ATOP-R&D

Human Factors Newsletter

April 6, 2005

Special Edition

13th Annual Symposium on Aviation Psychology (ISAP): The 13th ISAP will be held in Oklahoma City at the Oklahoma City Convention Center April 18-21, 2005. The FAA Civil Aerospace Medical Institute (CAMI) will be the local host. Wright State University and General Dynamics Advanced Information Systems will handle the technical details of the meeting including conference registration. Major sponsors are expected to be General Dynamics, the U.S. Air Force Research Laboratory, FAA, NASA and others.

Aviation Psychology is the field of study concerned with the role of human operators in aviation systems. First convened by the Aviation Psychology Laboratory in 1981, this symposium series is offered for the purposes of:

- presenting the latest research on human performance problems and opportunities created by changes in aviation systems and technology
- envisioning design solutions that best utilize human capabilities for creating safe and efficient aviation systems
- bringing together scientists, research sponsors, and operators in an effort to bridge the gap between research and application

While the symposium is aerospace safety oriented, anyone with an interest in human performance and behavior will find the experience invaluable. The objective is to provide a forum for critical examination of the impact of high technology on the role, responsibility, authority, and performance of human operations in modern aircraft and air traffic control systems all over the world.

Below is a list of presentations scheduled for the symposium. Abstracts are attached to this newsletter.

- **The Effect of Terrain-Depicting Primary-Flight-Display Backgrounds and Guidance Cues on Pilot Recoveries from Unknown Attitudes**
- **The Good, the Not-So-Bad, and the Ugly: Computer-Detected Altitude, Heading, and Speed Changes**

- **Relationship Between Use of an ATC Decision Support Tool and Controller Opinions**
- **A Human Factors Taxonomy for Ground Operations: JANUS-GRO**
- **A Summary of Unmanned Aircraft Accidents: Human Factors Implications**
- **An ATC "Mental Workout" Program**
- **Land of the Midnight Sun: Shedding Light on Differences in GA Accidents in Alaska Versus the Rest of the United States**
- **The ATC Operational Error Severity Index: An Empirical Analysis of the Effects of Loss of Horizontal and Vertical Separation**
- **ATCS Age and En Route Operational Errors: A Re-Investigation**
- **An Overview of Human Factors Associated with Rotorcraft Accidents in the United States**
- **ATC Operational Errors: Exceeding the Limits of Cognitive Capacities**
- **Metrics of Information Complexity for Air Traffic Control Displays**
- **General Aviation VFR-into-IMC: *Z*-Score Filtering of Demographic and Personality Variables, and the Personality Paradox**
- **The Air Traffic Selection and Training Battery: What it Is and Isn't (and How it has Changed and Hasn't)**
- **The Role of Institutional Review Boards in Aviation Research: It's the Law and It Makes Sense**
- **Implementing Electronic Flight Data in the Airport Traffic Control Tower**
- **Surprise and Unexpectedness in Flying: Factors and Features**
- **Dealing with Unexpected Events in Aviation: The Role of Domain Expertise**
- **Discussion Panel - Automation and Performance**
- **Program Update and Prospects for In-Flight Simulation Upset Recovery Training**
- **Evidence Against Crew Resource Management as a Cognitive Skill**
- **Automatic Text Analysis of ASAP Narrative Reports**
- **How High is High Enough? Quantifying the Impact of Air Traffic Control Tower Observation Height on Distance Perception**
- **A Systematic Approach to Addressing Human Factors Considerations in the Design of Flight Deck Components**
- **Analysis Methods for Determining the Safety Consequences of Mixed-Fleet Flying**
- **Human Factors Design of Electronic Documents**
- **Special Session: Maintenance Training & Communication Technologies**
- **Special Session: Maintenance Risk Management**
- **Improving Novice Flight Performance Using a Functional Avionics Display**

More information on human factors research can be found at the FAA Human Factors (ATOP-R&D) web site: http://www.hf.faa.gov

Paul Krois
FAA (ATO-P R&D Human Factors)

Comments or questions regarding this newsletter?
Please contact Bill Berger at (334) 271-2928
or via e-mail at bill.ctr.berger @faa.gov

13th Annual Symposium on Aviation Psychology

Abstracts

The Effect of Terrain-Depicting Primary-Flight-Display Backgrounds and Guidance Cues on Pilot Recoveries from Unknown Attitudes. Dennis B. Beringer and Jerry D. Ball, Civil Aerospace Medical Institute (CAMI).

A study was conducted to determine whether presentation of terrain on an electronic primary flight display (PFD) would negatively impact pilots' recoveries from unusual attitudes, and to determine whether guidance cues would eliminate any potential ambiguity in display interpretation. Each of 40 pilots was assigned to one of five conditions consisting of combinations of terrain depiction (none, full-color, brown) and guidance indications (pitch and roll arrows).

