003-05	*
West Africa (Banjul Accord)	Cape Verde, Gambia, Ghana, Guinea, Liberia, Nigeria, Sierra Leone	TBD	2003-05	*
East Africa (EAC)	Kenya, Tanzania, Uganda	TBD	*	*
Latin America	Argentina, Bolivia, Brasil, Chile, Cuba, Ecuador, Panamá, Paraguay, Uruguay, Perú	Airbus, Embraer, Canada	2001-06	$2.6M
South Asia	Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, Sri Lanka	FAA, Canada, Norway, EC, Boeing, Airbus	1998-07	$5M
South East Asia	Brunei, Cambodia, Hong Kong, Macau, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, Viet Nam, South Korea	FAA, EC, Boeing, Airbus	2002-06	$1.5M
North Asia	China, Japan, North Korea, South Korea, Mongolia	*	2003-07	$2M
Pacific Aviation Safety Office (PASO)	Australia, Cook Islands, Fiji, Kiribati, Marshall Islands, Micronesia, Nauru, New Zealand, Niue, Palau, Papua, New Guinea, Samoa, Solomon Islands, Tonga, Tuvalu, Vanuatu	TBD	*	*

Note. From "ICAO COSCAP Initiatives," by ICAO, 2004. TBD = To Be Determined, * = no data. CIS = Commonwealth of Independent States, UEMOA = West African Economic and Monetary Union, CEMAC = Economic and Monetary Community of Central Africa, EAC = East African Community.

Comparing the ten COSCAP organizations to the regional designations used in preceding figures and tables, three of the African COSCAPs are represented by one region, the Eastern and Southern African Region. The Asia and Pacific

Region represent the Pacific and three Asian organizations. This comparison shows that even these organizations have formed smaller groups within their respective regions. Organizational structures such as these allude to cooperation among countries that takes many forms given the distinctiveness of national interests and culture. Examining UEMOA, the first West African COSCAP, reveals that all nine countries are francophone. In contrast, the members of the Banjul Accord, the other West African organization, consist of both English and French speaking countries.

Regional organizations should have as their ultimate objective enabling full compliance by all of their Member States with the standards of the Convention. In practice, partnerships should encourage countries to assist neighbors (not necessarily a geographic relationship) with preparations for regular safety and security audits. Other regional oversight organizations have been formed outside the COSCAP organizations, including the Central American Aeronautical Safety Agency (ACSA). The ACSA was formed in 2000 out of the Central American Air Navigation Services, an international government corporation providing air navigation services (GW Consortium, 2004). ACSA was charged with providing safety oversight services to its member states, which are Belize, Costa Rica, El Salvador, Guatemala, Honduras, and Nicaragua. Its priorities include: compliance with the ICAO standards; ensuring uniform levels of safety among member states (based on the Eight Critical Elements), establishing standardized safety oversight systems and processes for the region; eliminating duplication; and establishing a cost effective system of safety oversight. The member states benefit from economies of scale such as sharing the cost of highly qualified technical personnel and developing commercial activity in the region.

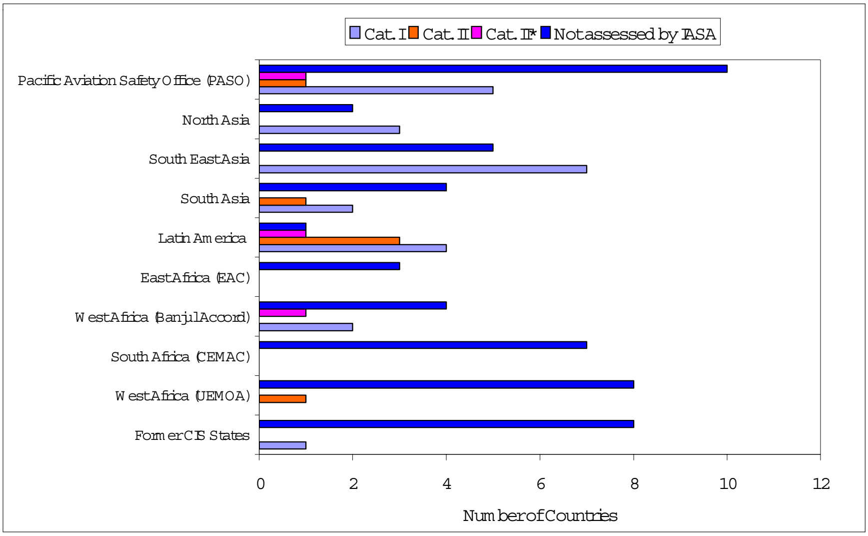


Figure 8. COSCAP Organizations Described by IASA Categories (FAA, 2003) (*Note.* From "Overview of the Federal Aviation Administration Flight Standards Service International Aviation Safety Assessment (IASA) Program," by FAA, 2003. Cat. 1 = Category I, etc.)

One method of determining the success of regional oversight organizations is through the examination of their audit results. Figure 8 shows the results of FAA IASA program assessments within specific geographic regions or regional organizations.

Figure 8 demonstrates that a number of countries have not been assessed by the IASA program. Thus, the authors are unable to determine the compliance status of such countries. Nonetheless, the data in Figure 8 provide useful information for identifying geographic areas that most urgently require additional resources or international attention.

Political, Safety, Security Elements, and Resources for a Model CAA

From the Summit program discussions and proceedings, the authors have developed a diagram for a model CAA that also identifies the key issues and challenges as noted by civil aviation officials (Figure 9). In addition to the safety and security elements, the GW Consortium argues that political elements, an adequate budget, technical, training, and human resources are just as important for building an effective CAA.

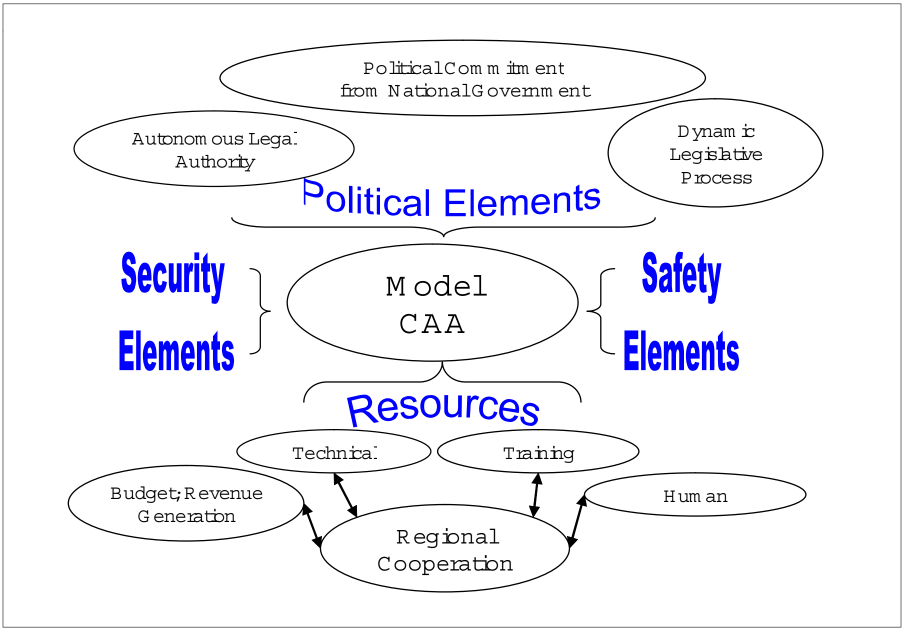


Figure 9. Political, Resource, Safety, and Security Elements Desired for a Model CAA

A model CAA is built on strong political commitment from the national government because this is essential to ensure an autonomous status and adequate resources. A dynamic, legislative process can establish and ensure that a CAA performs oversight functions while minimizing complicated and intractable political debate. The CAA should be an autonomous, legal entity free to complete its regulatory mission with a director that is free from political pressure.

The level of autonomy can vary, but the authority should be independent in practice. Ideal characteristics of this independence include; the legal authority to penalize, a separate budget, priority access to the Minister responsible for air transport, control of taxes and fees, as well as, direct use of all revenues in order to pay personnel competitive wages.

Adequate technical resources form an integral part of a model CAA because the authority must have the necessary technological advantages in order to ensure effective and efficient oversight of the complex aviation industry. Given the advancements of technology and their effect on regulations and due to natural turnover rates of qualified personnel, a CAA should have the capability to provide adequate training. Human resources includes a large enough pool of qualified technical personnel in fields such as general management, database management, legislation, security, airworthiness, operations, accident investigations, air traffic management, and airport operations.

For those countries that have the political commitment, but lack the adequate resources, the authors maintain that regional cooperation will help a CAA to attain the necessary resources. Regardless of form, successful cooperation can be characterized by political support from all member states, adequate financial, technical and human resources, adequate authority in the law, application of the critical elements, and a regulatory system that accounts for the harmonization of national variations.

Conclusions

The Convention on International Civil Aviation and its annexes have been signed by 188 countries in the world. The collection of Standards and Recommended Practices of ICAO has been developed over the past 60 years in a consensus process by the Member States whose aviation systems, governmental structures, laws, and oversight capabilities vary greatly. The complexity of the interconnected network of global aviation makes compliance with the SARPs by all Member States essential for the safety of the air transport system. A critical factor for successful compliance by Member States is the political commitment by the leadership of each country to civil aviation safety and security. The International Summits on the Aviation Safety and Security program has demonstrated that engaging Ministers and Director Generals is essential and effective in securing and sustaining such political commitment. Furthermore, the critical elements of safety and security oversight described herein have been shown to be applicable through out the world. Data presented in this paper show that while it is possible to develop and propose a model CAA structure, the key factors for sustaining an effective CAA for oversight are:

- Adequate budget and resources;
- Appropriate and robust laws and regulations to provide the necessary authority and enforcement mechanisms;
- Training programs; and
- Retention of qualified technical personnel.

References

Central Intelligence Agency. (2003). *The World Fact book 2003*. Retrieved November 21, 2003, from http://www.cia.gov/cia/publications/factbook/.

Federal Aviation Administration. (2003). *Overview of the Federal Aviation Administration Flight Standards Service International Aviation Safety Assessment (IASA) Program*. Retrieved October 18, 2003 from http://www.faa.gov/avr/iasa/iasabr15.htm

International Civil Aviation Organization. (2003). *2003 Air Traffic Remains Flat Improvements Expected in 2004*. Retrieved December 18, 2003, http://www.icao.int/icao/en/nr/2003/pio200319.htm

International Civil Aviation Organization. (2004). An Update on the of the ICAO Universal Safety Oversight Audit Programme. *Proceedings of the Seminar on Administration of the Technical Co-operation Projects of ICAO*. Singapore.

International Civil Aviation Organization. (1944). *Convention on International Civil Aviation (Chicago Convention)*. Montreal, Canada.

International Civil Aviation Organization. (2004)). *ICAO COSCAP Initiatives*. Retrieved January 21, 2004 http://www.icao.int/icao/en/tcb/TCB-Singapore-2004/Attachements/Presentations/HTMLICAO's%20COSCAP%20initiatives_files/frame.htm#slide0372.htm

International Civil Aviation Organization. (1999). *Safety Oversight Manual – Part A – The Establishment and Management of a State's Safety Oversight System*. Montreal, Canada.

Kotaite, A. (1998). *Universal audit programme conducted under auspices of ICAO would offer many benefits*. ICAO Journal, *53* (4) 5-6, 24-25.

National Transportation Safety Board. (1990). *Safety Recommendation.*. Washington, D.C., United States.

National Transportation Safety Board. (1993). *Scheduled 14 CFR Part 129: Foreign operation of AVIANCA (D.B.A. operation of AVIANCA).* Washington, D.C., United States.

The George Washington University Consortium. (2004). *Approaches to Successful ICAO Compliance – Regional Cooperation.* Washington, D.C., United States.

The George Washington University Consortium. (2003). *Critical Elements of Security Oversight.* Washington, D.C., United States.

The George Washington University Consortium. (2003). *Economic Benefit of Tourism and Safe and Secure Civil Aviation.* Washington, D.C., United States.

The George Washington University Consortium. (2003). *Overview of the Critical Elements of Security*. Washington, D.C., United States.

The George Washington University Consortium. (2004). Compilation of Minutes and Participant Evaluations. *Proceeding from International and Regional Summits on Aviation Safety and Security*. Washington, D.C., United States.

The George Washington Consortium. (2004). *Report on the Past Delegates Questionnaires for the International Summit on Aviation Safety and Security*. Washington, D.C., United States.

International Journal of Applied Aviation Studies, 4, Number 2

The Benefits of Implicit Learning When Training Ab-Initio Pilots

Daniel J. A. Rhind and Tony Head

Cranfield University
Bedfordshire, MK43 0AL
United Kingdom
Telephone: 01234 750111, Extension 4015
t.head@cranfield.ac.uk

Abstract

Twenty ab-initio pilots received either a traditional (explicit) training program or one that emphasized learning by experience (implicit). Trainees then completed six simulated approaches and landings (three of which involved a secondary task) in a flight simulator. The results showed that the participants in the implicit learning group performed as well or better than the explicitly trained participants did on seven of the eight performance variables. The implicit learners perceived the performance of the task as being associated with significantly lower level of cognitive load. They also performed significantly better on the secondary, performance duration estimate task. Recommendations are made regarding the potential benefits of implicit training programs, particularly those within safety critical environments.

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

Introduction

Pilots recognise that the landing phase of a flight is the leading cause of non-fatal aircraft accidents (Balfour, 1988; Nagel, 1988). Pilots and certified flight instructors (CFIs) also acknowledge that it is very difficult to explain how they perform certain aspects of their operation and the landing process. For example, pilots find it difficult to identify the cues and information which they consult when determining height Above Ground Level (AGL) as they approach the runway (Benbassat & Abramson, 2002a; Hasbrook, 1979). Due to the fact that it is difficult to articulate such skills, pilots are required to learn them through 'experience' (Benson, 1999; Thom, 1992). Furthermore, Benbassat and Abramson (2002b) also reported research which advocated a more experiential approach to the training of the skills involved in landing an aircraft.

Traditional approaches to pilot training usually are based upon fundamental theories of motor skill acquisition. These theories suggest that motor skills are initially acquired explicitly via cognitive processing which is verbally based. Through the learning process, the skill becomes automated (implicit) as the verbal rules associated with its performance are "forgotten" and task-related information is processed subconsciously (Anderson, 1983; Fitts & Posner, 1967). Within this process the learner is said to adopt a hypothesis-testing approach in that various techniques are used with successful attempts being remembered for future performance. This selective or S-mode learning style is theorised to lead to a small knowledge base which is easily verbalised (Berry & Broadbent, 1988).

More recent research has questioned the assumption that the learning of a motor skill must proceed from explicit to implicit, which underlies the more traditional theories of learning. Implicit learning involves the acquisition of a skill without a corresponding increase in the verbal knowledge associated with the skill (Maxwell, Masters, Kerr, & Weedon, 2001). The performer passively aggregates all task-related information and action-outcome contingencies which leads to a relatively large knowledge base that is not easily verbalised. This process is defined by Berry and Broadbent (1988) as Unselective or U-mode learning.

Cleeremans, Destrebecqz, & Boyer (1988) conducted a meta-analysis of studies which suggested that people can indeed acquire information regarding the underlying structures of situations incidentally. Simply through experience, an individual can implicitly acquire information about a given skill. They then are able to perform that skill without necessarily being able to explain how they performed it. A good analogy of this would be the ability to ride a bicycle. People can learn how to ride a bicycle without being directly instructed and even without observing other people perform the skill. It may be suggested that implicit learning would be most effective for the acquisition of other motor skills. In terms of Multiple Resource Theory (Wickens, 1984), the information received regarding the task would be in the spatial code. As a motor task is likely to be primarily spatial, this means that the acquired information will not need to be converted as it is already in the appropriate modality.The stimuli can be directly mapped onto the appropriate behaviour without the need for any conversions or conscious selections to be made (Schneider & Shiffrin, 1977). That is, the motor skill is learned through a process of observation and experience without complex verbal explanation. In contrast, if much of the information is presented in the verbal code then the trainee will need to convert it into spatial information before they can perform the motor task. This need for conversion may subsequently inhibit the learning process and may even distort the skill being learned.

More recently, research which investigated the use of a more implicit approach in the learning of a motor skill has been conducted. Maxwell et al (2001) conducted a study in which participants acquired a complex motor skill – golf putting. Half of the participants completed a secondary, random letter generation task (Baddeley, 1966), whilst putting, which was designed to prevent explicit learning. The other half of the participants formed the control group which performed a series of puts without a secondary task. It was assumed that this approach would facilitate a hypothesis-testing approach and lead to the acquisition of the verbal rules associated with the task. Both groups were found to have significantly improved over time, with the 'explicit learners' demonstrating a significantly larger amount of verbal knowledge regarding the skill. The performance of the 'implicit learners' was also found to be more robust against the potentially debilitating effects of psychological pressure relative to the 'explicit learners' Maxwell et al (2001) concluded from this that implicit learning can indeed precede explicit learning. Similar results were suggested in an earlier study (Hardy, Mullen & Jones, 1996).

Implicit and explicit learning can therefore be conceptualised as two distinct processing systems rather than being along a continuum. The importance of this distinction goes far beyond the notion that implicit learning can be just as effective as explicit learning. The significance of this new concept lies in the benefits which are associated with skills which have been implicitly learned. Firstly, in contrast to explicitly learned skills, implicit learning leads to skills whose performance does not require abstract working memory (Broadbent, 1984). This ensures that the skills can be completed without placing a demand upon the individual's memory processes and consequently the performance of the task may be perceived as being associated with a lower level of workload. Based upon a similar notion, implicitly learned skills are also more robust against the potentially debilitating effect of psychological pressure (Reber, 1989, 1993; Berry & Dienes, 1993). This can be explained via a concept known as the 'Response Selection Bottleneck' (RSB, as presented by Pashler, 1994).

The RSB theory suggested that when a person is performing a task, they first encode the task stimulus from the environment in which it is to be performed. From this information, a response is selected and finally the response action is executed. For example, a pilot may perceive that the aircraft is too low during its approach to the runway (encode task stimuli), choose to initiate a climb (select response) and then pull back gently on the control stick (response execution). In the case of a student pilot, this sequence of events will be processed explicitly as they have to 'think' about their actions. For an experienced pilot, however, this process will have become automated and hence will be processed implicitly.

Errors in performance have the potential to occur in dual task conditions. The RSB principle theorises that the encoding, response selection and response execution sequence can be processed for one action at a time. Therefore, when two actions are required simultaneously, an error/delay can occur if one attempts to perform them both explicitly. Thus the benefits of implicitly learned skills are that they are more robust against the debilitating effects of psychological pressure and they should be less stressful for the trainee to perform under dual-task conditions, e.g. whilst they are learning different or advanced landing techniques. These advantages become evermore important within a safety critical environment.

Some theorists have advocated a greater emphasis on experiential learning. For example, Knowles (1975) presented the Andragogy Theory which suggested that adults work best with self-directed, problem-solving learning. Furthermore, Rogers (1969) proposed the view that humans have a natural propensity to learn and that the teacher's role should focus on facilitating the learning process. In order to assess the possibility of adopting a more implicit approach to the training of ab-initio pilots, the current study was conducted. A control group received the typical (explicit) training brief which would be given whenever an ab-initio pilot was being instructed regarding the approach and landing stage of a flight. An experimental group was given a training programme which relied far more upon experience and was based upon facilitation (implicit learning) rather than demonstration and detailed instructions.

Hypotheses

The hypotheses to be tested were:

1. The implicit learners will perform as well as the explicit learners under test conditions.
2. The implicit learners will perform significantly better than the explicit learners on a secondary task whilst under test conditions.
3. The implicit learners will report a significantly lower level of perceived workload after completing the task relative to the explicit learners.

Method

Participants

Twenty participants took part in the experiment voluntarily. None had any prior flight training experience, ensuring that they did not already have the appropriate psychomotor sets in place. This allowed the efficacy of the two training approaches in developing the skills to be evaluated. Participants were randomly assigned to either the explicit or implicit training condition. This resulted in there being 10 participants in each of the two training conditions.

Flight Simulator

As a result of the safety critical nature of the task to be performed, it was not possible to train the participants within an actual aircraft and hence the Aerosoft Flight Simulator at Cranfield College of Aeronautics was used. The Aerosoft

Flight Simulator is a part-task trainer which uses standard PC technology for the three computers that run the Flight Model, Visual System and Instructor Station. The flight instruments are displayed electronically on a monitor in front of the user.The visual system uses a collumated display (set at infinity) to aid depth perception. The simulator is fully enclosed to minimise the influence of external distractions.

The simulator made it possible for the instructor to determine the parameters of the task in order to control for variables which could potentially have a confounding effect (e.g. the weather conditions and the aircraft's performance). The simulator was configured to perform like and display the cockpit instruments of a PA28 aircraft which is typical of the kind of aircraft which is used when training ab-initio pilots. A 'script' programme was written as part of the research such that the desired flight parameters could be set and the necessary flight data related to the dependent variables could be recorded. The features of the approach and landing programme were modelled on those of Cranfield Airport (EGTC). The flight commenced five nautical miles from the runway at an altitude of 2000 feet and an airspeed of 80 knots. The aircraft's Altitude, Airspeed, Heading and Glideslope error (deviation from optimal glideslope) were recorded at one second intervals throughout each flight.

Workload Measure

The NASA Task Load Index (NASA-TLX) developed by Hart and Staveland (1988) was used to provide a subjective measure of the workload associated with the performance of the flying task. The Prospective Duration Estimate technique (PDE) was used to provide an objective measure of the participants' task load and hence an indication as to their spare capacity whilst performing the task. The implementation of the PDE technique was based upon that used by Zakay and Shub (1998). The participants were instructed that when they heard the instructor say 'Go' they should mentally count for 14 seconds and then activate a switch when they felt that the time had passed. Once the switch was activated the actual time which had elapsed was recorded. The temporal distance between the duration time and 14 seconds was then calculated (deviation). The magnitude of this deviation is negatively correlated with the amount of non-temporal information being processed during the time estimation (Zakay, 1993).