Data were collected in the CAMI Advanced General Aviation Research Simulator in Piper Malibu configuration, and each pilot performed eight warm-up/familiarization recovery maneuvers using the basic electronic attitude-direction indicator (EADI) on the PFD. These trials were followed by 16 additional recoveries using the particular experimental PFD configuration assigned to each pilot. The eight maneuvers used consisted of (1) pitch up (20°); (2) pitch down (15°); (3/4) roll left or right (60°); (5/6) pitch up (20°) and roll left or right (60°); and (7/8) pitch down (15°) and roll left or right (60°). Half of the entry headings resulted in the recovery ending facing mountainous terrain higher than the aircraft altitude, and half ending the recovery facing level terrain lower than aircraft altitude. Pilot recovery times and initial response times were recorded for each trial. A recovery was considered completed when the pilot held the aircraft within ±2.5 degrees of zero-degrees pitch and ±5.0 degrees of zero-degrees bank for three seconds. All groups attained equivalent performance at the end of the familiarization period prior to starting the experimental trials.

Multivariate Analysis of Variance detected no significant differences between the displays during experimental trials. Univariate analyses were conducted to determine if type of maneuver resulted in any significant differences between display types. Again, no significant differences were found between displays and type of maneuver.

Additional trials were performed by most of the participants to address whether background terrain that was at a significantly higher elevation than the aircraft could cause recoveries to be made to the terrain horizon and not the zero-pitch line. Those participants flying a terrain-depicting display flew an additional pitch-up trial where the top of the mountainous terrain was 10 degrees above the zero-pitch line, but only one participant showed any indication of premature termination of the recovery. In addition, participants in all display groups flew one inverted recovery from approximately 165 degrees of bank, and there did not appear to be any consistent effects of display type on pilot performance.

In post-test interviews, pilots reported they were focusing their attention on the zero-pitch line, which was relatively prominent, and did not regard the terrain depictions, when present, as significant contributors to their recovery task. Many also made positive comments on the guidance cues despite the lack of any quantitative data to support favoring them. Researchers concluded that for this specific task, it appears the presence of a zero-pitch line of sufficient contrast (white with black borders) to all backgrounds allows pilots to perform recoveries adequately and that there appear to be no detrimental effects of the terrain depiction on unknown attitude recoveries.

The Good, the Not-So-Bad, and the Ugly: Computer-Detected Altitude, Heading, and Speed Changes. Elaine M. Pfleiderer, Civil Aerospace Medical Institute

Introduction. The relationship between communication events and controller workload has been well established. Unfortunately, a substantial amount of time and effort is required to transcribe and code these events. Alternative measures might be preferable if they could be obtained more easily. Manning, Mills, Fox, Pfleiderer, and Mogilka (2002) found that, relative to a set of computer-derived measures, communication events might not add enough unique information to the prediction of subjective workload to justify the effort involved in obtaining them. At the time the study was conducted, computer-derived measures of altitude, heading, and speed changes were not available. The present investigation compares altitude, heading, and speed clearances with computer-derived measures of altitude, heading, and speed changes.

Method. Two 20-minute samples of live air traffic data were collected from each of four sectors in the Kansas City en route airspace. Communications data were transcribed from audio recordings and coded (e.g., altitude, heading, and speed clearances). Altitude, heading, and speed changes for each of the 20-minute samples were computed from System Analysis Recordings (SAR) using the Performance and Objective Workload Evaluation Research (POWER) software system. The 20-minute samples were parsed into four-minute intervals, and the number of communications events and changes were tallied for each interval. In addition, Air Traffic Workload Input Technique (ATWIT) measures were provided by 16 subject-matter experts for each 4-minute interval for all samples.

Results. Multiple regression analysis of altitude, heading, and speed clearances on mean ATWIT scores yielded an $R = .59$ ($R^2 = .35$). Multiple regression of the number of computer-detected altitude, heading, and speed changes on mean ATWIT scores yielded the same results. Hierarchical multiple regression, employed to examine shared and unique variance of the two sets of measures, revealed that altitude changes alone could account for most of the variance in ATWIT scores ($R = .57$; $R^2 = .32$).

Discussion. Results suggest that computer-derived measures of altitude and heading changes may be a viable substitute for more labor-intensive assessment of communications events. However, ground speeds recorded by the Host Computer (and displayed on the controllers' radarscope) were too erratic to provide a valid measure of speed changes, and could not be recommended as an acceptable alternative for speed clearances. Concordance rates of clearances and computer-detected measures are discussed.

Relationship between Use of an ATC Decision Support Tool and Controller Opinions.
Carol A. Manning and Melanie Dennis. Civil Aerospace Medical Institute.

The purpose of this study was to assess the relationship between controllers' opinions about the User Request Evaluation Tool (URET), an air traffic control decision support tool, and their use of certain URET functions relative to alternate methods for obtaining and recording information. We expected that, compared with those who were less positive about URET, controllers who rated the tool more positively would use it more often and would use flight strip equivalents less often.

In 2002, 181 en route controllers using URET at six facilities were formally observed. URET display settings, usage, and use of flight strip equivalents were recorded. Controllers were also asked their opinions about the readability/usability of URET, changes in roles and communications between controller team members, their typical use of URET features, their perceptions of URET's effects on safety, workload, time required to perform tasks, and benefits provided to pilots. Dichotomous answers to opinion questions were used as independent variables in ANOVAs. Dependent variables were counts of activities performed using paper, the Host Computer, or URET.

Controllers who were observed to perform more URET tasks thought the system required less time to use and were more positive about its effect on safety than those observed to perform fewer URET tasks. They were also more likely than those who performed fewer URET tasks to indicate that they checked alerts and performed trial planning. Use of specific aircraft list (ACL) functions was also related to controllers' likelihood of using the URET decision support tools. For example, those who deleted gray entries from the ACL and annotated speeds and headings were also more likely than those who did not to check alerts. Moreover, Radar Controllers and Radar Associate Controllers who highlighted entries on the ACL were more likely to use trial planning than those who did not highlight entries on the ACL.

It is unclear whether controllers' familiarity with URET resulted in their positive opinions about the system or if having positive impressions made them want to use the system more frequently. It is also unclear whether increased use of the system would change controllers' opinions about

URET. Regardless, these results indicate that there is a relationship between system usage and positive opinions.

A Human Factors Taxonomy for Ground Operations: JANUS-GRO. Alfretia Scarborough, Julia Pounds, and Larry Bailey. Civil Aerospace Medical Institute.

Introduction. Since the FAA began tracking runway incursion data in 1988, vehicle and pedestrian deviations (VPDs) have accounted for an average of 21%. Although there are fluctuations in the percentages of VPDs for any given year, the overall trend for the past 15 years has remained fairly flat. Although considerable effort has been spent at addressing runway incursions (such as the airport surface detection equipment-Model 3 (ASDE-3) and the airport movement area safety system - AMASS), relatively little attention has been given to identifying the human factor causes of VPDs. We report on the initial development of a human error taxonomy designed to classify vehicle operator (VO) deviations based on their human factor causes. The taxonomy represents an extension of an existing air traffic control (ATC) operational error (OE) taxonomy to ground operations. The ATC OE taxonomy is called JANUS-ATC and the extension on which we report is called JANUS-GRO.

Method. Four volunteers, one each from the four ground operational specialties at the Will Rogers World Airport (Airport Operations, Airport Maintenance, FAA Airway Facilities Transportation System, and Aircraft Rescue and Fire Fighting) volunteered as subject matter experts (SMEs). The SMEs had an average of 12 years experience in their respective domains. The task for the SMEs was to review the JANUS-ATC taxonomy and to make modifications so that the taxonomy reflected a given SME's operational specialty. SME modifications involved redefining the JANUS ATC definition, adding categories/terms, and eliminating categories/terms that did not apply to a given ground operation. As an operational test, the JANUS-GRO taxonomy was used to categorize 166 VPD narratives for the two-year period 2000-2001.

Results and Discussion**.** Like JANUS-ATC, JANUS-GRO consists of two broad error categories: (1) factors directly related to VO performance, and (2) contributing factors that indirectly influence VO performance. Direct performance factors consisted of the task being performed, the mental processes that were involved, and the VO's compliance with standard operating procedures (SOP). Contributing factors consisted of organizational and supervisory influences as well as the contextual conditions that shape the VO's performance (such as ground traffic mix, equipment, and airport configuration). Over the two-year period, VO deviations accounted for 134 (81%) of the incidents. Of these, 96% were classified as direct performance factors, and four percent were classified as indirect performance factors. When the direct performance factors were further analyzed, 95% (122/129) were associated with failure to observe SOPs, and five percent (7/129) were associated with the mental processes of perception and vigilance. Of the indirect performance factors, 50% (2/4) were associated with VO training and 50% (2/4) were associated with equipment malfunctions.

Conclusion**.** The JANUS-GRO adaptations centered on the specificity of the VO tasks, the contextual framework surrounding the tasks, and the supervisory/organizational factors unique to a given VO. An initial JANUS-GRO analysis of VO deviations revealed that almost all are related to a failure to observe SOPs. However, it was unclear whether the noncompliance was

due to intent or whether they were the result of memory loss or inappropriate planning and decision-making. This points to the need for more in-depth investigations so that effective interventions can be developed that would serve to mitigate the factors underlying VO deviations.

A Summary of Unmanned Aircraft Accidents: Human Factors Implications.
Kevin W. Williams, Civil Aerospace Medical Institute.