Experimental Task

The flight commenced five nautical miles from the runway threshold and at a height of 2000 feet AGL. Participants were required to maintain an optimum speed of 80 knots throughout the flight and to fly the aircraft directly towards the runway on a bearing of 220 degrees. The participants' mission was to fly the aircraft at an Altitude of 2000 feet for the first nautical mile of the flight, at which point the aircraft intercepted the Glideslope of the instrument landing system. For the remainder of the flight, the participants were required to fly the aircraft as close to the optimum Glideslope as possible whilst still flying on a bearing of 220 degrees at a speed of 80 knots.

Experimental Design

An independent group's design was used with explicit or implicit training approach as the independent variable. As the participants were required to perform six simulated approaches and landings during which data was collected, a within-groups factor was formed.

Procedure

Volunteers with no prior flight experience were trained to land a PA28 in an Aerosoft Flight Simulator.All participants received a ten-minute standardised basic control and instrument guide which was based upon the information within Thom (1997). Thus all of the participants were told about the operation and function of the control column and the throttle. Furthermore, they were informed as to the location of the Heading Indicator,Air Speed Indicator,Artificial Horizon, Altimeter and the Glideslope Indicator.They all were instructed regarding the function of each display and how it should be read and interpreted. It was intended that this information would be equivalent to that which would be given to any trainee pilot before a first flying lesson. Participants then were randomly assigned to either the control (explicit) or experimental (implicit) learning condition. Participants within the control group received detailed instructions of the relationships between the primary flight controls and indicators. They were trained on the performance of the Basic 'T' and Selective Radial instrument scans and instructed to use each of these techniques throughout the simulated flight. For example, in terms of the performance of the Basic T scan, explicitly trained participants were informed that they should first focus on the Artificial Horizon. They then were told that they should check the display to the left before returning back to check the Artificial Horizon. Then they were instructed to check the indicator above the Artificial Horizon before returning to it and then they should finally check the indicator to the right. They were instructed to perform this Basic T scan whilst completing the task. In contrast, those within the experimental training condition were merely told to complete the task and were allowed to learn experientially without the instruction to follow strict procedural guidelines.

All participants then completed three practice approaches and landings during which no data was collected. For the first of these flights, the cloud base was set at 2500 feet, meaning that the runway was visible throughout the flight. For the second practice flight, the cloud base was lowered to 1000 feet. This was to encourage the participants to begin using the Instrument Landing System within the cockpit. For the final practice flight, the cloud base again was lowered to 500 feet, which meant that the runway was not visible for the majority of the flight. The cloud base was maintained at this altitude throughout all of the test flights. This ensured that the participants had to fly the simulator using the instruments panel rather than relying upon visual information.

Once the practice stage had been completed, the details of the mission were reiterated and the participants were informed that data was going to be collected from that point onwards. Participants completed three approaches and landings. Details of the mission were reiterated again to all participants before a final three approaches and landings were made. During each of these final trials, the PDE task also was concurrently performed at a random point in time. Once their trials had ended, the participants completed the NASA-TLX Task Load Index.

Measures

The data recorded consisted of the aircraft's Altitude, Airspeed, Heading, and Glideslope error. The simulator stored data at one second intervals for each of the dependent variables for each participant. The variables concerning Altitude, Airspeed and Heading were then computed in order to calculate the participants' deviations from the optimum values required by the task. In other words, the average amount by which the participants' performances differed from those stipulated by the task was calculated. The standard deviations of these differences also were determined. This process created four new performance variables: Altitude Error, Airspeed Error, Heading Error and Glideslope Error.

Results

Treatment of Data

Each of the four performance variables, for each participant, for each trial were summarised to give two measures of error: the within-trial arithmetic mean error and the within-trial standard deviation of error. The performances of the participants relative to the four mission objectives (Altitude, Speed, Heading and Glideslope) are displayed in Tables 1-4. The use of the arithmetic mean may be criticised because of minus signs, negative errors may cancel out positive errors. However, when the arithmetic mean error is combined with the standard deviation of error, it does have some advantages. The arithmetic mean error will give an indication as to whether or not there is a consistent error tendency whereas the standard deviation of error will show the variability within performance. Both of these measures of errors will have normally shaped distributions, which mean that they are appropriate for the application of parametric statistics. This statistical procedure has been implemented and justified within previous research involving flight simulation (Rees & Harris, 1995).

Altitude

No significant differences in the deviations from the optimum approach altitude of 2000 feet were found between the two training conditions, $F(1) = 0.01$, $p > .05$). However, there was a highly significant difference in the variability of performance on this criterion, $F(1) = 17.27$, $p < .01$). The participants within the implicit training condition displayed considerably less within-trial variability in performance ($M = 116.24$, $SD = 68.45$) relative to those participants who were trained in the explicit condition ($M = 177.80$, $SD = 94.67$). Further analysis of the mean deviations found no significant effect over trials, $F(5) = 2.15$, $p > .05$, nor was there a significant Trial x Training Condition interaction, $F(5) = 0.10$, $p > .05$. A significant trial effect was found; however, in the variability of performance, $F(5) = 2.34$, $p < 05$. Although the variability in performance of the implicit learners did reduce more quickly over trials, relative to the explicit learners, no Trial x Training Condition interaction was found, $F(5) - 0.53$, $p > .05$.

Table 1

Mean Within-Trial Deviations and Mean Variability from the Optimum Approach Altitude

	Deviations from the Optimum Altitude (feet)			
	Mean Error		SD of Error	
Trial Number	Explicit Mean SD	Implicit Mean SD	Explicit Mean SD	Implicit Mean SD
1	-271.63 (490.30)	-323.35 (544.93)	195.68 (61.02)	153.66 (89.48)
2	-345.46 (304.95)	-306.19 (289.41)	230.90 (145.58)	127.14 (82.92)
3	-145.77 (350.06)	-171.35 (218.41)	195.65 (122.79)	118.75 (66.30)
4	-69.40 (262.10)	-125.47 (210.14)	169.38 (69.18)	100.67 (34.80)
5	-129.22 (218.03)	-105.52 (255.67)	147.37 (44.73)	105.17 (58.78)
6	-114.23 (275.91)	-73.64 (110.56)	127.83 (65.20)	92.02 (63.43)

Speed

A significant difference was found in mean deviations from the optimum speed, $F(1) = 5.00$, $p < .05$). Participants within the explicit training condition flew the aircraft with a consistently lower level of error from the optimum speed of 80 knots ($M = 10.47$, $SD = 13.34$) relative to those in the implicit condition ($M = 17.24$, $SD = 18.71$). No significant differences were found in the variability of performance, $F(1) = 0.48$, $SD = p > .05$). No significant Trial x Training Condition interaction was found in the mean deviations from the optimum speed, $F(5) = 0.70$, $p > .05$, nor in the variability of performance, $F(5) = 1.64$, $p > .05$.

Table 2

Mean Within-Trial Deviations and Mean Variability from the Optimum Speed

	Deviations from the Optimum Speed (Knots)			
	Mean Error		SD of Error	
Trial Number	Explicit Mean SD	Implicit Mean SD	Explicit Mean SD	Implicit Mean SD
1	10.74 (11.51)	22.55 (13.47)	24.46 (5.16)	23.84 (5.44)
2	14.46 (10.37)	20.14 (15.82)	24.31 (4.08)	22.60 (6.58)
3	10.71 (22.01)	16.22 (15.62)	24.10 (6.50)	24.46 (5.47)
4	8.86 (14.67)	8.32 (32.15)	21.74 (6.05)	22.31 (4.13)
5	8.83 (7.58)	18.91 (14.30)	21.59 (5.89)	20.15 (3.61)
6	9.23 (12.01)	17.32 (15.69)	21.46 (4.32)	20.14 (7.55)

Heading

No significant differences in the deviations from the optimum heading were found between the performances of the two training conditions, $F(1) = 3.17$, $p > .05$. However, a significant difference was observed in the variability of performance, $F(1) = 5.96$, $p < .05$. Participants in the implicit training condition exhibited a significantly lower level of variability in performance ($M = 10.48$, $SD = 15.36$) relative to the participants in the explicit training condition ($M = 15.81$, $SD = 8.17$). No significant trial effect was found in the mean deviations from the optimum heading, $F(5) = 0.37$, $p > .05$.

Table 3

Mean Within-Trial Deviations and Mean Variability from the Optimum Heading

	Deviations from the Optimum Heading (degrees)			
	Mean Error		SD of Error	
Trial Number	Explicit Mean SD	Implicit Mean SD	Explicit Mean SD	Implicit Mean SD
1	4.44 (6.75)	-5.69 (20.63)	18.51 (10.30)	22.19 (30.29)
2	4.06 (5.35)	0.77 (3.50)	18.20 (8.52)	12.43 (9.35)
3	0.98 (7.59)	0.66 (5.05)	16.58 (8.64)	11.40 (16.35)
4	0.77 (2.70)	0.42 (1.34)	13.93 (6.25)	6.60 (3.20)
5	1.40 (3.50)	0.91 (3.06)	14.62 (7.85)	5.14 (3.20)
6	0.40 (1.05)	0.81 (2.45)	13.04 (6.60)	5.15 (4.00)

A significant trial effect was found in the variability of the participants' performances, $F(5) – 2.59$, $p < .05$. It was observed that the performances of the participants in both training conditions did improve on this criterion over time. However, it appeared that the improvement was far greater for those within the implicit training condition, although no significant Trial x Training Condition interaction was found $F(5) = 0.76$, $p > .05$). No significant Trial x Training Condition interaction was found in mean deviations from the optimum heading, $F(5) = 1.53$, $p > .05$.

Glideslope

No significant differences were found between the performances of participants in the two training conditions in the mean deviations from the optimum glideslope, $F(1) = 1.39$, $p > .05$, nor in the variability of performance, $F(1) = 2.71$, $p > .05$). Furthermore, no significant trial effect on mean deviations was found, $F(5)\ 0.13$, $p > .05$. However, a significant trial effect was found for the variability in performance with reference to the optimum glideslope, $F(5) = 2.84$, $p > .05$. Participants in both training conditions did improve over time, but overall the implicit learners did appear to learn more quickly, although again the Trial x Training Condition interaction was not significant (F(5) = 0.69, p > .05). No significant Trial x Training Condition effects were found for the participants' mean deviations, F(5) = 0.27, p > .05.

Table 4
Mean Within-Trial Deviations and Mean Variability from the Optimum Glideslope

	Deviations from the Optimum Glideslope (degrees)			
Trial Number	Mean Error		SD of Error	
	Explicit Mean SD	Implicit Mean SD	Explicit Mean SD	Implicit Mean SD
1	1.10 (2.04)	0.38 (2.87)	3.10 (1.62)	3.45 (2.74)
2	1.16 (1.32)	0.99 (0.97)	2.76 (1.33)	1.90 (1.04)
3	0.77 (1.57)	0.84 (0.82)	2.19 (0.82)	1.92 (1.27)
4	0.88 (1.25)	0.92 (0.87)	2.48 (1.75)	1.58 (1.16)
5	1.24 (0.82)	0.60 (1.15)	1.85 (0.62)	1.83 (1.58)
6	1.18 (1.65)	0.71 (0.71)	2.27 (1.34)	1.36 (1.61)

NASA-TLX Rating

An independent t-test was computed with the independent variable being the training condition (explicit or implicit) and the dependent variable being the participants' scores on the NASA-TLX. This revealed that there was a highly significant difference between the ratings of the two groups (t = 4.15, df = 18, p < 0.01). Participants within the explicit training condition provided significantly higher subjective workload ratings (M = 65.43, SD = 10.79) than the ratings given by the participants in the implicit training condition (M = 45.83, SD = 10.32).

Prospective Duration Estimate scores

The extent to which the PDE task interfered with primary task performance was assessed by comparing the participants' performances on trials 3 and 4 (the first trial which included the PDE task). An independent t-test revealed that there was no significant difference between the participants' performance on trials 3 and 4 on any of the 8 dependent variables. A summary of these results is shown in Table 5.

Table 5

Testing for any interference effects associated with the performance of the PDE task

	Trial 3		Trial 4		
Variable	Mean	SD	Mean	SD	t
Altitude Error	158.56	(284.28)	-97.43	(232.99)	-0.74*
Altitude Variability	157.20	(103.83)	135.02	(63.90)	0.81*
Speed Error	13.46	(18.79)	8.59	(24.33)	0.71*
Speed Variability	24.28	(5.85)	22.02	(5.05)	1.31*
Heading Error	0.82	(6.28)	0.59	(2.08)	0.15*
Heading Variability	13.99	(13.00)	10.26	(6.12)	1.16*
Glideslope Error	0.81	(1.22)	0.90	(1.05)	-0.26*
Glideslope Variability	2.05	(1.05)	2.03	(1.52)	0.07*

* = Not significant at the 5% level.

Table 6

Summary of deviations on the PDE task

Trial Number	Explicit	Implicit	t
	Mean SD	Mean SD	
4	1.27 (0.86)	-0.16 (0.52)	4.52**
5	1.33 (0.79)	-0.44 (0.42)	6.26**
6	0.69 (0.61)	-0.22 (0.56)	3.46*
PDE Scores	1.10 (0.64)	0.35 (0.30)	3.35**

* = Significant at the 5% level, ** - Significant at the 1% level.

In order to compare the performances of the two groups on the secondary task, a series of independent t-tests were conducted (where the independent variable was training condition, explicit or implicit, and the dependent variable was the participants' score on the PDE task). A summary of these results is shown in Table 6. Significant differences were found for each of the trials in which the PDE task was completed. When the overall PDE score was computed a highly significant difference was found ($t = 3.35$, $df = 18$, $p < .01$). The participants within the implicit training condition produced estimated durations with a significantly lower amount of deviation from the required time of 14 seconds ($M = 0.64$, $SD = 0.30$) relative to the participants within the explicit training condition ($M = 1.10$, $SD = 0.35$).

Relationship between workload measures: In order to determine whether or not there was a correlation between the subjective and objective measures of workload, a Pearson's Product Moment Correlation Coefficient was calculated. This revealed that a positive relationship was evident within the data between the participants' PDE score and the rating which they provided via the NASA TLX ($r = 0.48$, $p < 05$). Therefore, participants who reported a lower level of subjective workload were more likely to perform better on the secondary (PDE) task.

Discussion

Effectiveness of the Training Programmes

The present study investigated the potential benefits of teaching ab-initio pilots to land an aircraft (simulator) via a training programme which attempted to facilitate implicit learning. Firstly, it is important to establish whether or not any learning actually took place. The data indicated that significant improvements did occur in the level of variability within performance regarding Altitude, Heading and the Glideslope. Therefore, the participants within the control condition, using detailed instructions typical of what is currently used within pilot training, did exhibit a significant level of learning. The improvements made by those participants within the implicit training condition proved interesting: There was no direct need to inform them of the consequences of their behaviour or to positively/negatively reinforce them for performing well or poorly.There was also no need for participants to actually observe another person performing the task. One may conclude from the fact that as learning took place without the need for reinforcement, observation or a model, that some form of experiential/implicit learning had occurred.

Comparing the Training Programmes

Further analysis showed that the two groups did not differ in their level of Error in terms of Altitude, Heading or Glideslope. Nor did variability within their performances significantly differ in terms of Speed and the Glideslope. Therefore, the participants in the implicit training condition performed as well as the participants within the more traditional explicit training condition on five of the eight flight variables. This provided evidence for the acceptance of the first hypothesis in that, overall, the implicit training programme did promote performances which were comparable to the performances of participants within the control condition.

In fact, for some of the test variables, the implicitly trained participants actually performed significantly better relative to the explicitly trained participants. They were found to perform with a significantly lower level of variability in Altitude and Heading. These differences in performance may be due to the nature of the two training programmes used for each condition. As a result of receiving relatively less information, the implicit learners may have completed certain aspects of the task on a subconscious level. Rather than consciously thinking about flying the aircraft towards the runway, the implicitly trained participants may not have needed to consciously 'think' about flying straight, they just did it. This is, however, a very difficult conclusion to draw as inferences about the learning process must be made based upon the product (flying performance) as the cognitive process can obviously not be observed directly.

In terms of the Multiple Resource Theory (Wickens, 1984), it also could be argued that the implicitly trained participants received the majority of their training information via the spatial information code through observing and evaluating their own performance. There would have been no need, therefore, for these individuals to convert this information when accessing the spatial component of their working memory as the information already would have been in the correct modality (i.e. spatial rather than verbal). In contrast, the explicitly trained participants received the majority of their training information through verbal instructions and, hence, via the verbal information code. They then would have needed to convert this information into spatial instructions when accessing the spatial component of their working memory in order to practice and perform the task. Thus, the requirement to convert the information during task performance may have slowed down the learning process and hence explain why the implicit learners performed better on certain variables and appeared to be learning at a faster rate. The presumed additional information/cognitive processing workload imposed by such information conversion is reflected in the poorer PDE scores and the increased NASA TLX reported workload scores in the explicit learning group.

The procedural nature of the instructions given to the explicit learners may explain why the participants within the control condition did perform significantly better on the Speed Error element of the task. Therefore, although the implicitly trained participants exhibited a lower level of variability within performance on some test variables, they consistently were out-performed in terms of the difference between their actual speed and the optimum speed. An explanation for this could be that the implicitly trained participants did not view the speed element of the task as being as important as the aircraft's direction and height. They may have failed to regularly look at the speed indicator and instead concentrated on the direction and heading indicators. In contrast, the explicitly trained participants were given clear and detailed instructions of the procedures which they should follow during the task. This may have ensured that each indicator was checked regularly with each element of the task being equally important. It also could have been that variability in performance was predominantly a motor skill, whereas, the Speed Error variable was predominantly a procedural task which demanded that the Speed Indicator was checked regularly so that any required adjustments could be made. Therefore, it could be suggested that implicit learning could be more effective with perceptual motor tasks, such as visual flight rules (VFR) flying; whereas, explicit learning could be a more effective method for procedural tasks. By co-incidence, instrument flying is predominantly procedural in nature. As an aside, during the tasks, neither group of trainees verbally coached themselves, although a few did curse if they noticed a significant error.

Workload

The implicitly trained participants were found to perform significantly better on the Prospective Duration Estimate (PDE) task. Based upon the attentional model, (Zakay & Shub, 1998), one could conclude from this that the implicitly trained participants experienced a significantly lower level of workload and hence retained a higher amount of spare capacity when performing the task. Hayes and Broadbent (1988) suggested that performing implicit skills does not require abstract working memory.This would suggest that certain aspects of this task were being completed implicitly ensuring that a higher level of working memory was left available for use when performing the secondary task. The implicit learners may have viewed the instruments, recognised the adjustment which needed to be made and then activated the appropriate motor programme without having to consciously 'select' the movement which was required. This explanation would concur with the work of Schneider and Shiffrin (1977) who suggested that implicit skills are cognitively mapped direct from stimulus to response without the need for the selection stage to take place. Thus, less conscious information processing takes place and the learner has more spare capacity with which to carry out other tasks.

The participants in the explicit training condition, on the other hand, were given clear and detailed rules and procedures to follow when practicing and completing the task. This was designed to promote an explicit learning strategy in which the participants had to consciously think of the indicators which needed to be checked and then which control to move in order to make any necessary adjustments. The explicitly trained participants are likely to have needed to make a series of small selections as to what they needed to do next. The concept of the Response Selection Bottleneck, as presented by Pashler (1994) may go some way to aid an understanding of why this explicit form of training may delay the learning process and perhaps detracted from the participants' performance. For example, if a participant within the explicit training condition observed that a change in altitude was required, they still may be attempting to select the appropriate movement whilst also consciously trying to remember which indicator they should be looking at next. As the model suggested, two conscious explicit selections cannot be made simultaneously (especially during the learning phase of a skill), a delay could be caused between the change in altitude being made and them moving onto the next indicator.This may help to explain why the explicitly trained participants exhibited a higher degree of variability within performance and why they performed less well than the implicit group on the secondary PDE task. Thus, the second hypothesis can be accepted in that the implicit learners did perform significantly better than the explicit learners on a secondary task under test conditions. The successful application of the PDE technique within this research lends further support to those who have previously advocated PDE scores as being a useful secondary task for obtaining an objective measure of (conscious) workload and spare capacity (Zakay & Shub, 1998).

The difference in workload between the two groups indicated by their performances on the secondary task was supported by the participants' subjective ratings of workload (NASA TLX), which were found to be significantly lower for those within the implicit training condition. This again may be explained by the concept of the Response Selection Bottleneck (RSB) model as presented by Pashler (1994). The implicit learners theoretically should be able to complete more elements of the task implicitly and hence, by making fewer conscious selections. It could therefore be suggested that the explicit learners would have been more likely to be required to make simultaneous selections whilst performing the task. As the RSB model suggests, this cannot happen and so pressure will build up until the first selection is made. This will place a greater demand upon the participants' attentional resources and the pressure created by the need for simultaneous selections can be interpreted by the participants as representing a higher workload. The third hypothesis can be accepted in that the implicit learners did report a significantly lower level of workload associated with the performance of the task under test conditions.