The review and analysis of unmanned aircraft (UA) accident data can assist researchers in identifying important human factors issues related to their use. The most reliable source for UA accident data currently is the military. The military has a relatively long history of UA use and has always been diligent in accurately recording information pertaining to accidents/incidents. The purpose of this research was to review all currently available information on UA accidents and identify human error aspects in those accidents and what human factors issues are most involved.

Two primary sources of accident information were collected from the U.S. Army. The first was a summary of 56 UA accidents produced by the U.S. Army Aeromedical Research Laboratory and obtained from the U.S. Army Risk Management Information System (RMIS). The second was a direct query of the RMIS system of all UA accidents that occurred between January 1986 and June 2004. A total of 74 accidents were identified, the earliest of which occurred on March 2, 1989, and the latest on April 30, 2004.

Information regarding UA accidents for the U.S. Navy was collected from the Naval Safety Center. A summary of 239 UA mishaps occurring between 1986 and 2002 was received from the Naval Safety Center in Pensacola, FL (Kordeen Kor, personal communication).

Air Force accident/mishap information was collected from the Air Force Judge Advocate General's Corps Web site, http://usaf.aib.law.af.mil/. A total of 15 Class-A UA mishaps were retrieved from the Web site, covering the dates from December 6, 1999, to December 11, 2003. In addition, a complete accident investigation board report was received. Also, a summary of Air Force accidents and human factors issues was received electronically from Major Anthony P. Tvaryanas, the author.

Classification of the accident data was a two-step process. In the first step, accidents were classified into the categories of human factors, maintenance, aircraft, and unknown. Accidents could be classified into more than one category. In the second step, those accidents identified as human factors-related were classified according to specific human factors issues of alerts/alarms, display design, procedural error, skill-based error, or other. Classification was based on the stated causal factors in the reports, the opinion of safety center personnel, and personal judgment of the author.

The percentage of involvement of human factors issues varied across aircraft from 21% to 68%. For most of the aircraft systems, electromechanical failure was more of a causal factor than human error. One critical finding from an analysis of the data thus far is that each of the fielded systems is very different, leading to different kinds of accidents and different human factors

issues. A second finding is that many of the accidents that have occurred could have been anticipated through an analysis of the user interfaces employed and procedures implemented for their use. The current paper summarizes the various human factors issues related to the accidents.

An ATC "Mental Workout" Program. Julia Pounds, Randy Breedlove, Deborah Thompson, Dan Jack, Anthony Ferrante. Civil Aerospace Medical Institute.

Introduction. Despite air traffic controllers' reliance on perceptual and cognitive skills to perform their job, little is known about whether these skills can be further improved after a controller completes training and becomes proficient. This project evaluated the first National Air Traffic Professionalism (NATPRO) program, a new approach for operational controllers to test their attention skills. Previous research has shown that persons with self-reported experience playing video games were faster at routing and handing off aircraft in scenarios in a low-fidelity ATC task (Young, Broach, & Farmer, 1997). More recently, video game play was demonstrated to alter and improve a range of visual skills (Green & Bavelier, 2003). This study was designed to evaluate the potential performance gains associated with skill practice.

Method. A seminar and corresponding computer practicum focusing on attention processes were completed by 25 operational controllers in an FAA air route traffic control center. A pretest was administered prior to the seminar to program participants and 25 controllers who volunteered to serve as the control group. The post-test was administered after completion of three weeks of distributed practice of approximately one hour per day.

Results. Results showed that, compared to the control group, NATPRO participants improved their scores on memory (increasing 35%) and mental calculation tasks (increasing 16%) while performing the two tasks simultaneously with distractions. Participants also reported gaining knowledge about their mental skills, and anticipated benefits from further knowledge and skills practice.

Conclusion. The findings suggest that this approach to skill enhancement could add value to existing ATC programs. Expansion of this program to more ATC facilities is planned and program evaluation will continue. Efforts are underway to identify additional topics for the NATPRO program.

Land of the Midnight Sun: Shedding Light on Differences in GA Accidents in Alaska Versus the Rest of the United States. Cristy Detwiler, Albert Boquet , Douglas Wiegmann, Scott Shappell. Civil Aerospace Medical Institute.

Introduction. It has been said that people in Alaska fly private aircraft like we take taxis in the lower 48 states. Indeed, for some parts of Alaska, the only access may be by air. This presents quite a challenge to the general aviation (GA) pilot, given that Alaska is well known for its varied and often unique landscape, including but not limited to, seemingly endless mountains, glaciers, lakes, long coastlines, volcanoes and fjords. When this veritable obstacle course is considered with the temperamental weather and seasonal lighting conditions, even the most seasoned aviator will have to agree that Alaskan aviation represents some of the most difficult

flying in the U.S., if not the world. To address flying in this unique environment, several programs have been proposed, developed, and/or implemented to provide pilots in Alaska with improved technology and decision aids (e.g., Capstone and weather cameras to name a few). While these efforts will likely have an impact on Alaskan aviation safety, the question remains to what extent they will impact the accident rate and whether there are any human factors issues not being addressed by these programs. Prior research has typically focused on the demographics of aviation crashes (e.g., age, gender, experience), with little emphasis on specific human errors associated with GA accidents. With the introduction of the Human Factors Analysis and Classification System (HFACS), it is now possible to systematically examine the specific types of human error being committed by these aviators.