Limitations

There are some ways in which the methodology used within this research could be improved. Firstly, the participants only received one training session which lasted approximately one hour. It could be argued that it is difficult to draw too many conclusions based upon a relatively small amount of training; therefore, a longer training period could have been used. However, the trainees were ab-initio students and the study protocol was meant to reflect initial learning. Secondly, although the participants were randomly assigned to one of the training conditions, this still may not have ensured individual differences were controlled for. Despite the fact that none of the participants had any prior flight experience, presumably some of them would have been naturally better at the task than others and this could have had an effect on their performances. The alternative approach would have been performance/ability testing of each participant, with matched groups. However, this approach would have meant that the participants were no longer naive. Thirdly, the very nature of learning itself makes it difficult to accurately define what has taken place. One can only make inferences regarding the process based upon the product (i.e. task performance) and this may not always capture what has actually occurred. Finally, although some benefits of implicit learning were demonstrated within this study, it does involve the use of a flight simulator. It may be that any such benefits would not be present if the training took place in an actual aircraft with greater environmental variability. However, it could be hypothesised that such conditions would favour implicit learning!

Conclusions and Recommendations for Future Research

This study has re-visited the concept of implicit and explicit learning as two distinct processing channels rather than being along a continuum. The more traditional theories of the learning process (i.e. Fitts & Posner 1967; Anderson, 1983) require that explicit information is required before perceptual motor skills be acquired. The findings within this study do support those within Cleeremans et al's (1998) meta-analysis of the learning processes, in that people can learn incidentally simply through experience, without explicit instruction. This study also supports the findings of Maxwell et al (2001) in that the performances on a secondary task, in those who have been implicitly trained, appears to be significantly better than those who have received more traditional instruction based upon explicit learning. Further research is required in this field to determine the feasibility of adopting pilot training programmes which focus more upon implicit learning, with a view to benefiting from the associated advantages of a reduction in perceived workload and an increase in spare capacity. Such an approach might be more beneficial where the emphasis is upon learning perceptual motor skills rather than procedural skills. Such an approach could facilitate a reduction in the frequency of subsequent errors made by trainee pilots and ultimately the number of accidents. It could be concluded from this research that, for certain, yet to be defined tasks, the teacher role could be more a one of facilitation rather rule-based instructions. It also might be suggested that trainers should endeavour to avoid the temptation of providing an excess of verbal information when a relatively difficult motor skill is being taught as this may only have a negative impact by slowing the learning process. Too much verbal information also may make the task appear more difficult than it really is and reduce the trainees' spare capacity whilst performing the task. Both children and adults have a natural propensity to learn and, where appropriate, should be allowed to do so experientially through guided practice (Knowles, 1975; Rogers, 1969). The potential benefits of implicit learning could be advantageous to any safety critical situation in which a robust motor skill is to be learnt. Future research could explore the learning of different skills within safety critical activities other than flying, especially where multiple motor tasks and cognitive tasks must be carried out simultaneously. The results of this study suggested that moving towards training programmes which incorporate implicit learning could promote at least equivalent performances to traditional training, whilst maintaining a lower level of perceived workload and a higher degree of spare capacity in the trainee.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Baddeley,A. D. (1966). The capacity for generating information by randomisation. *Quarterly Journal of Experiment Psychology*, *18*,119-129.

Balfour,A. J. C. (1988). Accident investigation and its management. In J. Ernsting & P. Kind (Eds.) Aviation Medicine (2nd ed., pp. 697-702), Oxford, UK: Butterworth Heinemann.

Benbassat, D. & Abramson, C. I. (2002a). Landing flare accident reports and pilot perception analysis. *International Journal of Aviation Psychology, 12*, 137-152.

Benbassat, D. & Abramson, C. I. (2002b). Errorless discrimination learning in simulated landing flares. *Human Factors and Aerospace Safety, 2*(4), 319-338.

Benson, A. J. (1999). Spatial Disorientation – general aspects. In J. Ernsting, A. N. Nicholson & D. J. Rainford (Eds.) *Aviation Medicine* (3rd ed., pp. 419-454), Oxford, UK: Butterworth Heinemann.

Berry, D. C. & Broadbent, D. E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology, 79*, 251-272.

Berry, D. C. & Dienes, Z. (1993). *Implicit Learning: Theoretical and Empirical Issues*. Hillsdale: Lawrence Erlbaum Associates.

Broadbent, D. E. (1984). The Maltese cross: A new simplistic model for memory. *Behavioural and Brain Sciences, 7*, 55-94.

Cleeremans, A., Destrebecqz, A. & Boyer, M. (1988). Implicit Learning: News from the front. *Trends in Cognitive Sciences, 2*, 406–416.

Fitts, P. M. & Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks/Cole0.

Hardy, L., Mullen, R. & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of Psychology, 87*, 621-636.

Hasbrook, A. H. (1979). *Anatomy of landing: Cue by cue* (Tech. Rep. No. FAA-P-8740-26). Washington, DC: U.S. Department of Transportation.

Hayes, N. A. & Broadbent, D. E. (1988). Two modes of learning for interactive tasks. *Cognition, 28*, 249-276.

Knowles, M. (1975). *Self-Directed Learning*. Chicago: Follet.

Maxwell, J. P., Masters, R. S. W., Kerr, E. & Weedon, E. (2001). The implicit benefit of learning without errors. *The Quarterly journal of Experimental Psychology, 54A* (4), 1049-1068.

Nagel, D. C. (1988). Human error in aviation operations. In E. L. Weiner & D. C. Nagel (Eds.) *Human factors in aviation* (pp. 263-303), San Diego, CA: Academic Press.

Pashler, H. (1994). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance, 20*(2), 330-342.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *British Journal of Experimental Psychology*, 118, 219-235.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York, NY: Oxford University Press.

Rees, D. J. & Harris, D. (1995). Effectiveness of Ab Initio flight training using either linked or unlinked primary-axes flight controls. *The International Journal of Aviation Psychology, 3*(3), 291-304.

Rogers, C. (1969). *Freedom to Learn*. Columbus, OH: Merrill.

Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84* (2), 127-190.

Thom, T. (1992). *The pilot's manual flight training*. Frederick, MD: Center for Aviation Theory.

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

Zakay, D. & Shub, J. (1998). Concurrent Duration Production as a workload measure. *Ergonomics*, *41* (8), 1115-1128.

International Journal of Applied Aviation Studies, 4, Number 2

Training Development Papers

An Aviation Education Doctoral Program That Utilizes the Clinical Model to Enhance Research

Steven K. Marks

Oklahoma State University
Stillwater, Oklahoma 74078
Email: smarks@aesp.nasa.okstate.edu

and

John D. Vitek

Oklahoma State University
Stillwater, Oklahoma 74078
Email: jvitek@okstate.edu

Abstract

This article was designed to be a training development position paper. As such, it promotes the development of a training model that implements a new way of thinking and developing aviation education at comprehensive universities. As a position paper, formal operations that generate numbers and are supported by previous research studies are not the purpose. In this manner, a new innovative approach to the aviation education challenges of future can be discussed and critiqued along with our colleagues.

The authors describe a clinical model that would advance the aviation education discipline at comprehensive universities. The flight center is established as a clinic, similar to clinics in the health professions. It utilizes instructors to accomplish the significant role of flight training and minimizes the involvement of research faculty. The senior faculty in aviation programs, therefore, can be dedicated to research and training of graduate students. Because doc-

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

toral programs must be based in research, the nature of research is critical to program development. Examples of the six basic categories of research include basic, fundamental, exploratory, applied, programmatic, and industrial. Through such research the discipline of aviation can be established at comprehensive universities and strive to meet the demand of an industry that may be unlimited.

At the comprehensive university level, a differentiated faculty following a clinical model is one format that would succeed.

Introduction

This training development position paper proposes a model that aviation education research can prosper at the comprehensive university level. This new way of thinking promotes research at a level that has not been achieved to date.

At the research frontier, be that frontier the realm of outer space or the nature of airline travel between cities, the knowledge acquired from aerospace and aviation research is capable of changing human society. From experiments on flight in 1903 by the Wright brothers that proved human flight possible to the development of the International Space Station (Crouch, 1989), research experiments and tests in aviation during the last 100 years have changed how people move from place to place. A vast aviation industry has developed to facilitate travel, although the various forms of the airplane were also applied to military operations with deadly consequences. Every aspect of aviation requires research and education because people constantly strive to improve the quality of human experiences.

During the Wright brothers' experiments, it took days to sail from Europe to America but the Concorde, flying at supersonic speeds, routinely made the trip in hours. The Concorde was just one part of a complex economic/social empire designed to facilitate movement between places. The cost of maintaining this elite machine, however, contributed to its permanent removal from service in 2003. Aviation is a complex realm that must rely on educated citizens to continue the progress of the last century. What the future holds will be determined through research. Clearly, aviation education is an essential part of the industry that relies on technology for every aspect of moving people through the air from city to city. The purpose of this paper is to describe an aviation education program and document the types of research that will help the discipline achieve academic stature at comprehensive universities. Aviation education is an emerging discipline that now is recognized as more than pilot training (Truitt & Kaps, 1995, Johnson & Lehrer, 1995). "The sky is the limit" for aviation education when research is an integral part of the education process that leads to a doctoral degree.

Degree Programs in Aviation Education

As comprehensive universities develop doctoral programs in aviation, research acquires a level of importance for faculty to attain professional ranks and ultimately tenure. Contrast this goal with the primary objective of most undergraduate students in an aviation program – flight training (Karp, 2000). Because flight training rapidly moves from the classroom into individualized instruction in small aircraft, management of faculty resources will be critical for the successful establishment of a doctoral program. By using a clinical model for the flight program, the department can differentiate among faculty, thereby providing the critical element of time for faculty involved with research and the doctoral program. By clinical model, we mean that aviation should be patterned after a program such as the speech communication and disorder clinics in which people with speech disorders can be treated by professional staff that may or may not be active in the doctoral program of the discipline.

The speech communication and disorder clinical model is so successful, the American Speech and Hearing Association has incorporated clinical operations as part of their standard for accreditation of Graduate Education Programs in Audiology and Speech - Language Pathology (ASAA, 2004). Standard 3.0 identifies the curriculum (academic and clinical education) that is consistent with the program standards for entry into the professional practice. The program must offer appropriate courses and clinical experience on a regular basis so that students enrolled in the program may satisfy the requirements for a degree for entry into professional practice. The standard continues to identify the minimum contact hours necessary for each category and level of instruction. The FAA Practical Test Standards are a mirror image of this clinical format (FAA, 2004).

Upon obtaining a bachelor's degree in aviation with an emphasis on flight training, a graduate often lacks the necessary flight hours to become a professional pilot. By encouraging these students to enter a master's degree program and serve as flight instructors for the undergraduates, faculty can acquire time for research in contrast to spending hours in flight training. The MS students earn flight time as instructors, which helps them attain their personal goals of becoming professional pilots. Whereas some classroom instruction may be required of the faculty for the large number of students seeking flight training, the staff who run the clinic (i.e. flight training) assume the teaching burden from the faculty and generally are not involved in doctoral research. In another example, resident physicians in hospitals see patients to learn how to recognize and treat symptoms but they also relieve the faculty physicians from these tasks, which allow the faculty to perform research and enhance the reputation of the program.

The key to advancing the discipline is developing faculty who are active in research. Comprehensive universities expect all faculty to seek promotions and tenure and to have active research programs. At regional schools with aviation programs, research is generally not part of the reward system because of the large teaching load assigned to each faculty.As senior faculty develop at the comprehensive universities, however, a clear distinction among various programs will be possible. Although every flight training operation must meet minimum standards of the Federal Aviation Administration (FAA), departments with active research programs should be able to contribute higher standards to all aspects of aviation education.

How should programs be structured to enhance research in aviation? Johnson (1997) proposed the term Aeronology as the non-engineering aeronautical/ aerospace sciences degree that would use the research/practitioner model as the template for developing degree programs. To date, however, that term has not been embraced but doctoral programs in aviation are expanding. We believe that comprehensive universities need to adopt a clinical model to support emerging doctoral programs.

Doctoral Programs at Comprehensive Universities

Most state higher education institutions are classified into the three groups:

I. Comprehensive Graduate Universities: The functions of the comprehensive graduate universities include (1) lower-division and upper-division undergraduate study in a number of fields leading to the bachelor's or first professional degree; (2) graduate study in several fields of advanced learning leading to the master's degree; (3) graduate study in selected fields leading toward the doctor's degree; (4) Organized basic and applied research; (5) statewide programs of extension study and public service; (6) statewide programs designed to promote the economic development of the state, and (7) to the extent resources are available, to carry out limited programs and projects on a national and international scale.

II. Regional Universities: The functions of the regional universities include (1) lower-division and upper-division undergraduate study in several fields leading to the bachelor's degree; (2) a limited number of programs leading toward the first-professional degree; (3) graduate study below the doctor's level; (4) extension and public service responsibilities in the geographic regions in which they are located, and (5) responsibility for regional programs of economic development.

III. Two-Year Colleges: The function of the two-year college is (1) to provide general education for all students; (2) to provide education in several basic fields of university-parallel study for those students who plan to transfer and to a bachelor's degree granting university; (3) to provide one- and two-year programs of technical and occupational education that prepare individuals to enter the labor market; (4) to provide programs of remedial and developmental education for those whose previous education may not have prepared them for college, and (5) to provide formal and informal programs of study especially designed for adults and out-of-school youth to serve the community generally with continuing education opportunity.

The clear distinction among various colleges and universities, in terms of mission and objectives, should encourage professionals to acquire the research skills necessary to function at comprehensive universities. With comprehensive universities, however, the Department of Aviation must function differently than most departments. A clinical model is needed to promote program growth and development.

Doctoral degree programs in aviation education should be offered at the comprehensive university level. These institutions need the health clinic model to meet the demands of the flight training programs, as well as the research requirements of the doctoral degree. The flight center would be the clinic. At the airport, the clinic would be staffed and operated to prepare professional pilots and lead to the completion of the bachelor's degree. The students would meet all FAA Part 141 requirements. Run by a professional flight manager, the faculty would be at the master's level. The faculty rank of instructor would be appropriate because the terminal degree in the aviation field is the master's degree, although demand for faculty with doctoral degrees is growing.

Most universities define the instructor to be a person who should have earned a master's degree in his or her field and should have professional skills and expertise needed in the discipline. Such expertise should be certified by the professional organization of the discipline, or certifying agency. An instructor also demonstrates excellent performance in teaching and other assigned duties. The record of the instructor is maintained in accordance with the profession.

The remaining faculty that support the professional pilot degree outside the clinic are assistant professors, associate professors, and/or professor level faculty. The assistant professor would have earned the doctor's degree, shown the ability to exhibit the potential to grow in an academic career in accordance with the mission of the institution and the objectives of the academic unit. The associate professor would demonstrate that he/she is an accomplished teacher, and has a significant record of scholarly work in teaching, research, and service to the profession. The professor rank designates that a person's academic achievement merits recognition as a distinguished authority in his/her field. Professional colleagues, within the university and nationally, recognize the professor for their contributions to the discipline. A professor is an outstanding member of the academic community and sustains excellent performance in teaching, research, and service to the profession.

The transition from a flight-training program to a comprehensive degree program requires faculty to be dedicated to research. As a relatively new discipline, a variety of research is needed to advance the discipline. The following section on research differentiates among the types of research possible in aviation.

The Nature of Research

The acquisition of knowledge can occur in many ways, including well-designed experiments, through various experiences, or accidentally. Knowledge does not necessarily arise from speculation but requires evidence as the thoughts about the nature of the surface of the moon proved. When Armstrong and Aldrin landed on the moon in 1969, they acquired knowledge of the surface and were able to dispel all of the myths generated for millennia about the nature of the surface. Once new knowledge has been generated, the next step is communicating that knowledge to others, i.e., the process of education. It is not practical for everyone go to the moon to learn about the surface, therefore, education minimizes the cost of acquiring knowledge and the risks associated with acquiring evidence. Unfortunately, acquiring any new knowledge has increased in cost and complexity.

Education involves helping others acquire the knowledge that they need to function in society. Figure 1, a diagram of known versus unknown knowledge, is a conceptual perspective of the research frontier (modified from Marks, Vitek, Giardina, & McQueen, 2002). People are taught what is known but that information is not static because new knowledge is constantly being acquired through research along the frontier. Integrating this knowledge into education is necessary to continue the advancement of the research frontier. Unlike the scholars of 100 years ago who had a smaller knowledge base but had a greater breadth of knowledge, scholars today become very specialized in narrowly defined disciplines. The Wright brothers, for example, were craftsmen who experimented with a powered machine and proved people could use machines to fly. In contrast, a pilot today is a specialist in the systems of a particular aircraft with regard to operational procedures for takeoff, flight, landing, and emergencies. The procedures will vary based upon the system in each aircraft; hence, education and flight simulation are used to prepare the pilot to fly each type of plane. The ability to approximate reality in flight simulators provides a great education while minimizing the risk and costs associated with learning a flight system.

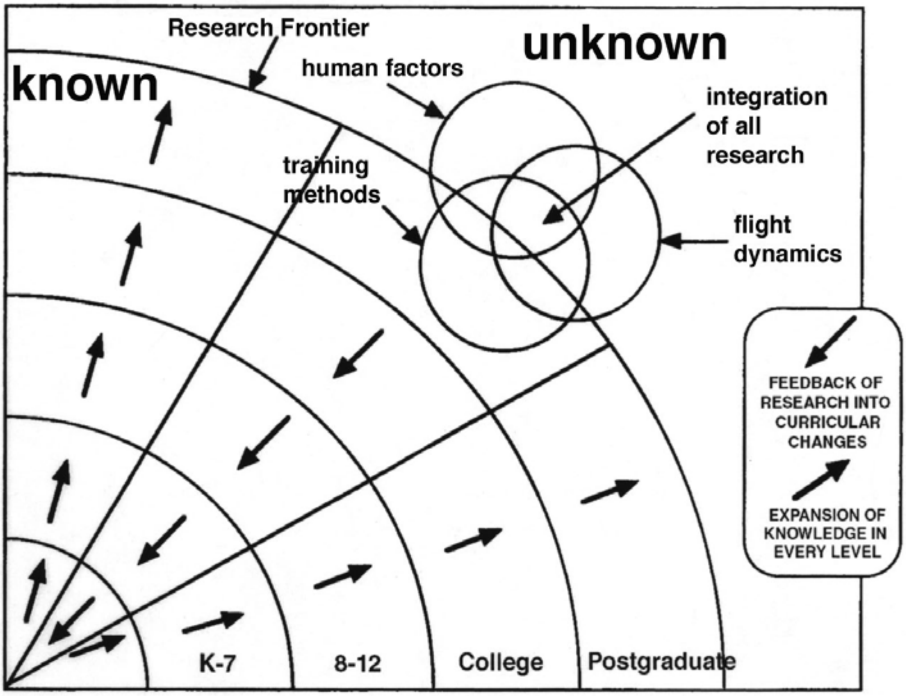


Figure 1. Known versus unknown knowledge.

Within education, faculty must perform research and publish to advance the knowledge base. Truitt and Kaps (1995) reviewed the status of scholarly journals in which aviation research can be published. When they prepared their article, computer searches were just beginning and finding existing literature was difficult. Whereas the literature remains scattered in a variety of journals, library search engines and numerous databases now facilitate acquisition of the knowledge as it is being published. Easy access to published literature is necessary because the knowledge base forms the common theoretical and conceptual bases of the discipline. Faculty contributes to the development of these bases and students must understand the foundations of the discipline from which they will generate new knowledge through research.

Whether the doctoral program is an EdD or PhD, research must be completed. Dill and Morrison (1985) stated these degrees might be similar. Most graduate colleges, however, restrict the amount of research performed for the EdD whereas few limits are placed on the research hours for the PhD. The nature of research is often the key because pure, applied and literacy objectives form the basis for the PhD, and the EdD is more related to application. In an effort to describe research available to fledgling aviation programs, we have modified the research classification used by NASA Communicating NASA's Knowledge (NASA, 1998) to describe suitable topics, which could form the basis for advancing aviation education as a discipline.

NASA has played a major role in aerospace education for at least 45 years. At some point, education specialists were hired to translate the new technical knowledge into education information and activities meaningful to students in K-16+. Why? A pipeline of future scientists is needed to continually expand the research frontier. What type of research is performed at the frontier and how this knowledge can contribute to aviation education and the development of professionals will be considered in the remainder of this paper.

Categories of Research

According to Communicating NASA's Knowledge (NASA, 1998), six categories of research exist for the expressed purpose of acquiring knowledge. *Basic, fundamental, exploratory, applied, programmatic, and industrial* are categories in which to aggregate the knowledge being generated. Examples are provided for each category of research to clarify the function and explain how each can enhance doctoral programs in aviation education.

Basic. Basic research seeks knowledge for the sake of that knowledge. Within the realm of the aviation/aerospace discipline, the best example may be the Hubble telescope. From its vantage point in space, data are captured without being influenced by the atmosphere. But what is the purpose of the data? Astronomers are seeking answers to many questions, such as when was the universe formed, or do black holes exist? How fast are galaxies moving? Does life exist on other planets? Whether we learn the answers to these questions or the numerous questions raised during analyses of the data will probably have little if any impact on human activity. Only our imagination limits the aims of basic research. In some circles, the knowledge may simply be called "trivia." This basic research, however, may have value in the future as other changes occur. Because we cannot predict what bit of new knowledge will initiate the next revolutionary change in the quality of life, basic research is considered a good investment. Basic research in aviation education may consist of new techniques of teaching flight training or the role of composite materials in future aircraft design.

Fundamental. Fundamental research seeks useful knowledge, such as the characteristics of the atmosphere. As cold and warm air masses mix, tornadoes may form. Predicting when and where these destructive winds might occur can save lives with advanced warnings. In another example, the location of the jet stream impacts the routes airplanes fly. Headwinds or tailwinds affect fuel usage, and therefore, where to fly to use or avoid these winds helps airlines save money. Along the flight route of every aircraft, atmospheric data are collected and transmitted to a central point to be shared with all other airplanes. Whereas we need fundamental knowledge of the atmosphere, its vast size and rapidly changing conditions present obstacles to easy solutions to understanding the forces at work. Within aviation education, fundamental research may enhance how learning occurs such that critical information is understood and remembered. What is the best method for training pilots about the complex atmospheric conditions they will face daily? Educators can work directly with scientists to devise programs to maximize learning.

Exploratory. Exploratory research attempts to identify perceived useful knowledge. Within the technique of remote sensing, scientists experiment with sensors to collect reflected or emitted energy at a scale that permits detailed analysis of the surface. Can the mineral content of bare rock be determined with a level of accuracy to locate ores that can be mined at a profit? One goal of using remote sensing is to save people from the tedious effort of field research. The techniques employed are often reliable for other purposes and through experimentation extended into new areas of knowledge acquisition. Radar, for example, was developed to detect enemy aircraft, but it has been vastly improved to interpret weather phenomena (such as precipitation and wind characteristics). In aviation education, methods of how to teach information can be tested to learn what works the best for the circumstances.

Applied. Applied research pursues practical objectives, such as experiments on the space station that are sent directly into K-12 classrooms. Applied research can have benefits that extend beyond the actual research. The psychological impact on students who watch experiments from space may be difficult to measure. Have these students been impressed by the experience and will it have a lasting impact on their choice of a career? How and when such decisions are made are difficult to document. If students do opt to pursue careers related to the space program, they would be fulfilling a hope that is embedded in the concept of education from space. The knowledge created by activities in space is useful, but motivating students toward science is also a goal of applied research.

Programmatic. Programmatic research seeks and provides knowledge for a mission. A question is posed and research is directed at answering that specific question, such as what sensors are capable of detected buried ice? The proposed research, however, may also be useful in related efforts such as developing the next generation of Space Shuttle. Educators might address the curricula of the aviation/aerospace programs to determine which topics are required versus those that should be electives.

Industrial. Industrial research attempts to achieve economic benefits, such as interpreting geologic structure to enhance mineral exploration. Many other examples can be given when one considers all the products in use today that were originally developed in support of space exploration. If such products can be developed for mass consumption, great profits are possible for the industries involved. Universities hope that some faculty research can generate profits through patents and licensing and thereby help underwrite research in a variety of disciplines. A highly successful training program could be licensed and result in resources for the university, program and faculty.

The nature of research influences the knowledge created and how aviation education is enhanced. However, aviation education is a relatively new discipline. In reviewing the 2003 University Aviation Association (UAA) Institutional Membership, 11% would be classified as comprehensive universities that offer students the opportunity to earn a doctoral degree. Fifty-eight percent of the members offer the Masters or Bachelor degrees and 31% are Community College or Technical Schools. When the terminal degree is offered as a Doctorate in Education (Ed. D.), the amount of research required for the dissertation is ten credits. This is significantly less than the minimum of 15 credits required for the dissertation in a Ph.D. program. Few comprehensive programs also mean fewer opportunities for people who have a doctoral degree in aviation and are seeking a university career. Aviation programs must be upgraded through research to enhance program quality and with the clinic format for flight training to provide faculty with time to perform research.

Conclusions

The key to advancing the discipline of aviation education is to provide aviation education with doctoral programs at comprehensive universities. Comprehensive higher education institutions are designed to promote the necessary research and involve students seeking the doctoral degree.

At the comprehensive university level, a differentiated faculty, following a clinical model, is one format that would succeed. The flight center would be the clinic that would lead to the bachelor degree. The remaining faculty would be assistant professors, associate professors, and/or professor level faculty. The transition from a flight-training program to a comprehensive degree program requires faculty to be dedicated to research.

In aviation education, a variety of research is welcome and necessary. This includes basics, fundamental, exploratory, applied, programmatic, and industrial research. The sky is the limit when research based aviation education doctoral programs are offered at comprehensive higher education universities. The clinical model would lead to a successful doctoral program.

References

American Speech-Language-Hearing Association (ASHA) (1999, January). *Standards for Accreditation of Graduate Education Programs in Audiology and Speech-Language Pathology.* Retrieved April 2, 2004, from the ASHA Web site: http://www.asha.org/about/ academia/accreditation/standards.htm.

Couch, T. D. (1989). *The bishop's boys: a life of Wilbur and Orville Wright*. W. W . Norton, New York.

Dill, D. D. & Morrison, J. L. (1985). EdD and PhD research training in the field of higher education: a survey and proposal. *The Review of Higher Education, 8* (2) 169-186.

Johnson, J. A. & Lehrer, H. R. (1995). The feasibility of developing a non-engineering aeronautical/aerospace science doctoral degree program in U. S. universities. *Journal of Studies in Technical Careers, XV* (4) 245-255.

Johnson, J. A. (1997). Curriculum design issues in developing a doctor of philosophy program in aeronology. *Journal of Air Transportation World Wide*, *2* (1) 81-92.

Johnson, J. A. (1999). An examination of the U. S. collegiate aviation workforce in preparing the next generation of aviation faculty members beyond 2000. *Collegiate Aviation Review, 17* (1) 31-39.

Karp, M. R. (2000). University aviation education: an integrated model. *Collegiate Aviation Review*, *18* (1) 1-11.

Marks, S. K., Vitek, J. D., Giardino, J. R., & McQueen, K. C. (2002). Creating curricular change: needs in grades 8-12 earth science. *Geomorphology*, *47* (2-4) 261-273.

National Aeronautics and Space Administration. (1998). *Communicating NASA's knowledge: A report of the communicate knowledge process team.* Executive office of the President, Office of Science and Technology Policy.

Truitt, L. J. & Kaps, R. W. (1995). Publishing aviation research: an interdisciplinary review of scholarly journals. *Journal of Studies in Technical Careers*, *XV* (4) 229-243.

U.S. Department of Transportation, Federal Aviation Administration. Practical Test Standards for Airplane. *Flight Standards Service*, Washington, D.C. 20591

International Journal of Applied Aviation Studies, 4, Number 2

Evaluation of Student Satisfaction with Flight Training Schools

Gregory L. Schwab

Aerospace Technology
Indiana State University
Terre Haute, IN 47809
(812)237-2641
gschwab@indstate.edu

Abstract

In the Department of Aerospace Technology at Indiana State University, many students complained to the faculty about their dissatisfaction with the two contracted flight schools' training programs. This study evaluated students' satisfaction with the flight schools. The Flight School Evaluation Survey was administered to all 66 students enrolled in the flight school program. Results showed that the quality of service provided by the two flight schools appeared generally adequate. However, survey results also showed that the flight schools demonstrated a lack of openness to improvement, some inadequacy of aircraft maintenance, and delays or denials of training. Recommendations to institute more consistent delivery of flight school training included regular departmental reviews of the program and institution of improvements; development of ground-based flight simulators for students, who experience flight training delays; periodic meetings of faculty, students, and flight school contractors; end-of-term student satisfaction surveys as part of the curriculum; and comparative surveys with students at both flight schools to assess each school's performance and follow-up with further improvements, if warranted.

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

Introduction

Nature of the Problem

At Indiana State University (ISU), the Department of Aerospace Technology (AST) faculty received numerous complaints about the students' flight training. This is outsourced under university contract to private flight training schools authorized by the Federal Aviation Administration (FAA) to conduct flight training under FAA Regulation Part 61, referred to as Part 61 in this paper. Part 61 prescribes the requirements for conducting flight training, including the issuance of pilot certificates and ratings, and, as related to this study, the conditions under which those authorizations are necessary (FAA, 2003).

Anecdotal evidence at ISU revealed that many students were dissatisfied with the service delivered by the university's two contracted flight schools, both of which operated at local airports. Students' complaints encompassed a variety of areas, such as maintenance delays and cancellations, nonstandard training procedures, and poor appearance of the flight schools' facilities. For confidentiality purposes, the two flight schools are identified throughout as Flight School A and Flight School B.

Purpose of the Study

The purpose of this study was to evaluate the two flight schools' quality of service through student feedback. Student satisfaction was determined by completion of the researcher-designed Flight School Evaluation Survey (FSES) (see Appendix). The results of the survey also assisted the ISU Department of Aerospace Technology to make decisions about the continued use of these flight schools.

Literature Review

Literature in three pertinent areas was examined. These were (a) customer satisfaction, (b) the student-flight school relationship, and (c) studies of flight training programs.

Customer Satisfaction

Customer satisfaction may be defined as the extent to which the buyer's expectations are met by the perceived performance of a product. Kotler and Armstrong (2001) explained, "If the product's performance falls short of expectations the buyer is dissatisfied. If performance matches or exceeds expectations, the buyer is satisfied or delighted" (p. 9).

Bellchambers and Neun (2001) claimed that customer satisfaction and quality are interrelated and essential to a company's reputation. Company reputation directly affects the "bottom line" of financial well-being. Companies should instill customers with confidence that the final product will meet the customers' requirements, be delivered correctly the first time, and not necessitate follow-up. In the case of flight training, it is crucial that the final product, the training itself, be delivered correctly the first time for obvious reasons of safety and survival.

Customer feedback is a key indicator for maintaining a quality management system for the company and improving future performance. Feedback, which can be gathered through surveys or interviews, should be solicited throughout the customer experience (Kotler & Armstrong, 2001). For service companies, quality and reliability require an ongoing effort (Yu, 2001). Bond and Fink (2001) assert that companies should allocate more resources to measuring customer satisfaction levels.

It is in the context of customer satisfaction that the present study was undertaken. Prior to this study, no formal measures of consumer satisfaction had been implemented for the specialized content of the ISU flight program.

Student-Flight School Relationships

The most common relationship a student experiences in flight training is with the flight instructor. This one-on-one relationship is critical for successful flight training to take place (Downs, 2000). Both the instructor and student have specific responsibilities that contribute to the success of training. Hiner (2000) identified several instructor habits that may hinder effective student flight training, such as frequent lateness, arriving unprepared for flight, shouting at students, not letting students fly the plane, and not listening to students. In such cases, where the instructor is not conducting the training properly, students' in-flight situational awareness could be compromised and the flight placed at risk (Davisson, 2000).

In a study of 106 flight training programs nationwide, Bryan (1996) found that students whose training was delayed for any reason, such as weather or maintenance delays or cancellations, and who failed to complete flight courses in timely fashion lost the effectiveness of concurrent lecture/laboratory courses. Bryan observed, "…such delays can lead to ever-increasing costs for the students. In addition, failure to complete their academic programs impedes their employment opportunities and interrupts their flight career progression" (p. 63).

Aspects other than instructors' inappropriate behavior also may be sources of complaint, such as uncooperative staff with whom students interact at the aircraft checkout counter. Flight school policies may affect students more punitively than "walk-in" customers—students may be required to obtain special permission from the flight schools to take passengers on flights, whereas other similarly certificated pilots using the same facilities have no such restrictions (Wurman, 2000). Such policies may result in students' feeling that they are unappreciated and their concerns are ignored, in comparison with nonstudent paying customers.

In relation to flight school policies, Wurman (2000) offered excellent advice. He suggested that business owners and managers discard outdated policies, remind instructors that the flight school is serving students, that institutions practice flexibility, teach the "why" behind policies that must be retained, and begin on the "right foot" with new students.

Studies of Flight Training Programs

Best practices for flight student recruitment were studied by Bowen, Carstenson, and Hansen (1999). These researchers examined 20 collegiate flight programs throughout the United States, with emphasis on recruitment of students from other same-campus university departments. Results demonstrated that good student recruitment develops from an active and ongoing commitment to student needs, as well as excellence in the classroom and flight training by the faculty and staff of the flight schools.

Luedtke and Papazafiropoulos (1996) examined retention issues as they relate to collegiate aviation. The researchers surveyed 102 current and former students at a flight-training program in a medium-sized state university in Nebraska. Luedtke and Papazafiropoulos found a direct relationship between student satisfaction and retention in collegiate aviation programs. The researchers concluded, "Just as in business, if customers do not receive the help and attention needed they will take their business elsewhere" (p. 46).

In summary, the literature supported the perspective that flight schools consider university students as "customers" and endeavor to please them at the risk of losing their business. Too often, the flight schools perceive students as less important than other customers. Informal student feedback to this effect led to the present study.

Methodology

Survey Development

A cross-sectional survey design was used for this study, conducted in the spring of 2001. The survey was developed with reference to the literature, Federal Aviation Regulation (FAR) Part 61, and input in aviation and education from ISU and industry representatives. Committees comprised of three department faculty members and three industry representatives were utilized for the development and drafting of the survey.The 25-item survey was divided into 4 sections; student demographics, training methods and procedures, aircraft maintenance issues, and flight school facilities.

Field-testing took place with a sample of 7 students, 10% of the total 66 students enrolled in the ISU flight-training program. After revisions for clarity and consistency, the final survey, the FSES, was administered by the researcher and an assistant to all 66 students (100% response rate) during a regularly scheduled classroom period.

The appendix reproduces the survey. Section A, Student Demographics, contains five items (1-5) asking students their class rank, how long they had been in training, how many flight hours they had completed, type of flight certificate held, and number of other flight schools attended. Section B, Comparison With Other Flight Schools and Progress, contains two items (6-7), asking how students felt about the ISU flight school in comparison to the flight schools previously attended, and if they felt their training was progressing as they thought it should. Section C, Training Methods and Procedures contains six items (8-13) asking, for example, if the flight school was open to suggestions for improvement, if the students felt safe flying with the flight school staff, and if the flight school had clearly stated training objectives.

Section D, Flight Instructors, contains eight items (14-21), asking, for example, if students felt that their instructors utilized training time well, if instructors were readily available for consultations, and if instructors informed them how they would be evaluated on each flight. Section E, Aircraft Maintenance and Facilities, contains four items (22-25), asking if students had been delayed or denied flight training on two or more occasions in the last 30 days due to a aircraft maintenance problem with an aircraft, if students felt the aircraft were adequately maintained, if students felt safe flying the flight school aircraft, and if the flight school facilities were well maintained.

Limitations

Three limitations were evident in this study. First, the study was not designed for application to other institutions with flight training programs. Thus, this study can be generalized only to settings highly similar to that at ISU. Second, although the selection of survey items was based on informed recommendations of the expert committees and the researcher, other relevant areas may have been omitted which could affect the results, such as students' personal flight experience. Third, the study sample was a relatively small convenience sample and the results may not be representative of other university aviation programs.

Results

Demographic Characteristics

Data analysis showed that of the 66 students, 39.40% (*n* = 26) attended Flight School A and 60.60% (*n* = 40) attended Flight School B. Of the total sample, 61 students were White, 93%; two were Hispanic, 1.5%; one was African American, 1%; and two were Japanese, 1.5%. A total of 61 were male, 93%, and five were female, 8%. All students lived on or near campus; all were within the traditional undergraduate age range from 18 to 21. None were career changers, part-time students, or commuters.

Credit hours and flight experience. Students' responses to Section A, Student Demographics, are shown in Tables 1- 4. Table 1 shows that close to half of the students at Flight School A had completed enough credit hours to have junior status, 42.30% (*n* = 11). However, the students at Flight School B were divided almost equally on number of credit hours: freshman, 22.50% (*n* = 9); sophomore, 25.00% (*n* = 10); junior, 25.00% (*n* = 10); senior, 27.50% (*n* = 11).

Table 1

Number of University Credit Hours Completed

	Facility				
	Flight School A			Flight School B	
Item 1: Credit Hours Completed	*n*	%		*n*	%
Freshman (0-31)	6	23.08		9	22.50
Sophomore (32-62)	3	11.54		10	25.00
Junior (63-93)	11	42.30		10	25.00
Senior (94 and above)	6	23.08		11	27.50

Note. Numbers in parentheses indicate number of credit hours completed.

The number of years in flight training is shown in Table 2. Over three-fourths, or 88.00%, of the Flight School A students had been in flight school for less than 3 years. A total of 32.00% ($n = 8$) had been in flight school less than 1 year and 28.00% each ($n = 7$) between 1 and 2 years and between 2 and 3 years.

The number of years in flight training for Flight School B students was similar, with 77.50% in the first three categories: 35.00% ($n = 14$) had less than 1 year, 20.00% ($n = 8$) had more than 1 year and not more than 2, and 22.50% ($n = 9$) had 2 years but less than 3. Also similar to Flight School A, 35.00% ($n = 14$) of the Flight School B students had less than 1 year of flight training.

The number of flight hours completed is shown in Table 3. Over a third of the Flight School A students, 38.46% ($n = 10$), had completed from 101-200 flight hours, compared to 22.50% ($n = 9$) of the Flight School B students. Moreover, approximately a third of the Flight School A students, 34.62 % ($n = 9$), had less than 100 hours, compared to close to half, 42.50% ($n = 17$), of the Flight School B students.

The number of flight certificates held (item 4) and other flight schools attended (item 5) are displayed in Table 4. All respondents held some form of flight certificates with private certificate the highest for Flight School A students, 44.00% ($n = 11$). Student and private certificates ranked similarly at Flight School A , 30.00% ($n = 17$) and Flight School B, 32.50% ($n = 25$). Although at Flight School B 17.50% ($n = 7$) were instrument certified, the other responses to this item for instrument, commercial, or CFI at both schools were between 12.00% ($n = 3$) and 8.00% ($n = 2$).

The number of other flight schools the students had attended (item 5) shows that Flight School A students had attended many other flight schools: 20.00% ($n = 5$) none, 36.00% ($n = 9$) one, 24.00% ($n = 6$) two, and 20.00% ($n = 5$) three or more. However, at Flight School B, almost half, 47.50% (*n* = 19), had not attended another flight school, and slightly less, 42.50% (*n* = 17), had attended one other flight school. Only two students each, 5.00% (*n* = 2), had attended two or three or more other flight schools.

Table 2
Number of Years in Flight Training

	Facility				
	Flight School A			Flight School B	
Item 2: Years in Flight Training	*n*	%		*n*	%
Less than 1 year	8	32.00		14	35.00
More than 1 year/ but not more	7	28.00		8	20.00
Two years but less than 3	7	28.00		9	22.50
Three years but less than 4	2	8.00		7	17.50
Four or more years	1	4.00		2	5.00

Note. Number of years in flight training is related to academic years. One student who attended Flight School A did not reply to this item.

Table 3
Number Flight Hours Completed

	Facility				
	Flight School A			Flight School B	
Item 3: Flight Hours Completed	*n*	%		*n*	%
Less than 100 hours	9	34.62		17	42.50
101 to 200 hours	10	38.46		9	22.50
201 to 300 hours	3	11.54		11	27.50
301 to 400 hours	0	0.00		0	0.00
More than 400 hours	4	15.38		3	7.50

Table 4

Flight Certificates and Other Flight Schools Attended

	Facility				
	Flight School A			Flight School B	
Item	*n*	%		*n*	%
Item 4: Flight Certificates Held					
Student certificate	6	24.00		12	30.00
Private	11	44.00		13	32.50
Instrument	3	12.00		7	17.50
Commercial	3	12.00		4	10.00
Certified Flight Instructor (CFI)	2	8.00		4	10.00
Item 5: Other Flight Schools Attended					
None	5	20.00		19	47.50
One	9	36.00		17	42.50
Two	6	24.00		2	5.00
Three or more	5	20.00		2	5.00

Note. One student at Flight School A did not reply to items 4 and 5.

Comparison with other flight school programs. Responses to Section B, Comparison with Other Flight Schools Programs (items 6-7), are shown in Table 5. Students' ratings of other flight schools attended in comparison with their present one showed a wide variation (item 6). As can be seen, Flight School A students' responses were divided fairly equally: worse, 30.77% (*n* = 8); about the same, 26.93% (*n* = 7); and better, 23.08% (*n* = 6). Only 19.22% (*n* = 5) had not attended another flight school.

As indicated in Table 5, almost a third of the Flight School A students, 30.77% (*n* = 8), considered their present school worse than others they had attended. In contrast, 10.26% (*n* = 4) of the Flight School B students rated their present school worse than previous ones. Only 5.13% (*n* = 2) rated their present school about the same, and 15.38% (*n* = 6) rated it better. However, compared to 19.22% (*n* = 5) of the Flight School A students, close to three-fourths of the Flight School B students, 69.23% (*n* = 27), had not attended another flight school.

Table 5

ISU Flight Schools Compared to Other Flight Schools

	Facility				
	Flight School A		Flight School B		
Item	*n*	%		*n*	%
Item 6: Comparison of School Ratings					
Worse	8	30.77		4	10.26
About the same	7	26.93		2	5.13
Better	6	23.08		6	15.38
Had not attended another flight school	5	19.22		27	69.23
Item 7: Flight Training Progress					
Behind where I thought I would be	19	73.09		23	58.97
About where I thought I would be	5	19.22		14	35.90
Ahead of where I thought I would be	2	7.69		2	5.13

Note. One student from Flight School B did not reply to items 6 and 7.

Table 5 also shows students' perception of their progress in flight training (item 7). Most of the Flight School A students, 73.09% (*n* = 19), believed they were behind where they thought they would be; 19.22% (*n* = 5) believed they were where they thought they would be; and 7.69% (*n* = 2) believed they were ahead of where they thought they would be.

The majority of Flight School B students, 58.97% (*n* = 23), believed they were behind where they thought they would be. However, more Flight School B students than Flight School A students believed they were where they thought they would be, 35.90% (*n* = 14), and only 5.13% (*n* = 2) believed they were ahead of where they thought they would be.