Method. Five independent pilot-raters, using HFACS, examined over 15,500 GA accidents having at least one aircrew error. Comparisons of Alaska with the rest of the U.S. (RoUS) included traditional demographic and environmental variables, as well as the human errors committed by aircrews.

Results. Overall, differences in the types of unsafe acts (decision errors, skill-based errors, perceptual errors, and violations) committed by pilots involved in accidents in Alaska and those in the RoUS were minimal. Skill-based errors continue to have the greatest impact on the accident data, resulting in an average of 70 to 80 percent of the accidents being associated with at least one skill-based error. However, a closer inspection of the data revealed notable variations in the specific types of unsafe acts that occurred. For instance, the decision to land, take-off, or taxi from unsuitable terrain was more than 12 times more likely to occur in Alaska than the RoUS (odds ratio = 12.280; M-H=605.228, $p<.000$). Likewise, skill-based errors like the loss of directional control (odds ratio = 1.782; M-H=46.863, $p<.000$) and compensation for wind conditions (odds ratio = 2.800; M-H=122.482, $p<.000$) were more likely to occur in Alaska accidents. Additionally, VFR flight into IMC violations were over twice as likely in Alaska than in the RoUS (odds ratio = 2.395; M-H=15.815, $p<.000$).

Conclusions. These data provide valuable information for those government and civilian programs tasked with improving GA safety in Alaska and the RoUS. Furthermore, the overall results suggest that interventions being developed for Alaskan aviation would likely have similar impacts in the RoUS.

The ATC Operational Error Severity Index: An Empirical Analysis of the Effects of Loss of Horizontal and Vertical Separation. David Schroeder, Larry Bailey, and Scott Goldman. Civil Aerospace Medical Institute.

Introduction. In April 2003, the Department of Transportation Office of the Inspector General (IG) made several recommendations following their review of the methods used to index the severity of air traffic control (ATC) operational errors (OEs). The current OE severity index (SI) uses a 100-point scale to classify OEs into four categories: low, low moderate, high moderate and high severity based on five measures, including the extent to which aircraft had less than standard separation. In their review of the severity ratings for the 13-month period from May 2001 through May 2002, the IG expressed concern that some OEs appeared to be misclassified into a lower severity level. In particular, 65 incidents were identified in which an OE was not

classified as high severity despite even though aircraft appeared to be within 30 seconds of a collision. Existing regulations call for controllers to separate aircraft based on horizontal and vertical separation standards to maintain safety versus a time-based approach. In response to the report, air traffic requested that we conduct a study to examine the existing procedures used to classify OE severity.

Method**.** From May 1, 2001 through April 30, 2000, 2,390 OEs were extracted from the FAA's OE database. We determined the number and percentage of OEs that fell within each of 10 loss of separation categories (10% or less through 91% to 100% of the separation standard) for both vertical and horizontal separation. Determinations were made for each of three separation standards (Low En route, High En route, and Terminal). For example, at high en route altitudes, controllers are required to separate aircraft by either 2,000 feet vertically or five miles laterally. Finally, we determined the average SI score for OEs that were present in each of the 10 loss-of-separation categories. Separate calculations were made for loss of vertical and horizontal separation.

Results. Compared to vertical separation, the results revealed the relatively greater importance (i.e., stronger relationship) of loss of horizontal separation in determining the overall SI score. SI scores increased in a relatively linear fashion for each 10% loss in horizontal separation. These results were similar for each of the separation standards. However, the picture for the loss of vertical separation was more complex. The results suggest that once there has been a 50% loss in the prescribed vertical separation, the continued loss of vertical separation contributes little, if any, to the overall increase in the SI score.

Conclusion**.** The results clearly illustrate the important role of the loss of horizontal and vertical separation for determining OE severity. Results suggest that controller strategies used to separate aircraft are more likely to utilize direction and space versus altitude. More extensive analyses involving the other factors in the OE SI reinforce the importance of the existing approach to calibrate OE severity versus a time-based analysis.

ATCS Age and En Route Operational Errors: A Re-Investigation. Dana Broach. Civil Aerospace Medical Institute.