Training methods and procedures. Flight School A students' responses to Section C, Training Methods and Procedures (items 8-13), are shown in Table 6. In a combined percentage, over half the Flight School A students, 53.84%, either strongly disagreed or disagreed (both 26.92%, *n* = 7) that the flight school was open to student suggestions for improvement (item 8). Again in a combined percentage, over three-fourths, 76.92%, agreed (65.38%, *n* = 17) or strongly agreed (11.54%, *n* = 3) that the flight school staff was visible on a daily basis to help with training (item 9). Over half, 53.85%, agreed (46.16%, *n* = 12) or strongly agreed (7.69%, *n* = 2) that the school provided adequate instruction for their preferred tempo of training (item 10). Similarly, half, 50.00%, agreed (42.31%, *n* = 11) or strongly agreed (7.69%, *n* = 2) that they would recommend the school to others (item 11), although 19.23% (*n* = 5) strongly disagreed and 11.54% (*n* = 3) disagreed on recommending the school.

Table 6
Flight School A: Training Methods and Procedures

	Rating									
	Strongly Disagree		Disagree		No Opinion		Agree		Strongly Agree	
Item	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
8. Open to improvement	7	26.92	7	26.92	5	19.23	6	23.08	1	3.85
9. Staff visible and assists	1	3.85	3	11.54	2	7.69	17	65.38	3	11.54
10. Adequate instruction	0	0.00	7	26.92	5	19.23	12	46.16	2	7.69
11. Would recommend to others	5	19.23	3	11.54	5	19.23	11	42.31	2	7.69
12. Feel safe flying with staff	0	0.00	1	3.85	4	15.37	14	53.85	7	26.93
13. Clearly stated objectives	3	11.54	7	26.92	8	30.77	5	19.23	3	11.54

Note. Two students did not respond to items 10 and 12.

With regard to feeling safe flying with the flight school staff (item 12), four-fifths of the students agreed (53.85%, *n* = 14) or strongly agreed (26.93%, *n* = 7) that they felt safe. Concerning whether the school provided clearly stated training objectives (item 13), 30.77% agreed (19.23%, *n* = 5) or strongly agreed (11.54%, *n* = 3) that it had. However, 26.92% (*n* = 7) had no opinion and disagreed.

Table 7 reports the Flight School B students' responses to Section C. Close to half the students, 45.00%, disagreed (30.00%, *n* = 12) or strongly disagreed (15.00%, *n* = 6) that the flight school was open to student suggestions for improvement (item 8). Varying percentages had no opinion (27.50%, *n* = 11) or agreed (25.00%, *n* = 10) or strongly agreed (2.50%, *n* = 1) on this issue. Four-fifths, 80.00%, agreed (60.00%, *n* = 24) or strongly agreed (20.00%, *n* = 8) that the flight school staff was visible on a daily basis to help with training (item 9). Three-fourths, 74.37%, agreed (58.99%, *n* = 23) or strongly agreed (25.38%, *n* = 6) that the school provided adequate instruction for their preferred tempo of training (item 10). Similarly, 71.80% agreed (41.03%, *n* = 16) or strongly agreed (30.77%, *n* = 12) that they would recommend the school to others, although 17.95% (*n* = 7) had no opinion (item 11).

Table 7

Flight School B: Training Methods and Procedures

	Rating									
	Strongly Disagree		Disagree		No Opinion		Agree		Strongly Agree	
Item	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
8. Open to improvement	6	15.00	12	30.00	11	27.50	10	25.00	1	2.50
9. Staff visible and assists	1	2.50	1	2.50	6	15.00	24	60.00	8	20.00
10. Adequate instruction	2	5.26	3	7.89	4	10.53	23	60.53	6	15.79
11. Would recommend to others	1	2.56	3	7.69	7	17.95	16	41.03	12	30.77
12. Feel safe flying with staff	0	0.00	0	0.00	3	7.89	12	31.58	23	60.53
13. Clearly stated objectives	2	5.26	6	15.79	9	23.68	16	42.11	5	13.16

Note. Two students did not respond to items 10 and 12.

With regard to feeling safe flying with the flight school staff (item 12), over four-fifths of the students, 92.11%, agreed (31.58%, *n* = 12) or strongly agreed (60.53%, *n* = 23) that they felt safe. Concerning whether the school had clearly stated training objectives (item 13), the majority, 55.27%, agreed (42.11%, *n* = 16) or strongly agreed (13.16%, *n* = 5) that it had. However, 23.68% (*n* = 9) had no opinion and 21.05% disagreed (15.79%, *n* = 6) or strongly disagreed (5.26%, *n* = 2).

Flight instructors. Section D of the survey, Flight Instructors (items 14-21), addressed respondents' opinions regarding their experiences with instructors. As Table 8 shows, two-thirds of the Flight School A students, 68.00%, agreed (44.00%, *n* = 11) or strongly agreed (24.00%, *n* = 6) that their instructors used the training time well (item 14). Slightly more students, 69.24%, agreed (46.16%, *n* = 12) or strongly agreed (23.08%, *n* = 6) that instructors were readily available for consultations (item 15). A higher percentage, four-fifths, 80.00%, agreed (64.00%, *n* = 16) or strongly agreed (16.00%, *n* = 4) that the instructors understood students' training needs (item 16). Almost two-thirds, 64.00%, agreed (48.00%, *n* = 12) or strongly agreed (16.00%, *n* = 4) that instructors were prepared for each session (item 17). However, 28.00% disagreed (20.00%, *n* = 5) or strongly disagreed (8.00%, *n* = 2) that instructors were prepared.

Table 8

Flight School A: Instructor Evaluations

	Rating									
Item	Strongly Disagree		Disagree		No Opinion		Agree		Strongly Agree	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
14. Used training time well	1	4.00	3	12.00	4	16.00	11	44.00	6	24.00
15. Available for consultations	2	7.69	2	7.69	4	15.38	12	46.16	6	23.08
16. Understood my training needs	1	4.00	1	4.00	3	12.00	16	64.00	4	16.00
17. Prepared for each session	2	8.00	5	20.00	2	8.00	12	48.00	4	16.00
18. Identified evaluation format	1	4.00	6	24.00	4	16.00	11	44.00	3	12.00
19. Summarized major points after flight	1	4.00	4	16.00	1	4.00	13	52.00	6	24.00
20. Stimulates interest in flying	1	4.00	4	16.00	5	16.00	9	36.00	7	28.00
21. Had two or more instructors in past 30 days	10	41.67	3	12.50	5	20.83	3	12.50	3	12.50

Note. One student did not answer items 14, 16, 17, 18, 19. Two students did not answer item 21

Regarding student evaluations by instructors, the majority, 56.00%, agreed (44.00%, *n* =11) or strongly agreed (12.00%, *n* = 3) that their instructors identified how they would be evaluated prior to flight (item 18). Three-fourths, 76.00%, agreed (52.00%, *n* = 13) or strongly agreed (24.00%, *n* = 6) that instructors summarized the major points after each flight (item 19). Almost two-thirds, 64.00%, agreed (36.00 %, *n* = 9) or strongly agreed (28.00%, *n* = 7) that the instruction stimulated their interest in flight training (item 20). The majority, 54.17%, strongly disagreed (41.67%, *n* = 10) or disagreed (12.50%, *n* = 3) that they had two or more instructors in the 30 days prior to the survey (item 21). One-fourth, 25.00%, agreed or strongly agreed that they did have two or more instructors (12.50%, *n* = 3 each).

As Table 9 shows, almost two-thirds of the Flight School B students, 65.79%, agreed (44.74%, *n* = 17) or strongly agreed (21.05%, *n* = 8) that their instructors used the training time well (item 14). Over four-fifths, 82.06%, agreed (61.55%, *n* = 24) or strongly agreed (20.51%, *n* = 5) that instructors were readily available for consultations (item 15). Almost the same percentage, 82.05%, agreed (56.41%, *n* = 22) or strongly agreed (25.64%, *n* = 10) that the instructors understood students' training needs (item 16). A similar percentage, 82.06%, agreed (64.11%, *n* = 25) or strongly agreed (17.95%, *n* = 7) that instructors were prepared for each session (item 17).

Table 9

Flight School B: Instructor Evaluations

	Rating									
	Strongly Disagree		Disagree		No Opinion		Agree		Strongly Agree	
Item	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
14. Used training time well	1	2.63	5	13.16	7	18.42	17	44.74	8	21.05
15. Available for consultations	1	2.56	1	2.56	4	12.82	24	61.55	8	20.51
16. Understood my training needs	2	5.13	2	5.13	3	7.69	22	56.41	10	25.64
17. Prepared for each session	1	2.56	3	7.69	3	7.69	25	64.11	7	17.95
18. Identified evaluation format	5	12.82	6	15.38	4	10.26	20	51.28	4	10.26
19. Summarized major points after flight	3	7.69	2	5.13	2	5.13	23	58.97	9	23.08
20. Stimulates interest in flying	1	2.56	4	10.26	4	10.26	18	46.15	12	30.77
21. Had two or more instructors in past 30 days	13	33.34	8	20.51	7	17.95	8	20.51	3	7.69

Note. One student did not respond to items 16, 17, 18, 19, 20, and 21. Two students did not respond to items 14 and 15.

Regarding student evaluations by instructors, close to two-thirds, 61.54%, agreed (51.28%, *n* = 20) or strongly agreed (10.26%, *n* = 4) that their instructors identified how they would be evaluated prior to flight (item 18). Over four-fifths, 82.05%, agreed (58.97%, *n* = 23) or strongly agreed (23.08%, *n* = 9) that instructors summarized the major points after each flight (item 19). Slightly over three-fourths, 76.92%, agreed (46.15%, *n* = 18) or strongly agreed (30.77%, *n* = 12) that the instruction stimulated their interest in flight training (item 20).

The majority, 53.85%, strongly disagreed (33.34%, *n* = 13) or disagreed (20.15%, *n* = 8) that they had two or more instructors in the 30 days prior to the survey (item 21). Slightly over one-fourth, 28.20%, agreed (20.51%, *n* = 8) or strongly agreed (7.69%, *n* = 3) that they did have two or more instructors.

Aircraft maintenance and facilities. Section E of the survey, Aircraft Maintenance and Facilities (items 22-25), addressed respondents' opinions regarding the flight schools' aircraft maintenance and facilities. As Table 10 shows, almost three-fourths, 73.08%, of the Flight School A students strongly agreed (65.39%, *n* = 17) or agreed (7.69%, *n* = 2) that training had been delayed or denied in the 30 days prior to survey administration (item 22). A similar percentage, 73.07%, strongly disagreed (38.46%, *n* = 10) or disagreed (34.61%, *n* = 9) that the aircraft were adequately maintained (item 23). Slightly over half, 56.00%, agreed (52.00%, *n* = 13) or strongly agreed (4.00%, *n* = 1) that they felt safe flying in the aircraft (item 24), although 24.00% disagreed or strongly disagreed (12.00%, *n* = 3 each) that they felt safe. Four-fifths of the Flight School A students, 80.77%, agreed (50.00%, *n* = 13) or strongly agreed (30.77%, *n* = 8) that the facilities were well maintained (item 25).

Table 10

Flight School A: Maintenance and Facilities

	Rating									
	Strongly Disagree		Disagree		No Opinion		Agree		Strongly Agree	
Item	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
22. Delayed/denied training in last 30 days	2	7.69	3	11.54	2	7.69	2	7.69	17	65.39
23. Aircraft adequately maintained	10	38.46	9	34.61	1	3.85	5	19.23	1	3.85
24. Feel safe in aircraft	3	12.00	3	12.00	5	20.00	13	52.00	1	4.00
25. Facilities well maintained	2	7.69	1	3.85	2	7.69	13	50.00	8	30.77

Note. One student did not respond to item 24.

Table 11 shows that, unlike the Flight School A students, over four-fifths, 84.62%, of the Flight School B students strongly disagreed (56.41%, *n* = 22) or disagreed (28.21%, *n* = 11) that training had been delayed or denied in the 30 days prior to survey administration (item 22). A slightly larger percentage, 89.75%, strongly agreed (61.54%, *n* = 24) or agreed (28.21%, *n* = 11) that the aircraft were adequately maintained (item 23). Almost all students, 92.31%, agreed (61.54%, *n* = 24) or strongly agreed (30.77%, *n* = 12) that they felt safe flying in the aircraft (item 24). Almost four-fifths of the Flight School B students, 79.48%, agreed (51.27%, *n* = 20) or strongly agreed (28.21%, *n* = 11) that the facilities were well maintained (item 25).

Table 11

Flight School B: Maintenance and Facilities

	Rating									
	Strongly Disagree		Disagree		No Opinion		Agree		Strongly Agree	
Item	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
22. Delayed/denied training in last 30 days	22	56.41	11	28.21	3	7.69	2	5.13	1	2.56
23. Aircraft adequately maintained	1	2.56	1	2.56	2	5.13	11	28.21	24	61.54
24. Feel safe in aircraft	0	0.00	0	0.00	3	7.69	12	30.77	24	61.54
25. Facilities well maintained	0	0.00	4	10.26	4	10.26	20	51.27	11	28.21

Note. One student did not respond to items 22-25.

Discussion

Student responses showed wide variation in both flight experience and opinions about the flight school training. Differences were discernible between the responses of students at Flight School A and Flight School B.

Comparison with Other Flight Schools and Programs

Responses to Section B of the FSES, Comparison with Other Flight Schools and Programs, indicated that most students at both Flight School A and Flight School B were divided fairly equally between feeling that the ISU flight schools were worse, about the same, or better than previous schools attended (see Table 5). Regarding progress, most students at both schools felt they were behind where they thought they would be (see Table 5). Students may have rated the flight schools similarly because of their limited experiences, especially since the Flight School B students lacked previous experience with flight schools. Nevertheless, it is interesting that most students felt that they were behind where they thought they would be. Perhaps the students' expectations as customers of the flight school were not being met.

Training Methods and Procedures

Responses to Section C, Training Methods and Procedures, showed that most students at both flight schools felt the staff was not open to suggestions for improvement (see Tables 6-7). However, the majority agreed that flight school staff were available for training and the school provided adequate instruction. In addition, the majority of students at both schools would recommend the school to others and felt safe flying (see Tables 6-7). With regard to the flight schools providing clearly stated objectives, Flight School B students rated this facility higher than Flight School A students (see Tables 6-7). These responses seem to indicate that Flight School B students experienced overall satisfaction with their training. Their high satisfaction may have resulted from of their lack of knowledge and experience in what to expect minimally from flight schools.

Flight Instructors

Responses to Section D, Flight Instructors, showed that students at both flight schools on the whole were satisfied with their individual instructors. Students rated instructors high on all elements (see Tables 8-9). Responses to this section were the most uniform for both flight schools.

Aircraft Maintenance and Facilities

Responses to Section E, Aircraft Maintenance and Facilities, showed that the majority of Flight School A students felt training had been delayed or denied in the 30 days prior to survey administration. These students also felt that the aircraft were not adequately maintained (see Table 10). However, somewhat paradoxically, these students also felt safe in the aircraft and viewed the facilities as well maintained (see Table 10). In contrast, well over the majority of the Flight School B students did not feel training had been delayed or denied in the previous 30 days and did feel that the aircraft were well maintained (see Table 11). Similar to Flight School A students, almost all of the students felt safe flying in the aircraft and that the facilities were well maintained (see Table 11).

These results indicated that students' levels of satisfaction with both flight schools were similar for many items in the FSES. Overall, students were satisfied with the adequacy of instruction, safety factors, flight instructors, and maintenance of the aircraft and facilities. However, students were dissatisfied with their own progress, with the schools' not being open to suggested improvements, with inadequate aircraft maintenance, and with the schools' delays or denials of training.

No studies were located that specifically assessed collegiate aviation programs for student satisfaction. However, present results generally confirm findings in the literature on customer satisfaction and student-flight instructor relationships. The areas of student dissatisfaction found in the present research supports Bellchambers and Neun's (2001) observations that customer feedback is a key indicator of service delivery. The extent of student satisfaction also bears out Yu's (2001) conclusion that reliability was more important than customization. In the present study, students' lack of satisfaction in some of the surveyed areas demonstrates their expectations of basic service.

High student satisfaction with flight instructors as primary representatives of the schools (Downs, 2000) adds strength to Hiner's (2000) assertion that if the relationship is of high quality, customers' accusatory behavior in adverse conditions tends to be low. In the present survey, students had relatively few complaints overall and did not indicate that instructors behaved irresponsibly in the ways outlined by Hiner.

However, students' assessment of management's lack of openness to suggestions for improvement confirms previous findings on the importance of maintaining customer relationships (Bellchambers & Neun, 2001; Bond & Fink, 2001). Moreover, students' feelings that training was delayed or denied and that aircraft maintenance was somewhat inadequate point to the conclusion that the flight schools were not meeting FAA requirements or ISU contractual obligations.

Conclusions

The findings of this study indicated that the quality of service provided by the two contracted flight schools to Indiana State University flight program students appears generally adequate. This conclusion contradicted students' verbal complaints to faculty. Nevertheless, when the results of survey item were examined individually, the quality of flight school service required much improvement in certain areas. These were flight schools' lack of openness to improvement, some inadequacy of aircraft maintenance, and delays or denials of training. Both flight schools must maintain the same level of training standards as directed by FAR Part 61. However, Flight School A received a substantially greater number of verbal student complaints, and survey results did bear out these complaints, in contrast to Flight School B. These appear in items 6-25 (see Tables 5-11).

Several implications may be seen from the study results. Because of both student verbal feedback and survey results, reexamination of the entire flight school program is warranted by the ISU Department of Aerospace Technology. Implementation of student satisfaction surveys each term may be one way to monitor the flight schools' service delivery. Survey administration would not only greatly reduce students' complaints but could also decrease the risk of accident or injury by early identification of flight safety problem areas.

Recommendations

A number of recommendations may be made from the study results. First, the Department of Aerospace Technology should review the entire flight school program on a regular basis. Second, an alternative training program, such as use of ground-based flight simulators should be developed for students who experience delays in flight school training. Such a program would meet student requirements for continuous training and act as a backup system for aircraft maintenance delays.

Third, study results should be distributed to all students, department staff and faculty, and flight school managers. Feedback would be solicited concerning the results, and this feedback could be incorporated into a subsequent report and improvement of flight training. Fourth, monthly meetings should be scheduled to include faculty, student representatives, and flight school contractors for assessment of new implementation as well as guidance for continuous improvement.

Fifth, a more comprehensive survey based upon the FSES should be developed. This survey should be administered to students after the improvements are in place, and the results compared with the present study results. This comparison would act as an ongoing evaluation of the flight school training. Finally, separate studies should be conducted with students at both flight schools to determine whether significance differences exist in their responses. If these are found, additional exploration should be made of the source of these differences.

The recommended systematic surveys and resultant reports could substantially enhance communications among the students, department staff and faculty, and flight school managers and instructors. Exchange of information and sharing of problems could contribute to much improved flight school instruction, and fulfillment of FAA requisites and contractual obligations at the ISU flight-training program.

Summary

In the context of students as customers to be satisfied and delivered an excellent product, this study identified ISU flight school students' satisfactions and important dissatisfactions with the two flight schools providing training. Study results, as well as the recommendations offered, call for a more proactive response by the Department of Aerospace Technology. The department should institute continuous monitoring of student flight training and take steps toward improvement so that the ISU flight training program may consistently meet both the mandated FAA requisites and the ISU flight students' needs and requirements. It is hoped that this study will serve as a model for other aerospace departments to become more aware of, monitor, and make the necessary improvements in their flight training programs to deliver the safest and most high-quality flight training to their students.

References

Bellchambers, G. R., & Neun, D. J. (2001, August). Building on customer satisfaction. *Professional Remodeler*, *5*, 72-76.

Bond, E. U. & Fink, R. L. (2001, April). Meeting the customer satisfaction challenge. *Industrial Management, 43*, 26-31.

Bowen, B., Carstenson, L., & Hansen, F. (1999). Recruiting from within: Action-oriented research solutions to internal student recruitment in collegiate aviation education. *Journal of Air Transportation Worldwide, 4* (1), 14-25.

Bryan, J. L. (1996). An analysis of student programmatic delays in postsecondary flight training programs: A national study. *Journal of Air Transportation Worldwide, 1* (1), 63-83.

Davisson, B. (2000, May). Checklist for learning: 17 things you must know. *Flight Training*, 23, 20-24.

Downs, E. C. (2000, May). Take charge of your learning: Make every lesson count. *Flight Training*, 23, 28-30.

Federal Aviation Administration. (FAA). (2003). *Part 61-Certification: Pilots, Flight Instructors and Ground Instructors*. Retrieved September 3, 2003, from the www.faa.gov/avr/afs/fars/far-61.txt

Hiner, R. (2000, April). Seven things instructors do to irritate students. *Flight Training, 23*, 16-18.

Kotler, P., & Armstrong, G. (2001). *Principles of marketing* (9th ed.). Upper Saddle River, NJ: Prentice Hall.

Luedtke, J. R., & Papazafiropoulos, I. (1996). Retention in collegiate aviation. *Journal of Air Transportation Worldwide*, *1* (1), 39-50.

Yu, L. (2001). What really makes customers happy? *MIT Sloan Management Review, 42* (4), 19-31.

Wurman, R. (2000, June). Rethinking school policies. *Flight Training, 23*, 52-53.

Appendix

Flight School Evaluation Survey (FSES)

The purpose of this study is to gather your views as how you perceive the Flight School is meeting the needs of Indiana State University Department of Aerospace Technology Students. Your sincere responses will help the department accomplish the departmental goals and meet your needs. Please ensure you have read and signed the survey consent form prior to beginning this survey.

The term *Flight School* is used generically to indicate where you are currently receiving your flight training as a student.

Please indicate your answer to each of the following questions reflecting on only the question being asked. If you feel you don't want to answer a particular question, leave it blank and move on to the next question. Please darken in the circles as you feel best describes your opinion.

Please indicate at which airport you currently receive your flight training by darkening the appropriate circle.

○ Flight School A ○ Flight School B

A. Student Demographics

The following questions will help us gather information about where you are in training and general background information.