Public Law 92-297, passed in 1971, requires mandatory retirement of air traffic control specialists (ATCSs) at age 56. This law applies to ATCSs hired on or after May 16, 1972. This provision was based on testimony that as controllers age, the cumulate effects of stress, fatigue (from shift work), and age-related cognitive changes created a safety risk (U.S. House of Representatives, 1971; U.S. Senate, 1971). This hypothesis has been considered in two studies of en route operational errors (OEs). The study by the Center for Naval Analyses (CNA, 1995) found no relationship between controller age and OE involvement. However, Broach (1999) reported that the probability of involvement in an OE increased with age. The method used in both studies – linear regression – was less than ideal for rare events such as OEs. Moreover, experience was confounded with age in both studies. Therefore, the purpose of this study was to test the hypothesis that controller age, controlling for experience, was related to the occurrence of OEs using a method appropriate for rare events.

A total of 3,054 usable en route OE records were extracted from the FAA Operational Error/Deviation System (OEDS) for the period FY1997 through FY2003. These records were matched with air route traffic control center (ARTCC) non-supervisory controller staffing records for those years, resulting in a database of 51,898 records. Poisson regression was used to model OE count as a function of the explanatory variables age and experience. Experience was defined as tenure with the FAA as an ATCS. Tenure was recoded into six discrete categories to simplify the analysis. Age was recoded into two groups: age 55 and younger; and age 56 and older to specifically assess the risk for controllers older than the mandatory separation age. The data were aggregated by fiscal year, age group, and tenure group to create a cross-classification table suitable for analysis using the SPSS® version 11.5 General Loglinear (GENLOG) procedure.

Overall, the Poisson regression model fit the data poorly (Likelihood Ratio $\chi 2 = 283.81$, $p < .001$). To assess the effect of age across experience, the two age groups were contrasted. The Generalized Log Odds Ratio was used to estimate the odds ratio for age, that is, the odds of OE involvement for older (GE age 56) controllers (see SPSS, 1999, p. 202 – 203). The odds of OE involvement for older controllers (GE age 56) were 1.02 times greater than the odds for younger (LE age 55) controllers, with a 95% confidence interval of 0.42 to 1.64. This range of odds indicated that neither age group was less or more likely to be involved in an OE, controlling for experience.

The analysis does *not* support the hypothesis that older controllers are at greater risk of involvement in an OE. These results suggest that the original rationale for the mandatory retirement of controllers may need to be re-examined. Additional research on age and ATCS performance is recommended.

An Overview of Human Factors Associated with Rotorcraft Accidents in the United States. Albert Boquet, Cristy Detwiler, Doug Wiegmann, Scott Shappell. Civil Aerospace Medical Institute.

Introduction**.** The practicality of the helicopter has enabled it to be used in a variety of different settings in aviation. These include, but are not limited to, search and rescue, heavy hauling, logging operations, recreational flying, sightseeing, and passenger transport, to name a few. This investigation used the Human Factors Analysis and Classification System (HFACS) to analyze U.S. rotorcraft accidents from 1990 to 2000.

Method**.** Five GA pilots were recruited from the Oklahoma City area as subject matter experts and received roughly 16 hours of training on the HFACS framework. All five were certified flight instructors with a minimum of 1,000 flight hours in GA aircraft (mean = 3,530 flight hours) as of June 1999 when the study began. Using narrative and tabular data obtained from both the NTSB and the FAA NASDAC, the pilot-raters were instructed to classify each human causal factor using the HFACS framework. Of the rotorcraft accidents during this time frame, 1,148 General Aviation (GA) and 98 Commercial (C) rotorcraft met the criteria for analysis.

Results. The majority of the accidents associated with human error involved skill-based errors (SBE - 75.5%) and decision errors (DE - 27.9%). The analysis of phase of flight revealed that

the highest percentage of accidents occurred during landing (25%) followed by maneuvering (13%). Surprisingly, 12% of accidents occurred during the cruise phase of flight and another 11% occurred during takeoff. Approximately nine percent of accidents occurred during hover, seven percent during approach, with taxi, climb, and descent accounting for 7% collectively. In comparing general aviation to commercial operations, there were significantly greater SBEs noted for GA (73%) vs. Commercial (49%). The next highest category of errors was DE, however, there was a higher percentage of DE among commercial operations vs. GA (32% and 26% respectively). The same pattern held true for V with Commercial revealing roughly 18% V vs. 5% for GA.

Discussion**.** For both GA and Commercial, the most common human factor associated with accidents was SBEs. Roughly three quarters of the accidents analyzed involved these "stick and rudder" mistakes such as breakdown in visual scan or failing to trust instruments in IFC conditions. It is worth noting that as rotorcrafts have become increasingly reliable, the leading factor associated with rotorcraft accidents is SBE. A similar trend is noted in GA fixed-wing aircraft where it is estimated that as much as 80% of accidents can be traced to human error. Based upon these data, it is clear that efforts should be made to address the skill level of those individuals flying rotorcraft. This area has the greatest room for improvement, and may yield real benefits for the proper intervention strategies. The high percentage of DE also indicates that training aimed at increasing pilot abilities to assess and evaluate their environments may provide tangible gains in rotorcraft safety. Finally, a more fine-grained analysis is called for to better assess what types of errors are being made and also to look for differences among the different types of operations that rotorcraft are involved in.

ATC Operational Errors: Exceeding the Limits of Cognitive Capacities. Jing Xing, Lawrence Bailey. Civil Aerospace Medical Institute

Operational errors (OEs) in air traffic control are made by novice controllers and those with vast experience. Since air traffic control requires a high level of cognitive processing, this study was intended to elucidate the association of OEs and the capacity limits of cognitive processing in the brain. For this purpose, we developed a memory-attention model and used the model to analyze 93 runway incursion OEs related to controller-controller communications. The analysis showed that roughly 60% of the OEs might have been associated with exceeding the capacity limits of attention and memory. We identified seven types of capacity limits as potential factors contributing to the OEs: (1) inattention; (2) attentional blink; (3) memory overload; (4) disruption of memory consolidation; (5) habit interference; (6) goal interference; and, (7) similarity interference. The results suggest that controllers might be able to prevent certain types of OEs simply by being aware of those limitations. The model is preliminary and needs to be validated before being applied to a broader range of OE analysis.

Metrics of Information Complexity for Air Traffic Control Displays. Jing Xing. Civil Aerospace Medical Institute.

Information complexity associated with visual displays is a bottleneck that limits their use. While automation tools are designed to bring new functions to users and increase their capacities, they

also create new tasks associated with acquiring and integrating information from displays. In particular, a complex display increases information load to human operators and reduces usability. Thus the efficiency of the tool largely depends on the complexity of displayed information. To evaluate the costs and benefits of an automation system, it is important to understand how much information is shown on the display, and whether the information is too complex for users to process. In this paper, we present a set of observable metrics to assess information complexity of visual displays. The metrics count information complexity as the combination of three basic factors: numeric size, variety, and relation. Each factor is evaluated by the functions at three stages of brain information processing: perception, cognition, and action. Ideally, these measures provide an objective method to evaluate automation systems for acquisition and design prototypes.

General Aviation VFR-into-IMC: *Z*-Score Filtering of Demographic and Personality Variables, and the Personality Paradox. William Knecht. Civil Aerospace Medical Institute.

Does pilot personality affect risk-taking with weather? Armchair logic says "Yes," while data often say "No." In this work, we apply the technique of *z*-score filtering (slice analysis) to pilot takeoff decisions made in the face of simulated adverse weather seen at taxiway level. Such a filtering technique might prove useful, provided emphasis is kept to maintain experiment-wise reliability. Statistical and methodological problems with personality data are discussed. The results of this particular data set showed a strong effect of weather on takeoffs, as measured by visibility, cloud ceiling, and the interaction of the two. But, despite best efforts, no strong effect of personality could be found in this data set. Theoretical reasons are discussed as to why it may be difficult to show that personality predicts behavior.

The Air Traffic Selection and Training Battery: What It Is and Isn't (and How It Has Changed and Hasn't). Raymond E. King and Andrew R. Dattel. Civil Aerospace Medical Institute.

The Federal Aviation Administration (FAA) has developed a new selection procedure, the Air Traffic Selection and Training (AT-SAT) computerized test battery, to help select Air Traffic Control Specialists. AT-SAT is an aptitude test and not a test of air traffic control knowledge. Of the 264 applicants who have taken AT-SAT, 155 responded to a job announcement, while 109 previously passed the OPM (pre-employment) test and had to achieve a passing score on AT-SAT before they were admitted into training at the FAA Academy. Of the 155 job-announcement applicants, 131 (84.52%) achieved a passing score of 70 or greater (termed a "qualifying score"), while 24 applicants (15.48%) failed to achieve a minimum score of 70. Those who had been prescreened with the OPM test fared a bit better, with 104 (95.41%) achieving a qualifying score; five (4.59%) applicants failed. Current research efforts include equating a parallel form, re-hosted on a Windows 2000 operating platform, with the assistance of research participants from the US Army, US Navy, and US Air Force. Another recent project was focused on re-weighting the subtests and adjusting the overall constant to address issues of potential adverse impact, without compromising validity. A greater concern in this effort was to ensure that AT-SAT performance would predict job performance rather than just success or failure in training. Despite this re-weighting effort and updating of the operating platform, the

content of the battery remains unchanged. Future efforts will involve a longitudinal validation to compare performance on AT-SAT with success in training and on the job.

The Role of Institutional Review Boards in Aviation Research: It's the Law and It Makes Sense. E.S.Stein. William J. Hughes Technical Center.