1. By number of credit hours completed, would the university consider you a:

○ Freshman (0-31)
○ Sophomore (32-62)
○ Junior (63-93)
○ Senior (94 and above)

2. How long have you been in flight training (in academic years)?

○ Less than one year
○ More than one year but not more than two year
○ Two years but less than three years
○ Three years but less than four years
○ Four years or more

3. How many flight hours have you completed?

○ Less than 100 hours
○ 101-200 hours
○ 201-300 hours

- ○ 301-400 hours
- ○ More than 400 hours

4. Indicate the flight certificate you currently possess (answer only once for the highest certificate held).

- ○ Student certificate
- ○ Private
- ○ Instrument
- ○ Commercial
- ○ Certified Flight Instructor (CFI)

5. How many other flight schools have you attended?

- ○ None
- ○ One
- ○ Two
- ○ Three or more

B. *Comparison with Other Flight Schools and Programs*

6. If you have attended other flight schools, how do you feel the Flight School compares to the other flight schools you have experienced?

- ○ Worse
- ○ About the same
- ○ Better
- ○ I have not attended another flight school

7. In your opinion, how is your flight training progressing as compared to where you thought you would be at this time?

- ○ Behind where I thought I would be
- ○ About where I thought I would be
- ○ Ahead of where I thought I would be

C. *Training Methods and Procedures*

The following questions will help us gather information about the Flight School's training procedures and methods. Your input will assist us in identifying any factors we should consider.

Using the scale provided, please indicate your level of agreement with each of the following statements.

1. Strongly Disagree   2. Disagree   3. No Opinion   4. Agree   5. Strongly Agree

Appendix (Continued)

8. The Flight School is open to student suggestions for improvement.

1 2 3 4 5

9. The Flight School staff is "visible" on a daily basis to assist me in my training.

1 2 3 4 5

10. The Flight School provides adequate instruction for my preferred tempo of training.

1 2 3 4 5

11. I would recruit my friends and acquaintances to complete their flight training at the Flight School.

1 2 3 4 5

12. I feel safe flying with Flight School staff.

1 2 3 4 5

13. The Flight School has clearly stated training objectives.

1 2 3 4 5

D. *Flight Instructors*

14. The instructors use training time well.

1 2 3 4 5

15. The instructors are readily available for consultation with students.

1 2 3 4 5

16. The instructor seems to know when I don't understand the material.

1 2 3 4 5

17. The instructor is well prepared for each flight training session.

1 2 3 4 5

Appendix (Continued)

18. The instructor tells me how I will be evaluated before a flight.

1 2 3 4 5

19. The instructor summarizes major points after each flight.

1 2 3 4 5

20. The instruction provided continues to stimulate my interest in flight training.

1 2 3 4 5

21. I have had two or more flight instructors in the past 30 days.

1 2 3 4 5

E. *Aircraft Maintenance and Facilities*

The following questions will help us gather information about the Flight School's aircraft and facilities.

22. I have been delayed or denied flight training on two or more occasions in the last 30 days due to a maintenance problem with an aircraft.

1 2 3 4 5

23. I feel the aircraft are adequately maintained.

1 2 3 4 5

24. I feel safe flying in the Flight School aircraft.

1 2 3 4 5

25. Flight School facilities are well maintained.

1 2 3 4 5

International Journal of Applied Aviation Studies, 4, Number 2

Teaching Pilots Judgment, Decision-Making, & Critical Thinking

Charles L. Robertson

University of North Dakota
John D. Odegard School of Aerospace Sciences
19 Vail Circle
Grand Forks, ND 58201
(701) 772-3033
(701) 777-4061
Email: robertso@aero.und.edu

Abstract

Teaching pilots judgment, decision-making, and critical thinking involves teaching higher-order thinking skills. In few other settings is the development of higher-order thinking skills more critical than in aviation, where approximately 75% of all aviation accidents are caused by pilot errors. This paper discusses the methods and strategies for teaching higher-order thinking skills and the learning theory supporting these methods. Teaching higher-order thinking skills involves emphasizing methods and strategies for developing cognitive skills using problem-based learning. The instruction must be student-centered, utilize authentic, real world situations, active learning, cooperative learning, and be customized to meet the individual learner's needs. Cognitive skills are learned like other cognitive knowledge and should be based on cognitive and constructivist learning principles. Specifically, information processing, knowledge structure, and situated learning theories underpin the development and transfer of these skills. Improving higher-order thinking skills will lead to better judgment, decision-making, and critical thinking; hence, fewer aviation accidents.

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

In few other settings is the development of higher-order thinking skills (HOTS) more critical than in aviation. In fact, "it is estimated that approximately 75% of all aviation accidents are human factors related" (*Aviation Instructor's Handbook* (AIH), 1999, p. 9.8). Historically, these accidents were reported as pilot error, meaning an action or decision made by the pilot was the cause of, or was a contributing factor which lead to, the accident (AIH, 1999, p. 9.8). "Teaching pilots to make sound decisions is the key to preventing accidents" (AIH, 1999, p. 9.9).

"Higher order thinking skills are essential and must be taught. Recent findings of cognitive research provide a better understanding of how people learn and how they solve problems, from which new teaching strategies are emerging" (Kerka, 1992). Effective methods for teaching HOTS are already in use in disciplines outside aviation; understandably, their use in aviation will occur only if the aviation community recognizes their value in improving pilot judgment training and in reducing accidents. Teaching HOTS effectively involves emphasizing higher-order thinking strategies using problem-based learning (PBL) instruction, authentic, based on real world problems, student-centered, active learning, cooperative learning, and customized instruction to meet the individual learner's needs (Carr, 1990; Cotton, 1991; Howe & Warren, 1989; Kerka, 1992; Reigeluth, 1999). To adopt these new teaching strategies in aviation, a closer look at the learning theories that support the development and transfer of HOTS will be required. It also will be necessary to define HOTS and to look at how they are taught in aviation and in other disciplines. This examination will begin with definitions and a discussion about teaching HOTS.

Definitions

Before an examination of the different learning theories can be made, a working definition and a clear understanding of HOTS are needed. It is generally accepted that higher-order thinking skills are the cognitive process and cognitive skills involved in making a rational decision on what to do or what to believe (Ennis, 2000). In contrast, according to Cotton (1991), there is no universally accepted definition of higher-order thinking, creative thinking, critical thinking, or decision-making. Thomas and Albee (1998) asserted that "critical/creative/ constructive thinking is closely related to higher-order thinking: they are actually inseparable" (What is Higher-Order Thinking section, ¶ 9). Alvino (1990) offered a "Glossary of Thinking-Skills Terms" which: "Are widely—though not universally—accepted by theorists and program developers. [...] Bloom's Taxonomy - categorizes thinking skills from the concrete to the abstract—knowledge, comprehension, application, analysis, synthesis, and evaluation. The last three are considered higher-order skills." Thus, HOTS are analysis, synthesis, and evaluation skills (Alvino, 1990; Cotton, 1991; Reigeluth & Moore, 1990).

Higher-order thinking skills include analysis, synthesis, and evaluation and they describe the thinking skills used in judgment, decision-making, and critical thinking. HOTS are learned in a similar fashion and are support by the same learning theories. It seems realistic to examine the learning theories as they collectively apply to cognitive skills. According to Reigeluth and Moore (1999), Bloom's higher-order thinking skills "are all taught through basically similar methods" (p. 55). It should be true also that, if HOTS are learned through similar methods, thus similar learning theories should support these methods.

Teaching HOTS in aviation and non-aviation settings should be the same; however, the current guidance in aviation omits any references to teaching cognitive skills. With the exception of the omitted emphasis in teaching the cognitive skills, they are the same phenomenon. In flight education, judgment, decision-making, and critical thinking are taught as aeronautical decision-making (ADM), while they are taught as HOTS outside aviation. According to *Advisory Circular 60-22* (AC60-22) (1991) and *AIH* (1999), "ADM is a systematic approach to the mental process used by aircraft pilots to consistently determine the best course of action in response to a given set of circumstances;" conversely, HOTS are both the cognitive process and skills for deciding what to do (Ennis, 2000; Howe & Warren, 1989). The lack of guidance means cognitive skills typically are not emphasized in pilot training; thus, they are not being taught as effectively as they need to be to reduce the number of pilot error accidents.

Traditionally, the literature in aviation reflects the omission of any reference to teaching or the development of cognitive skills (Bell & Mauro, 1999; Buch & Diehl, 1984; Deitch, 2001). However, a number of authors have begun to present and discuss the value of teaching cognitive skills in addition to the cognitive process in ADM (Cohne & Freeman, 1996; Connolly, 1990; Jensen, 1988; Ryder & Redding, 1993; Shebilske, Regian, Winfred, & Jordan, 1992; Wiggins, 1997). Such reports raise an important concern about why the current guidance and training materials do not reflect the need to teach cognitive skills along with the cognitive process. While it could be argued that higher-order thinking is far more complex than simply determining where to land the airplane, the underlying skills needed in making decisions (analysis, synthesis, and evaluation) are the same regardless of the complexity of the problem or the setting. The issue in aviation is whether pilot judgment can be improved by enhancing both the cognitive process and skills. It is reasonable to conclude that the strategies and methods used to teach these cognitive skills elsewhere must be adopted by the aviation community to improve the pilot's ability to make good judgments and to gain the maximum benefit of ADM training.

The Requirements for Teaching Thinking Skills

The requirement for teaching HOTS can be identified by examining the current teaching methods and strategies used in disciplines outside of aviation. However, Cotton (1991) said, "There is no one best way to teach thinking skills" (Programs, Strategies, and Training are Important section, ¶3). The research supports instruction in many specific skills and techniques using various instructional approaches to promote development and enhancement in thinking skills. To foster the development of thinking skills, the instruction redirection/probing/ reinforcement, asking higher-order questions, lengthening wait-time (Cotton, 1991). She drew this conclusion after reviewing 56 documents, including 33 reports of research studies or reviews of which 23 were descriptive, theoretical, or guidelines documents or they were concerned with research in areas other than the effectiveness of programs and practices. The implication from these papers is that any strategy or technique employed to facilitate learning thinking skills can be effective, if properly administered. These strategies involve engaging the learner in some form of mental activity, examining that mental activity, and then challenging the learner to explore other ways to accomplish the task or the problem (Landa, 1999).

In contrast to the strategies recommended by Landa, the *AIH* (1999) said "the best way to illustrate this concept [poor judgment chain] is to discuss specific situations which lead to aircraft accidents or incidents…a scenario which can be presented to students to illustrate the poor judgment chain" (p. 9.8). "By discussing the events that led to this incident, instructors can help students understand how a series of judgmental errors contributed to the final outcome of this flight" (p. 9.9). This difference between the strategies offered by Landa and the *AIH* is Landa's approach actively engages the learner in mental activities, examination, and evaluation, while the AIH directs the instructor to illustrate the poor judgment chain so the pilot can passively understood. According to *AIH* (1999), "ADM training focuses on the decision-making process and the factors that affect a pilot's ability to make effective choices" (p. 9.9). Nevertheless, it could be argued that the scenario presented by the instructor will provide the pilot with an example of how to solve this problem and this example could be recalled later to decide what he or she should do to break a similar poor judgment chain. However, it does not teach the pilot how to handle unfamiliar error chains.

This is the critical difference between teaching judgment in aviation and outside aviation. In aviation, the many scenarios are presented to the student pilot as worked examples demonstrating how the expert would solve a problem or a series of problems. Outside of aviation, this approach may be referred to as case studies. The difference would occur when the instruction outside aviation would include instruction and practice in applying these techniques to new situations. In other words, teaching the learner to solve ill-defined, ill-structured, complex problems. This approach assumes well developed analysis, synthesis, and evaluation skills, which ADM does not address while HOTS does. A closer look at teaching HOTS follows.

Teaching higher-order thinking skills effectively involves customizing the examination and exploration of the mental activity to meet the individual learning needs of the learner. Kerka (1992) said:

> Learning is characterized as an active process in which the learner constructs knowledge as a result of interaction with the physical and social environment. Learning is moving from basic skills and pure facts to linking new information with prior knowledge; from relying on a single authority to recognizing multiple sources of knowledge; from novice-like to expert-like problem solv-ing[(Thomas, 1992)]. (What Strategies Develop These Skills section, ¶ 1)

Howe and Warren (1989) added, "There needs to be a shift in many classes, from a teacher-centered classroom to a student-centered classroom in which students can be involved in collecting and analyzing information, paired problem solving, cooperative learning settings, simulations, debates, and critical reporting sessions" (What Does Research Indicate Regarding Teaching Critical Thinking section, ¶ 6). In addition to the approaches offered above, Landa (1999) said that three strategies can be used to facilitate the learning of thinking skills; they are guided discovery, expository teaching, and a combination strategy. Landa (1999) described the guided discovery strategy as (a) giving the learner a task or problem and have them perform it, (b) help the learner formulate a method (detailed set of instructions) to follow to perform the task, (c) have the learner examine the mental activity, and then (d) challenging them to explore other ways to accomplish the task or the problem. Teaching HOTS effectively involves emphasizing HOTS strategies in PBL which includes problem solving-, case study-, and scenario-based instruction (Reigeluth, 1999). In addition, Cotton (1991) said, "Educators are now generally agreed that it is in fact possible to increase students' creative and critical thinking capacities through instruction and practice" (Introduction section, ¶ 9). Ristow (1988) and Presseisen (1986) reiterate, students can learn HOTS, if schools will concentrate on teaching them how to do so.

In contrast to the guidance provided above for teaching HOTS, the current guidance for aviation omits any discussion of providing simple to complex or concrete to abstract scenarios. It also omits any guidance involving active engagement by the pilot in learning or in learning activities or in actually engaging in problem solving or decision-making. The guidance implied an instructor-centered learning environment under the complete control of the instructor. It also omits any involvement in collecting and analyzing information, paired problem solving, cooperative learning, simulations, debates, and critical reporting. The scenario-based instruction does not provide the pilot the opportunity to use guided discovery or discovery learning when the pilot is encouraged to learn anything beyond the application and correlation levels of learning. The *AIH* (1999) emphasized instead, learning the decision process which it refers to as the "DECIDE" model – (a) detect-the fact that a change has occurred, (b) estimate-the need to counter or react to the change, (c) choose-a desirable outcome for the success of the flight, (d) identify-actions which could successfully control the change, (e) do-the necessary action to adapt to the change, and (f) evaluate-the effect of the action. According to the *AIH* (1999), "a problem is perceived first by the senses, then is distinguished through insight and experience" (p. 9.11). This implied that the cognitive skills needed in ADM are learned elsewhere or through experience and not learned through instruction.

Before the appropriate learning theory is examined, the requirements for facilitating the transfer of HOTS from the instructional setting to their application should be discussed. Transfer of knowledge relates to first learning and storing information in long-term memory (LTM) and then retrieving or recalling that information from LTM in the application of that information. The information involves declarative knowledge (knowledge about things) and procedural knowledge

(knowledge about how to do things) (Clark, 1999). Critical in the learning setting is avoiding inert knowledge; that is, knowledge that is learned and cannot be recalled later. Avoiding inert knowledge typically involves relating information to the environment where the knowledge is to be applied and falls into either near- or far-transference (Alessi & Trollip, 2001; Clark, 1999). Near- and far-transference does not depend on the instructional model, but rather on the nature of the problem, scenario, or case presented. Developing authentic and realistic problems in the learning setting with very similar circumstances to those occurring in aviation will promote near-transfer. Near-transfer teaching methods involving practicing or drilling step-by-step procedures may be used because there is little variance in the application of the procedure. In contrast, far-transfer teaching methods may be needed in other abnormal or emergency situations, particularly when ill-structured, ill-defined complex problems are involved and where the pilot must use extensive judgment, must use a different approach, or there are no set steps or set procedure established. According to Clark (1999), "use schema-based instructional models to teach far-transfer tasks including problem solving tasks that (a) use a schema-based training design, (b) provide varied context examples and problems, and (c) teach related process knowledge."

The transference of knowledge from the learning environment to its practical application is not a separate problem from the learning and application of HOTS. It is the problem addressed in the literature in teaching HOTS. That is, learning the cognitive skills underlying the decision-making process in a learning environment so they may be recalled from LTM and applied when an ill-defined, ill-structured, complex problem-solving situation occurs in aviation is the instructional challenge besetting the development of HOTS. The learning requirements discussed above about teaching HOTS are the near- and far-transference challenges in learning.

Learning Theories Supporting Thinking Skills

Now that the definition, relationship, and requirements or the "what" side of the teaching issue have been addressed, it is time to consider how HOTS are learned. In other words, what learning theories explain how higher-order thinking skills are learned? Again, this is important because for teaching methods to be effective they must be based on learning theory (Alessi & Trollip, 2001: Carnegie, 2002; Reigeluth, 1999). After a brief overview of the behavioral, cognitive, and constructivist learning principles and a discussion of their ability to support the requirements of teaching HOTS, specific learning theories supporting these instructional strategies and methods will be discussed. To this end, Alessi and Trollip (2001) said, "no universal agreement exists on how learning occurs. How psychologists have viewed the principles of learning has changed significantly throughout the 20th century" (p. 16). Driscoll (2000) said:

> Despite the differences among the learning theories…they do share some basic, definitional assumptions about learning. First, they refer to learning as a persisting change in human performance or performance potential. This means that learners are capable of actions they could not perform before learning occurred and this is true whether or not they actually

> have an opportunity to exhibit the newly acquired performance… Second, to be considered learning, a change in performance or performance potential must come about as a result of the learner's experience and interaction with the world (p. 11).

The persisting change in human performance or performance potential resulting from the learner's experience and interaction with the world can be explained by behavior, cognitive, or constructivist learning theories or a combination of one or more of these theories and the specific learning theories within these theories.

Behavioral

Behaviorism appears to have been built on the foundation begun by Ebbinghaus' verbal learning experiments and the work of Pavlov and Thorndike (Driscoll, 2000). In fact, according to Driscoll, Ebbinghaus is credited with ushering in a new era of interest in the study of learning, when he began experimenting with the notion that if ideas are connected by the frequency of their association, then learning should be predictable. Thorndike, on the other hand, believed sensation and impulse, rather than ideas association was important. This led Thorndike to propose the Law of Effect. Meanwhile, Pavlov's experiments led to classical conditioning. These early works formed the groundwork for Skinner's radical behaviorism. Skinner's work refined and demonstrated that a particular pattern of reinforcement or punishment resulted in different rates of learning or degrees of retention, based on the principle of association, Law of Effect, classical conditioning, and operant conditioning. Central to this theme is the belief that learning is always an observable change of behavior and it is a result of connecting certain responses with a given stimuli (Alessi & Trollip, 2001; Behaviorism, 2001; Carbenell, 2001; Huitt, 1997; Murphy, 1997; Operant Conditioning, n.d.; Operant Conditioning, 1996).

The behaviorist learning theory "maintains that learning should be described as changes in the observable behavior of a learner made as a function of events in the environment," (Alessi & Trollip, 2001) and it includes Pavlov's classical conditioning. According to Alessi and Trollip (2001):

> The basic principle of classical conditioning is that repeatedly pairing a neutral stimulus with a natural stimulus (one that elicits a natural response) causes the neutral stimulus also to elicit the response. The implication is that humans learn many behaviors because of their pairing with basic human needs and responses, such as the need for food, sleep, reproduction, and the like (p. 18).

When classical conditioning is coupled with operant conditioning, the use of rewards and punishments, the behavior modification can be more efficient and effective. Criticism of this approach argued it ignores important unobservable aspects of learning (such as thinking, reflection, memory, and motivation) (Alessi & Trollip, 2001).

> Decades of learning research has demonstrated that classical and operant conditioning principles do not predict all learning outcomes. Theories of motivation, memory, transfer, and the like have promoted instructional methods that behavioral techniques would not… The outcomes of education and training must include more that just learner achievement. They must include learner satisfaction, self-worth, creativity, and social values… People must be adaptive and lifelong learners, must have the confidence necessary to change with their environment, and must be able to work collaboratively with others… These goals are values were marginally recognized by behavioral approaches to education… Behavioral principles such as positive reinforcement, corrective feedback, and spaced practice are appropriate in interactive [settings] (Alessi & Trollip, 2001, pp. 36-37).

Traditionally, the behavioral principles have been used to explain how aviators learn various flight procedures including responses to changing flight conditions, abnormal and emergency situations. If HOTS are to be taught in the future to enhance pilot judgment, then unobservable behavior or cognitive learning will need to be employed, in addition, to this traditional approach.

Cognitive

In contrast to the behaviorist view that learning affects observable behavior only, cognitive learning theory involves the mental processes of learning. That is, cognitive learning "places emphasis on the unobservable constructs, such as the mind, memory, attitudes, motivation, thinking, reflection, and other presumed internal processes" (Alessi & Trollip, 2001; Cognitive Learning Theory, n.d.; Cognitive Learning Theory Terms, n.d.; Information Processing Theory, 1996). According to Driscoll (2000),

> In cognitive information processing view, the human learner is conceived to be a processor of information in much the same way a computer is. When learning occurs, information is input from the environment, processed and stored in memory, and output in the form of some learned capability (p. 76).

The cognitive learning theory addresses the process occurring inside the learner's mind and the internal processes of learning (Reigeluth & Moore, 1999; Alessi & Trollip, 2001). According to Alessi and Trollip (2001, p. 19): "Cognitive psychology places emphasis on unobservable constructs, such as the mind, memory, attitudes, motivation, thinking, reflection, and other presumed internal processes." Cognitive learning is dominated by the information-processing approach. "The areas of cognitive theory that are most important to [instructional] design are those relating to perception and attention, encoding of information, memory, comprehension, active learning, motivation, locus of control, mental models, metacognition, transfer of learning, and individual differences (Anderson, 1980; 1981; Anderson, 1977; Berger, Pezdek & Banks, 1986; Bower & Hilgard, 1981; Gagné, Yekovich, & Yekovich, 1993; Kozma, 1987)" (Alessi & Trollip, 2001, p. 20). Conversely, "the cognitive approach has undervalued the powerful principles of reinforcement. Cognitive educators spoke of collaboration, communication, and transfer… [but] they did not do a very good job of translating such principles into practice in the learning environments they created" (Alessi & Trollip, 2001, p. 37).

Cognitive learning principle effectively describes most of the learning issues missed in the behavioral learning. They seem to support the HOTS requirements; that is, a mixture of cognitive and behavioral principles will need to be used. The cognitive learning principles support the cognitive skills overlooked by the behaviorist, while the procedural responses taught in response to various stimuli are effectively supported by the behavioral principles. Nevertheless, there are some requirements for teaching HOTS that are better explained by the constructivist learning principles. Therefore, the constructivist learning principles must also be discussed before examining the strengths and weaknesses of the behavioral and cognitive approaches.

Constructivist

The constructivist approach, on the other hand, asserts, "Learning is a process of people actively constructing knowledge, where traditional instructional methods, such as memorizing, demonstrating, and imitating, are considered incompatible with the notion that learning is a process of construction" (Alessi & Trollip, 2001, p. 32). According to Reigiluth and Moore (1999), the following principles or suggestions typically are promoted as ways to accomplish the goal of allowing learners to actively construct their own knowledge:

> (a) Emphasize learning rather than teaching. (b) Emphasize the actions and thinking of learners rather than of teachers. (c) Emphasize active learning. (d) Use discovery or guided discovery approaches. (e) Encourage learner construction of information and projects. (f) Have a foundation in situated cognition and its associated notion of anchored instruction. (g) Use cooperative or collaborative learning activities. (h) Use purposeful or authentic learning activities. (i) Emphasize learner choice and negotiation of goals, strategies, and evaluation methods. (j) Encourage personal autonomy on the part of learners. (k) Support learner reflection. (l) Support learner ownership of learning and activities. (m) Encourage learners to accept and reflect on the complexity of the real world, and. (n) Use authentic tasks and activities that are personally relevant to learners (Alessi & Trollip, 2001, p. 32).

Furthermore, constructivists maintain that traditional methods [tutorial and drill instruction] produce knowledge that does not transfer well or inert knowledge. They suggested that methodologies such as hypermedia, simulation, virtual reality, and open-ended learning environment are of more benefit to learners, allowing them to explore information freely, apply their own learning styles, and use software as a resource rather than as a teacher (Alessi & Trollip, 2001).

> Growing research evidence indicates that constructivist methods work better only for learners with well-developed metacognitive skills. Some evidence also indicates that constructivist techniques are very time consuming. … Constructivist techniques are good for some types of learning, some situations, and some learners, but not all (Alessi & Trollip, 2001, p. 39).

Strengths and Weaknesses of Behavioral, Cognitive, and Constructivist

It appears the behavioral principles do little to support the learning requirements of the cognitive skills element of higher-order thinking. However, HOTS are defined as both the cognitive process and the cognitive skills involved in making a rational decision on what to do or what to believe (Ennis, 2000). Therefore, when the cognitive process portion of this definition is considered, behavioral principles support the procedures that are carried out as a response to a stimulus. In fact, most of the current normal, abnormal, and emergency procedures (cognitive processes) training in aviation are conducted using the behavioral principles with extensive use of positive reinforcement, corrective feedback, and spaced practice (AIH, 1999). While many normal, abnormal, and emergency situations can be taught and practiced effectively, many more situations cannot be anticipated and taught; thus, the aviator is also required to employ cognitive skills in many situations where he or she was not specifically trained nor had previously experienced. Understandably, behavioral principles do nothing to explain how the cognitive skills are learned when either unanticipated or multiple responses are required, possible, and appropriate; in other words, when the pilot is faced with ill-defined ill-structured complex problems.

Behavioral principles do not adequately address the transfer of learning problem occurring between the training setting and the application of decision-making in-flight. Behavioral principles would produce near-transference learning that could be applied in aviation; in fact, normal, some abnormal, and some emergency procedures have been taught effectively where an identifiable stimulus (system malfunction) resulted in a set procedure response. However, the extensive judgment and problem-solving skills coupled with the lack of observable behavior and the cognitive nature of far-transference means the behavioral principles do not support it.

Consequently, the remaining discussions in this section will focus on cognitive and constructivist principles, because these learning theories appear to support and explain the learning of the cognitive skills element of HOTS. Indeed, Reigeluth and Moore (1999) said, "cognitive learning theory has contributed the most to understanding how best to teach and test this type of learning [higher order thinking skills]" (p. 55). Additionally, the points offered by Driscoll (2000) and Perkins (1991) are worthy of note; that is, "there is no single constructivist theory of instruction" (Driscoll, 2000, p. 375) and:

> Constructivist theory rests on the assumption that knowledge is constructed by learners as they attempt to make sense of their experiences. Learners, therefore, are not empty vessels waiting to be filled, but rather active organisms seeking meaning. Regardless of what is being learned, constructive processes operated and learners form, elaborate, and test candidate mental structures until a satisfactory one emerges (Perkins, 1991a, p. 376).

Driscoll helped clarify Alessi and Trollip's account of the constructivist theory. The cognitive and constructivist learning theories provide a basis for grounding a wide range of instructional designs which support all learning situations where the cognitive skills are taught. Learning theories provide a theoretical basis for the instructional designs and provide insight into what the instructional design needs to do to promote effective learning. The strengths and weaknesses of the learning theory underpinning the instructional design will affect the effectiveness of the learning. Therefore, examining the specific learning theories supporting HOTS is necessary to gain a complete understanding of the development and transference of cognitive skills supporting HOTS.

Specific Learning Theories Supporting HOTS

Cognitive research has shown the learning of HOTS is not a change in observable behavior but the construction of meaning from experience (Johnson & Thomas, 1992; Thomas, 1992). This implies the constructivist theory provides the best support of higher-order learning. Thomas' 1992 assertion that there are three types of cognitive theories upon which teaching strategies should be based has gone unchallenged. According to Thomas (1992), these three cognitive theories are (a) information processing theory, (b) knowledge structure theories, and (c) social history theory. However, Clark (1999) pointed out that while individual meaning construction facilitates thinking skills, there is little support for building a common set of knowledge and skills among learners in constructivist instructional designs, "the uniqueness of constructed knowledge is acknowledged" (p. 181). Thus, the procedural task and the stimulus-response support of the behaviorism and the information processing theory of the cognitive theory are required along with the constructivist theory to support fully the development of HOTS.

The information processing theory explains how the mind takes in information, knowledge structure theories depict how knowledge is represented and organized in the mind, and social history theory explains the vital role of cultural context in the development of individual thinking. Because, the social history theory explains the development of individual thinking as it may apply to one's social responsibility and it addresses the role of previous schema in long term memory (Thomas, 1992); therefore, the role of cultural context provides a better explanation for making a rational decision on what to believe rather than for what to do. Lave's situated learning theory should be used in aviation. Lave's situated learning theory recognizes the role of previous schema, the schema's affect on learning new knowledge, how knowledge needs to be presented and learned in an authentic context. It pays particular attention to the settings and applications that normally involve that knowledge and how learning requires social interaction and collaboration as well (Lave, 1996). Thus, the situated learning theory is better suited to explaining and supporting learning of HOTS.

The first two cognitive theories, information processing and knowledge structure theories and constructivist theory, situated learning theory, collectively support the learning activities discussed in the beginning of this section and provide the theoretical underpinnings of teaching and learning thinking strategies (Bell & Mauro, 1999; Carr, 1990; Cotton, 1991; Diehl, 1991; Kerka, 1992; Peirce, 2001; Splitter, 1995). Hence, they are the principal learning models that support the instructional methods and teaching strategies discussed as requirements of teaching HOTS above including guided discovery, expository, problem-based learning, simulation, tutorials, team-building, redirection, probing, reinforcement, and high-order questions approaches to learning. These learning models provide a choice of theoretical foundations upon which instructional designs for different learning settings can be based and range of learning models the instructor may chose from to customize the instruction to the individual learner's needs. Furthermore, they provide a starting point for addressing the learning models or theories and their role in facilitating the learning of HOTS.

The strategies for teaching HOTS discussed in the previous section employ the same learning theories used in acquiring other cognitive skills. The learning theories supporting learning cognitive skills are not either cognitive or constructivist rather they are from both. In some situations the learning process is driven by information-processing which emphasizes perception and attention, encoding of information, memory, comprehension, active learning, motivation, locus of control, mental models, metacognition, transfer of learning, and individual differences. In other situations, the process will be driven by individual construction of meaning, situated learning, and collaborative learning.

Unanswered Questions

Many questions about learning and teaching HOTS remain. These questions are:

- Will improvements in HOTS improve pilot judgment and eliminate pilot-error type accidents in general aviation?
- Does decision-making require domain specific knowledge to be learned to the higher-order thinking levels of learning?
- Should the exploratory instructional method be used in aviation?
- What is the appropriate mix of learning theories needed to learn HOTS efficiently and effectively?

Indeed, will improving the underlying HOTS improve pilot judgment and decision-making and subsequently result in a reduction of pilot-errors and the general aviation accident rate? To answer this question, additional research is needed. This is the next step in an on-going research project being conducted in a joint effort among the FAA, industry, and two major aviation universities. However, before this research can be done, effective methods for teaching HOTS need to be identified and applied. This paper was limited to identifying the teaching methods and strategies that should be used to improve ADM training to determine the learning requirements. Other papers will need to answer this question.

Next, is judgment and decision-making a direct result of learning knowledge at the higher-order thinking levels? Typically, the literature addressed judgment and decision-making as a cognitive process and cognitive skills except in the aviation literature. HOTS are domain specific, based on knowledge within that domain, and not easily generalized to other domains. Furthermore, it suggested learning within the domain is enhanced when thinking strategies are taught with subject content. These are subjects worthy of additional research and would be important in a number of areas where judgment and decision-making abilities are important in reducing risk or where rationally deciding what to do is important. Again, these subjects are beyond the scope of this paper.

Can the exploratory method of instruction touted by the constructivist as the most effective way to facilitate individual meaning making be used effectively in aviation? The literature of the exploratory method said that the exploratory method is not well suited to settings where standardization or specific content is required, because the specific content selected for learning must controlled by the learner's desires and will likely only include the content the learner wants that meets the individual goals and needs (Clark, 1999). The exploratory method will likely have limited application in pilot training in favor of the guided discovery or other more regimented instructional approaches such as problem-based learning. In other words, the quality of individual learning can be realized, while ensuring all desired outcomes are meeting the requirements of aviation. Clark (1999) suggested it is only "appropriate when a specific predetermined set of knowledge and skills is not an essential outcome." (p.180). In aviation, a specific predetermined set of knowledge and skills typically is considered an essential outcome; thus, it is unlikely this method will be used significantly or beyond a supporting role in aviation.

Finally, what is the appropriate mix of learning theories and concepts needed to facilitate efficient and effective learn of HOTS? This comparison does not provide an instructional design for teaching HOTS in aviation; thus, it does not answer this question. Until the research establishes the value of teaching HOTS in reducing accidents by improving judgment and decision-making skills, research on maximizing the learning should be delayed. Again, this question is beyond the scope of this comparison.

Conclusion

The cognitive skills needed to make good judgments and decisions are teachable. The aviation community needs to incorporate the instruction of these skills into its aeronautical decision making training to reduce the number human factors caused accidents. Cognitive skills are being taught outside of aviation as HOTS and they are taught by integrating thinking skills strategies in combination with other learning activities. In other words, to enhance judgment and decision-making abilities, the learner must improve his or her HOTS. They can be taught effectively and efficiently with instructional designs, which support the specific requirements of teaching the strategies and methods used in these other disciplines.

The requirements for teaching HOTS are instructional approaches designed to promote development of thinking skills including specific mental operations, redirection, probing, reinforcement, asking higher-order questions, and lengthening wait-time. Along with emphasizing higher-order thinking strategies that include engaging learners in some form of mental activity, examining that mental activity, challenging the learner to explore other ways to accomplish the task or solve the problem, and then having the learner determine which way is best. HOTS also involve customizing the examination and exploration of mental activity to meet the individual learning needs in an active process, constructing knowledge as a result of interaction with the physical and social environment in student-centered classrooms where students are involved in collecting and analyzing information, paired problem solving, cooperative learning settings, simulations, debates, and critical reporting sessions. These strategies could be guided discovery, expository teaching, or a combination strategy presented in a PBL design.

The transference of knowledge from the learning environment to its practical application is not a separate problem from the learning and application of HOTS. It is the problem addressed in the literature for teaching HOTS. That is, learning the cognitive skills underlying the decision-making process in a learning environment so that they may be recalled from LTM and applied when an ill-defined, ill-structured, complex problem solving situation occurs in aviation is the instructional challenge besetting the development of HOTS.

Near-transfer teaching methods involving practicing or drilling step-by-step procedures may be used because there is little variance in the application of the procedure. For example, in an aircraft abnormal or emergency situation, where no real thought is required to handle the situation and a simple "maintain aircraft control, identify, verify, and then complete the appropriate checklist" will do. In contrast, far-transfer teaching methods may be needed in other abnormal or emergency situations, particularly with ill-structured, ill-defined, and complex problems where the pilot must use extensive judgment, must use a different approach, or there are no set steps or set procedure in deciding what to do. Developing authentic and realistic problems with very similar circumstances to those occurring in-flight will promote near-transfer. On the other hand, when problems have somewhat different circumstances, judgment, or unique problem solving is required, then training methods are needed to promote far-transference. The training methods to promote far-transference include beginning with near-transference type problems, progressing toward more abstract and complex problems, and finally, continuously relating each problem to the environment where these ill-defined, ill-structured, and complex problems will be encountered.

How HOTS are learned can be explained and supported, in part, with a blend of learning theories including behavioral, cognitive, and constructivism. The behavioral learning theory supports learning the mental process and the procedural processes employed in normal and in the typical abnormal and emergencies situations. However, the behavioral learning theory provides little support to how ill-defined, ill-structured, and complex problem solving is learned. These problems are better explained with learning theories that support cognitive learning, namely cognitive and constructivist theories.

In some situations the learning process is driven by information-processing which emphasizes perception and attention, encoding of information, memory, comprehension, active learning, motivation, locus of control, mental models, metacognition, transfer of learning, and individual differences. Moreover, in other situations, the process will be driven by individual construction of meaning, situated learning, and collaborative learning. Hence, mixing instructional designs that are based on different learning theories should allow the educator to take advantage of the strengths of each learning theory and enhance the development and transfer of the cognitive skills beyond any single based design.

Enhancing the development and transfer of HOTS is influenced by the learning theories' ability to accommodate the requirements of teaching these skills. Cognitive skills are learned like other cognitive knowledge and should be based on cognitive and constructivist learning principles. Specifically, information processing, knowledge structure, and situated learning theories underpin the development and transfer of these skills; thus, they promote the development and transfer of the necessary thinking skills needed to solve ill-defined, ill-structured, complex problems.

This paper suggested that pilot judgment, decision-making, and critical thinking could be taught by teaching higher-order thinking skills. Higher-order thinking skills are taught by emphasizing the methods and strategies to teach cognitive skills with problem-based learning instruction. Cognitive skills are learned like other cognitive knowledge and should be taught from simple to complex and from concrete to abstract. The paper did not elaborate on problem-based learning instructional designs, which is the subject of subsequent papers on improving pilot thinking skills. The paper attempted to establish that the methods and strategies used in teaching higher-order thinking skills could be used in aviation to teach pilot judgment, decision-making, and critical thinking. Improvements in thinking skills should lead to reducing the number of aviation accidents by reducing the number of pilot bad judgments.

References

Advisory Circular 60-22 (AC60-22) (1991). *Aeronautical decision making.* (NTIS No. DOT/FAA/AC60-22).

Alessi, S. M. & Trollip S. R. (2001). *Multimedia for learning: Methods and development*. 3rd ed. Boston: Allyn and Bacon.

Alvino, J. (1990). A glossary of thinking-skills terms. *LEARNING 18*/6, 50.

Aviation Instructor's Handbook (1999). (NTIS No. FAA-H-8083-9). Washington, DC: U.S. Government Printing Office (GPO).

Barrett, E. C. (2003). *Cognitive learning theory.* Retrieved March 15, 2004, from Syracuse University, Instructional Design, Development & Evaluation Department web site http://web.syr.edu/~ebarrett/ide621/cognitive.htm

Behaviorism (2000). Retrieved March 14, 2004, from http://www.funderstanding.com/behaviorism.cfm

Bell, B. & Mauro, R. (1999). *Training in judgment and aeronautical decision-making.* Retrieved June 24, 2003 form http://www.comet.ucar.edu/aviationsafety/judgement_aero_dm.pdf

Buch, G. D., & Diehl, A. E. (1984). An investigation of the effectiveness of pilot judgment training. *Human Factors*. October, Vol. 26 No. 5.

Carbenell, L. (2001). *Behaviorism: What kind of questions do the behaviorists ask?* Retrieved March 14, 2004, form http://www.my-ecoach.com/idtimeline/behaviorism.html

Carnegie, C. (2002). *Flight performance – level 3: Human factors-section 6*. Retrieved June 30, 2003, from Florida International University, Aeronautics Learning Laboratory for Science Technology and Research (ALLSTAR) web site, Http://www.allstar.fiu.edu/aero/HumFac06.htm

Carr, Kathryn S. (1990). *How can we teach critical thinking?* Retrieved June 24, 2003, from University of Illinois at Urbana-Champaign, ERIC Clearinghouse on Elementary and Early Childhood Education (ERIC/EECE) web site, http://ericps.ed.uiuc.edu/eece/pubs/digests/1990/carr90.html

Clark, R. C. (1999). *Building expertise: Cognitive methods for training and performance improvement.* Washington DC: International Society for Performance Improvement.

Cohne, M. & Freeman, J. (1996). Thinking naturally about uncertainty. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting.* (pp. 179-183). Santa Monica, CA

Connolly, T. (1990). Pilot decision-making training. AFHRL Technical Paper 88-67. Daytona Beach, FL. (NTIS No. ADA 221349).

Cotton, K. (1991). Teaching higher-order thinking skills. Retrieved January 11, 2003, from Northwest Regional Educational Laboratory, School Improvement Research Series (SIRS) web site http://www.nwrel.org/scpd/sirs/6/cu11.html

Diehl, A. E. (1991, November) Does cockpit management training reduce aircrew error? *Paper Presented at the 22nd International Seminar International Society of Air Safety Investigators.* Canberra, Australia. Retrieved January 14, 2003 from http://s92270093.onlinehome.us/crmdevel/resources/paper/diehl.htm

Deitch, E. L. (2001). Learning to land: A qualitative examination of pre-flight and in-flight decision-making process in expert and novice aviators. Retrieved June 24, 2003, from Virginia Polytechnic Institute and State University, Digital Library and Archives web site http://scholar.lib.vt.edu/theses/available/etd-12122001-183807/unrestricted/00-Contents.pdf

Driscoll, Marcy P. (2000). *Psychology of learning for instruction* (2nd ed.), Needham Heights, MA: Allyn & Bacon.

Ennis, Roberts H. (2000). *A supper-streamlined conception of critical thinking.* Retrieved November 11, 2003, from http://www.criticalthinking.net/SSConcCTApr3.html

Howe, R. W. & Warren, C. R. (1989). *Teaching critical thinking through environmental education.* (ERIC Identifier: ED324193)

Huitt, W. (2004). The cognitive system. *Educational Psychology Interactive*.

Valdosta, GA: Valdosta State University. Retrieved March 14, 2004, from http://chiron.valdosta.edu/whuitt/col/cogsys/cogsys.html.

Information Processing Theory (1996). Retrieved March 14, 2004, from http://www.educationau.edu.au/archives/cp/04h.htm

Jensen, R. (1988). Creating a'1000 Hour' pilot in 300 hours through judgment training. In *Proceedings of the Workshop on Aviation Psychology*. Newcastle, Australia: Institute of Aviation, University of Newcastle.

Johnson, S. D. & Thomas, R. (1992). Technology education and the cognitive revolution. *Technology Teacher 51*, No. 4 (January 1992): 7-12 (EJ435370).

Kearsley, G. (n.d.). Operant Conditioning (B.F. Skinner). *Explorations in Learning & Instruction: The Theory Into Practice Database.* Retrieved March 14, 2004, from http://tip.psychology.org/skinner.html

Kerka, S. (1992). Higher Order Thinking Skills in Vocational Education. (ERIC Digest No. 127). ERIC Clearinghouse on Adult, Career, and Vocational Education, Columbus, Ohio. Office of Educational Research and Improvement (ED), Washington, DC. (ERIC Document Reproduction Service No. ED350487)

Landa, Lev N. (1999). Landamatics instructional design theory and methodology for teaching general methods of thinking. In Charles M. Reigeluth (Ed.). *Instructional-design theories and models: A new paradigm of instructional theory* (Vol. 2, pp. 341-369). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Lave, J. (1996). Situated learning. *Learning Theories: Situated Learning* [on-line]. http://www.educationau.edu.au/archives/CP/04k.htm

Millhoff B. L. (2002). *Cognitive learning theory terms.* Retrieved March 15, 2004, from University of Guam, College of Education Course 451 web site http://www.uog.edu/coe/ed451/tHEORY/learnterm_c.html

Murphy, E. (1997). *Constructivism: From philosophy to practice.* Retrieved March 14, 2004, from http://www.stemnet.nf.ca/~elmurphy/emuphy/cte2b.html

Operant Conditioning (1996). *Learning Theories: Situated Learning* [on-line]. http://www.educationau.edu.au/archives/CP/04j.htm

Peirce, W. (2001) *Strategies for teaching thinking and promoting intellectual development in online classes.* Retrieved January 11, 2003 from Maryland Consortium of Community Colleges for Teaching Reasoning, Reasoning Across the Curriculum Program at Prince George's Community College. http://academic.pg.cc.md.us/~wpeirce/MCCCTR/ttol.html

Perkins, D. N. (1991, May). Technology meets constructivism: Do they make a marriage? *Educational Technology, 31*, 18-23.

Presseisen, B. Z. (1986). *Critical thinking and thinking skills: State of the art definitions and practice in public schools.* (ERIC Identifier: ED268536).

Reigeluth, C. M. (1999). What is instructional-design theory and how is it changing? In C. M. Reigeluth (Ed.). *Instructional-design theories and models: A* *new paradigm of instructional theory* (Vol. 2, pp. 5-29). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Reigeluth, C. M. & Moore, J. (1999). Cognitive education and the cognitive domain. In C. M. Reigeluth (Ed.). *Instructional-design theories and models: A new paradigm of instructional theory* (Vol. 2, pp. 51-68). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Ristow, R. S. (1988). The teaching of thinking skills: Does it improve creativity? *Gifted Child Today 11*/2, (pp. 44-46).

Ryder, J. & Redding, R. (1993). Integrating cognitive task analysis into instructional systems development. *Educational Technology Research and Development. 41* (2). 1042-1629.

Shebilske, W., Regian, J., Winfred, A., & Jordan, J. (1992). A dyadic protocol for training complex skills. *Human Factors, 34* (3). 369-374.

Splitter, L. J. (1995). *On the theme of "teaching for higher order thinking skills."* Retrieved June 24, 2003 from http://www.chss.montclair.edu/inquiry/summ95/splitter.html

Thomas, M. and Albee, J. (1998). *Higher-order thinking strategies for the classroom.* Retrieved June 23, 2003, from http://members.aol.com/MattT10574/HigherOrderLiteracy.htm

Thomas, R. G. (1992). *Cognitive theory-based teaching and learning in vocational education.* Information Series No, 349. Columbus: OH (ERIC Identifier: ED345109).

What is higher order thinking? (n.d.). Hammond, LA: Southeastern Louisiana University, Department of Teaching and Learning. Retrieved January 11, 2003, from http://www.selu.edu/Academics/Education/TEC/think.htm

Wiggins, M. (1997). Expertise and cognitive skills development for ab-initio pilots. In R. a. Telfer, & P. J. Moore (Eds.). *Aviation Training: Learners, Instruction, and Organization.* Brookfield, VT:Avebury.

International Journal of Applied Aviation Studies, 4, Number 2

Preliminary Implications for Academic Professionals of Aviation Student Myers-Briggs Type Indicator (MBTI) Preferences

Mary N. Kutz

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

Dovie M. Brown

16108 Silverado Drive
Oklahoma City, OK 73013
Phone: 405-348-1388
Email: deebrown1@cox.net

David B. Carmichael

1136 Fox Lake Lane
Edmond, Oklahoma 73034-7309
E-Mail Address: dcarmichael2@cox.net
Phone: 405-348-8050

and

Mahmood Shandiz

11113 Woodbridge Road
Oklahoma City, OK 73162
E-Mail Address: mshandiz@okcu.edu
Phone: 405-720-2420

Abstract

This article is based on some of the results of previous research conducted in 2003 for the purpose of looking at the Myers-Briggs

---

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530, P.O. Box 25082, Oklahoma City, OK 73125.

Type Indicator (MBTI) preferences of college students, specifically aviation majors and business majors. Another aspect of the research compared the preferences of aviation Professional Pilot undergraduate majors with the preferences of Aviation Management undergraduate majors. The findings of the study revealed that the ESTJ aviation management students and the ESTP professional pilot students share three of the four preferences. Although the two groups of students have much in common, the single letter of difference could have meaningful implications for aviation instructors in developing effective teaching and learning strategies in the classroom.

Introduction

An effective instructional strategy is often essential to learning retention, especially in the aviation classroom. Student learning in the aviation environment with its numerous subspecialties and associated complexities is affected by not only the unique personal preferences and attitudes of students, but also their individual responsiveness to instruction which may be contingent upon the preferences and style of the instructor.

Woods (1993) suggested that Bloom's taxonomy, Perry's model of attitude toward learning, and Myers-Briggs Type Indicator (MBTI) typology be combined to enhance current and future teaching/learning strategies targeted toward a specific audience. Based on that approach, an understanding of aviation student MBTI preferences might be the logical place from which to launch future inquiry. Numerous studies of the MBTI with students in other occupational specialties such as engineering have been helpful in understanding the impact of career-specific preferences on learning (Thomas, 1998). If research supports the notion that there is a tendency for persons with the same MBTI type to exhibit similar career interests, MBTI preferences could serve as the scaffolding for a slowly evolving mosaic of the aviation student. Aviation instructor preferences and associated instructional strategies combined with knowledge of aviation student preferences could offer valuable implications for academic professionals in the field. An early effort to examine MBTI preferences of aviation students at Oklahoma State University (OSU) produced some interesting preliminary findings in that regard.

Statement of the Problem and Purpose

The MBTI remains one of the most widely used inventories in many professions and occupational settings and has traditionally focused on education and counseling (Lauderdale & Thomas, 1994). However, at this point it has not been as widely used in the field of aviation education to individualize learning and aid classroom instructors in this highly volatile field. There has been some focus on learning styles and even preferences of aviation students in such fields as crew resource management. However, an overall inventory of aviation student preferences using the MBTI across aviation subspecialties potentially offers valuable insight for academic professionals. An improved understanding of the modern aviation student across a variety of dimensions including the way they focus attention, the way they take in information, make decisions and orient themselves to the environment may be important to aid in improving classroom instructional techniques. That understanding also sets the stage for further research with other reliable instruments in order to put together a clearer picture of the aviation student and those strategies that would most effectively engage them in meaningful learning experiences.

Psychological Type Theory and the MBTI

The MBTI is based in part on the concepts of psychological type developed by Jung and the later work of Briggs and Briggs-Myers. The Myers-Briggs Type Indicator, or MBTI as it is most commonly called, and the acronyms associated with its 16 personality types have become familiar to many as a way of examining preferences and differences in people. A single type is comprised of a combination of four preferences identified by the MBTI as follows: (E) for *Extravertion* or (I) for *Introversion* used to describe how people gain their energy; (S) for *Sensing* or (N) for *Intuition* used to describe how people take in information; (T) for *Thinking* or (F) for *Feeling* describes how people make judgments or reason; and (J) for *Judging* or (P) for *Perceiving* which deals with attitude and how people orient themselves to their environment. Each of the separate types has no significance pertaining to good or bad qualities, but does have significance in terms of differences in preference which can have significant meaning for educators as do the various combinations in their implications for classroom teaching strategy (Myers, 1991).

Oklahoma State University Aviation Student Preferences

An attempt at understanding MBTI scores of aviation students began in 2003 with a small study of the MBTI scores of students enrolled in the Professional Pilot undergraduate degree option and students enrolled in the Aviation Management undergraduate degree option at Oklahoma State University. The preliminary findings of that study indicated that there was no significant difference between the aviation management and pilot training students on focusing attention, information taking, decision making, and environment orientation indicators (see Table 1). The OSU study represented a sample of 118 aviation students from a small population of 176, which were randomly selected from aviation management and professional pilot classes. It was the first in a series of samples to be collected from aviation student populations. Although the data represented only an early beginning of much needed long range research, the preliminary data indicated that aviation students in both major options had similar dominant preferences with only slight difference in the environmental orientation indicator. Although the significance of those findings will be known only after much more extensive research, the tentative implications are interesting and somewhat enlightening.

Table 1

Relation between Aviation Students' Major and MBTI-Focusing Attention Indicator

Major/EI	Extraversion	Introversion	Total
Management	30	24	54
Pilot	35	29	64
Total	65	53	118

$X^{2=}.0089$, df=1, p=.92

The overall dominant MBTI type of the professional pilot majors was ESTP *(Extroversion, Thinking, Sensing, Feeling)*, compared to ESTJ *(Extroversion, Thinking, Sensing, Judging)* for aviation management students. The dominant indicator of how aviation students of both major options focus attention was *Extraversion*, representing well over 55 percent of the respondents for both major options. The dominant indicator of how aviation students of both options take in data was *Sensing*, once again representing over 55 percent of the students sampled. The decision-making preference indicator of aviation management students was 57 percent *Thinking*, while the pilot indicator was 64 percent *Thinking.* It was the environmental orientation indicator that showed a slight difference in that approximately 52 percent of the aviation management students used *Judging* as their dominant preference, compared to 59 percent of the pilots who used *Perceiving* as their dominant preference.

Implications of Preliminary Findings of Aviation Student MBTI Profiles

The preliminary findings of this study pointed to the possibility that modern aviation students, regardless of major option, may prefer to focus their attention and find energy in things and people. As *Extraverts*, they prefer interaction with others and are action oriented. The Center for Applications of Psychological Type (CAPT) reported that the majority of university faculty are *Introverts* preferring to find energy in the inner world of ideas, concepts and abstractions (Brightman, 2003). Teachers of an opposite dominant preference may struggle with allowing opportunities for *Extraverts* in the classroom to utilize discussion and possibly provide opportunities for students to explain it to themselves and others in an effort to ensure that they understand the material. Teachers might consider providing opportunities to work in groups, either inside or outside of class. It is important to recognize that even though *Extraversion* may be the dominant way aviation students process data, the remaining 44 percent of the students prefer *Introversion* and need some internal processing time to connect material and see the larger picture.

Table 2

Relation between Aviation Students' Major and MBTI Information Taking Indicator

Major/SN	Intuition	Sensing	Total
Management	24	30	54
Pilot	28	36	64
Total	52	66	118

$X^{2=}0.0057$, df=1, p=.94

Aviation students in the OSU study were predominantly *Sensing* in that they preferred to take in information through their senses. This is consistent with the general population of students in other fields in that CAPT data indicated that the majority of undergraduates are *Sensing students.* On the other hand, the majority of university faculty fall into the *Intuitive* category (Brightman, 2003).

Sensing students like detail gleaned from their senses, while *Intuitive* students look for the big picture and watch for patterns and relationships. It is important to teach students how to relate material to previously learned material by comparing and contrasting to real-life problems or cases. *Intuitive* students need that linkage but *Sensing* students can also learn to appreciate it. The tools needed for *Intuitive* types are helpful to both (Myers, McCaulley, Quenk, & Hammer, 1998).

The preliminary results of this initial study at OSU pointed to the possibility that aviation students across the board are more *Thinking* oriented than *Feeling.* They prefered to make decisions impersonally based on analysis and logic, while *Feeling* students prefer to make decisions based on emotion or human values and needs. CAPT data bases indicated that the majority of university faculty also have preferences for *Thinking*. In this one dimension, the dominant preference of student and faculty are similar (Brightman, 2003).

Table 3
Relation between Aviation Students' Major and MBTI Decision Making Preference Indicator

Major TF	Feeling	Thinking	Total
Management	23	31	54
Pilot	23	41	64
Total	46	72	118

X^2= 0.5453, df=1, p=.46

Thinking students like clear, practical objectives, while *Feeling* students like group and team activities. Although there is a clear dominance in this dimension, it is important for instructors to remember that the *Feeling* students need some group and team activities for meaning and retention. It is important also for *Thinking* aviation students to be exposed to the values and feelings of people in order to raise awareness of the emotional side of aviation.

One of the dominant complaints from aviation leadership in the Kutz (1998) study was that students leaving the educational arena come with weak communication skills and interpersonal skills. Aviation students often are quite candid that they just want to fly and often do not recognize that their technical background must include team and communication skills – even in the air. For this reason, crew resource management classes have become required curriculum in most aviation academic and training environments. Requiring *Thinking* students to conduct interviews or work in teams to solve problems necessitates that they apply both knowledge and understanding of human values to a particular problem or exercise. They learn to deal with both logic and subjective feeling orientations that they will ultimately face in the real world. Students must learn that decisions are not always best made using logic but may require consideration of the feeling and experiences of others.

It was in their orientation to the environment that aviation management, students and professional pilot students parted company somewhat. One of the surprises of the preliminary findings of this study was that the majority of aviation management students were *Judging* or decisive and deadline oriented, while the majority of pilots were *Perceptive* or spontaneous and liked to think in terms of possibilities. Deadlines for *Perceivers* often are stretched while they procrastinate and seek more data. CAPT reported that the majority of university faculty have a preference for *Thinking*. On this preference, they are similar to the aviation management types, but opposite of pilot types. One method recommended for teachers in dealing with *Perceptive* students is to break assignments into chunks with several shorter deadlines to keep the *Perceivers* on target and provide intermediate feedback (Myers, 1991).

Table 4
Relation between Aviation Students' Major and MBTI Environmental Orientation Indictor

Major/JP	Judging	Perceiving	Total
Management	28	26	54
Pilot	26	38	64
Total	54	64	118

X^2=1.4873, df=1, p=.22

The ESTJ preferences of aviation management students and the ESTP preferences of professional pilot students of the OSU study were consistent with the findings of Tieger and Barron-Tieger (2001) in their commentary on career choices. They found that ESTP's like work that is active, full of adventure and fun, and allow for risk taking. They also like to respond to unplanned situations using unconventional approaches, where they can negotiate satisfactory solutions. They found that flight instructors, flight engineers and commercial helicopter pilots fit into this profile.

Tieger and Barron-Tieger (2001) found that ESTJ's enjoy careers that let them work systematically and use time and resources efficiently to achieve a logical conclusion. They like straight-forward assignments that allow them to using strong reasoning powers. They found that ESTJ's fit well as a project manager, administrator, supervisor, compliance officer or in other occupations that require organization.

At first glance one would assume from the findings of the OSU study that the ESTJ management students and the ESTP pilot students would have much in common as they shared three of the four preferences. However, even the single letter of difference could have meaningful implications for aviation instructors.

Elliott and Sapp (1988) found that ESTP college students generally prefer "collaborative" learning. They like to work with others and are flexible and tolerant. However, they quickly tire of theories and conceptual thinking. They want to take action to solve a problem and learn best by hands-on experience. They are more effective with application of learning through scenarios and simulations than with lectures and tests. Myers, et al (1998) stated that they do best in careers needing realism, action, and adaptability.

ESTJs, on the other hand, are decisive and realistic with little patience with inefficiency. They take command and provide leadership in order to get things done on schedule. They prefer structure in the classroom and to follow a schedule. They like immediate results and may jump to conclusions before all of the facts are in (Myers, et al, 1998).

Planning for instruction in the aviation classroom would require a strong mix of structure and collaborative learning to provide balance and maximize learning.

The Significance of the MBTI Profile

Although the preliminary data in this study is certainly not conclusive and the significance of it cannot be generalized to other populations, it provides a beginning baseline for an emerging profile of the aviation student population in two degree options at one of a limited number of universities offering aviation degrees. As the data is accumulated from subsequent populations of students and is expanded to faculty populations, the implications of those preferences will become clearer and more meaningful. In the meantime, it is important to understand that the MBTI was designed to understand the whole person rather than four distinct preferences and that whole person concept is extremely important to teachers to nurture the academic growth of students.

It is important also to remember that in every class, there are students who must work in their auxiliary style while teachers work with the dominant or preferred style of the other students. It is challenging enough for aviation students to learn the voluminous amount of new material they are required to learn without forcing some of them to continually operate in the discomfort of an auxiliary mode in the classroom. Faculty who offer a variety of learning activities improve the odds of making learning stick.

Faculty and students may have opposite preferences that make it especially difficult when the instructor must work outside his or her own preference. Nevertheless, it is important that the preferences of students be honored so that no one group of students is in the auxiliary mode at all times. For example, lecture for many professors is a preferred classroom instructional strategy. However, failure to recognize that extroverted students need opportunities for expression in order to learn, may result in burnout, absenteeism, and a variety of other symptoms of weak instruction.

Preliminary findings of the OSU study pertaining to aviation student preferences provided only a beginning or foundation for an emerging profile of the aviation learner. Research regarding teaching and learning preferences must ultimately be expanded well beyond the MBTI to develop a broad understanding of aviation students and incorporate learning differences found in the aviation classroom into modification of instructional materials and strategies which accommodate those differences.

References

Brightman, H. J. (2003). *GSU master teacher program: on learning styles*. Atlanta, GA: Georgia State University. Retrieved from http://www.gsu.edu/~dschjb/wwwmbti.html

Elliott, G. R. & Sapp, G. L. (1988). The relationship between the Myers-Briggs Type Indicator and the Grasha-Riechmann Student Learning Styles Questionnaire. *Journal of Psychological Type, 14*, 46-50.

Kutz, M. (1998). *Characteristics of successful aviation leaders of Oklahoma.* Unpublished doctoral dissertation. Stillwater, OK: Oklahoma State University.

Lauderdale, W. & Thomas, L. (1994). Applications for the Myers-Briggs Type Indicator. *Proceedings of the annual meeting of the Mid-South Educational Research Association.*

Myers, I. B. (1991). *Introduction to type. A description of the theory and Applications of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press, Inc.

Myers, I. B., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual, a guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press, Inc.

Thomas, E. L. (1998). A comparative study of 1975 and 1993 entering freshman Engineering students at Auburn University using the Myers Briggs Type Indicator. Unpublished doctoral dissertation. Auburn, AL: Auburn University.

Tieger, P. D. & Barron-Tieger, B. (2001). *Do what you are.* Boston: Little, Brown & Co., 305, 239-241.

Woods, D. (1993). Models for learning and how they're connected: relating Bloom, Jung, and Perry. *Journal of College Science Teaching, 22*, 250-254.

International Journal of Applied Aviation Studies, 4, Number 2

Book Review

Human Factors in the Training of Pilots
By Jefferson M. Koonce

Reviewed by
C Haley-Seikel
College of Continuing Education
University of Oklahoma
Norman, OK

and

Jim Hamm
College of Continuing Education
University of Oklahoma
Norman, OK

If allowed to define the word *disciplines* for this book review as academic and aviation technical branches of learning, then, *Human Factors in the Training of Pilots* (Koonce, 2002) is a book which allows readers from both of the defined disciplines to understand each other's point of view. Academic researchers have investigated human factor principles and have shown how they apply to the training environment. However, it has always been difficult for the aviation technical discipline to embrace the reasoning behind the research. Pilots have been the best at understanding the roles they perform in the air, but to transfer this knowledge to another has presented challenges for those who are certified to teach. Pilots have different techniques for accomplishing procedures. Researchers can identify the approach to teaching techniques. Joining both these disciplines into one have presented challenges with each branch using

---

Requests for reprints should be sent to Kay Chisholm, FAA Academy, AMA-530,
P.O. Box 25082, Oklahoma City, OK 73125.

their language and processes for reaching the end results. To transfer the "whys" and "why nots" in the aviation pilot training classroom in a consistent effective manner involves both disciplines.

There are numerous books available which talk about human factors, many of which focus on flight and ground instruction for pilots, but we have not seen one that provides such in-depth study of the many factors influencing learning in this environment. There is probably not a more hostile environment for learning than a hot, noisy, cramped, and sometimes scary training airplane. Because of this, the effective flight instructor needs to know every trick in the book to accomplish the training objective. There are many "How To Fly" books aimed at various levels of pilot training (from every imaginable technical perspective) but they lack the practical application viewed from the practitioners experiences. *Human Factors in the Training of Pilots* utilizes the greater good of both the *whys* and the *why nots*. Koonce indicates that his purpose for writing the book is "...with hope that it may develop an awareness of human factors principles and how to utilize the knowledge of human factors in the training and evaluation of pilots (p.i)." It is in the opinion of the reviewers that he met his purpose.

The book is neatly sectioned into two parts. The first part is titled Human Factors and the second- Applications and Hints From the Years. Additionally, the compiled references section is probably worth the price of the book alone. It lists everything from the necessary FAA Advisory and Circulars to the romance of Richard Bach's novels. Any conscientious pilot or aviation instructor wishing to develop their personal aviation library could use this as their checklist. In the Human factors section, Koonce describes and defines human factors using practical language and gives examples of how these work in aviation. The writings do not offend or overwhelm the reader with long and tiring academic terms. Instead, Koonce uses plain language to defend the academic points in an understandable and reasonable manner. This allows even the novice aviation technical reader to agree or disagree with the human issues within the learning arena. Even the definition of human factors was dissected and applied to the aviation technical discipline. Two other techniques Koonce uses in the human factors section writings is the novel/story telling process and truisms. Again, this enables the most novice reader (regardless of discipline) the ability to comprehend and enjoy this section.

As an example, on page 14, Koonce discusses what motivates someone to fly. He uses a poem which has circulated in the aviation field for over 50 years. An 8-year-old child wrote this poem. Koonce could have discussed in length extrinsic vs. intrinsic motivation and the research that supports both concepts, but instead he gets his point across with practical examples. In addition, throughout the book are aviation truisms. These small but pointed statements support theories and concepts. They tend to give the reader a sense of confirmation to the points. Everything from cognitive theories to mapping the retina are covered in the human factors section. In the process of discussing these concepts, one can quickly realize the purpose of knowing this information when applying these principles to aviation training.

The second section of the book (Applications and hints from the years) is filled with empirical insight for nearly every phase of flight from Preflight to How to Crash an airplane. The book is filled with valuable hints and tips that surely have been handed down by flight instructors through the ages.

This book is a must-read for all professional flight instructors and definitely a should read for any aviation enthusiast. However, no competent book reviewer could complete a review without at least one negative…so here it is. The title stinks! It sounds and looks too much like a textbook (I know, one could say all books are textbooks) and that just might keep the target readers from adding it to their shopping cart-what a shame that would be.