Research in medicine and social sciences often involves the participation of human participants, who, under the rules in place today, volunteer their time and understand both the benefits and risks associated with the research. This was not always the case. Rules, regulations, and laws currently require oversight by organizations referred to as Institutional Review Boards (IRBs). These boards exist to protect the participants, ensure their ethical treatment, and encourage good research. IRBs enhance the quality of research planning, and the IRB process should be part of every researcher's timeline for completion of his/her projects.

Implementing Electronic Flight Data in the Airport Traffic Control Tower. T.R.Truitt. William J. Hughes Technical Center.

The Federal Aviation Administration (FAA) is investigating the potential effects of implementing electronic flight data systems (EFDSs) at Airport Traffic Control Towers (ATCTs). I use existing task analyses, published literature, and recent field observation data to determine the basic functionality of flight progress strips (FPSs) in the ATCT. I identify gaps in the research and formed a general set of principles to guide the design of an EFDS prototype. Given the proper application of principles for design and automation, the EFDS should maintain some of the basic functionality and benefits of the FPSs, reduce workload related to flight data entry, tracking, and sharing, and provide new features that will enhance controller performance and encourage use. I present possible risks and outcomes that are likely to accompany an EFDS in FAA ATCTs.

Surprise and Unexpectedness in Flying: Factors and Features. Janeen A. Kochan, Eyal Breiter, Florian Jentsch. University of Central Florida.

This database analysis was conducted to determine which factors, or combination of factors, play a part in creating an unwanted outcome due to surprising or unexpected events encountered by pilots. The purpose of this study was to identify likely precursors to perceived surprising and unexpected events, and, to advance our understanding of the overt behaviors and misbehaviors found in response to these events. This study also sought to determine if there were any significant differences between commercial air carrier and general aviation flight operations in regard to surprising and unexpected events. The results of this study indicated that the involvement of surprise or unexpectedness can indeed have a detrimental effect on the outcome of the flight. We also found indicators of the processes and mechanisms leading from surprise to an unwanted outcome.

Dealing with Unexpected Events in Aviation: The Role of Domain Expertise. Janeen A. Kochan, Florian Jentsch. University of Central Florida,

Unexpected events, particularly those resulting in the loss of control of an aircraft, can have devastating effects on the outcome of a flight. Pilot reactions to unexpected events vary for a

number of reasons, including previous experience with similar events, emotional connotation, and the contextual surround of the event. Events which are sufficiently discrepant from currently activated schemata and action scripts create a state of surprise. The level of surprise is based on the expected probability of the event, the strength of one's activated schemata, and the event's relationship to ongoing action scripts. Once the process of surprise has been triggered, one's ongoing cognitive and behavioral processes are altered (Meyer, Reisenzein, & Schützwohl, 1997; Meyer, Neipel, Rudolph, & Schützwohl, 1991) which in turn increases the potential for the consequence of the event to cascade toward an unwanted outcome (Wickens, 2003; Woods & Patterson, 2001).

The purpose of this study was to investigate the influence of domain expertise on the cognitive mechanisms underlying human perceptions, processing, and reactions to unexpected or surprising events. The research questions addressed in this study emerged from the intersection of psychological and physiological theories of cognition and emotion in applied settings. The overarching questions were: can we identify specific knowledge and skills which enhance one's ability to deal with unexpected events? Do these skills differ from domain expertise? Does domain expertise improve or deter one's reactions and responses to unexpected events?

A process model of surprise, adapted from Meyer, Reisenzein, and Schützwohl (1997) provided the theoretical bases for testing the research hypotheses. Specifically, differences in perception, the use of available informational cues, and the ability to exhibit cognitive flexibility in response to unexpected events were evaluated for pilots possessing more or less domain expertise. Pilot expertise level was determined by a mathematical formula developed pursuant to a factor analytic study by Doane, Sohn, and Jodlowski, (2004). Experimental design provided manipulation of cue accessibility dimensions (Kahnemann, 2003) of salience (subtle versus overt), number (few versus numerous), and priming (subtle versus overt). Cognitive flexibility dimension was measured by script shifting as evidenced by flight plan continuation or flight diversion requests.

Results demonstrated that domain expertise facilitates the acquisition speed, perception, and utilization overt cues. However, level of expertise was *not* related to subtle cue perception or use. Additionally, and most surprisingly, domain expertise was expected to delay or prevent script shifting when pilots were faced with an unexpected event, however no significant differences were found between the different levels of expertise groups. Results of the study suggest that one key to dealing with unexpected events is the ability to be cognitively flexible, and recognize when a shift in action scripts and/or modification of schemata are required.

Discussion Panel - Automation and Performance. Michelle Harper, Janeen A. Kochan, University of Central Florida; Lance Sherry, Melanie Diez, George Mason University

For several decades, it has been recognized that a pilot's role has shifted due to the introduction of automation to the cockpit. While the goal of automating the cockpit has been to provide more time for other flight critical tasks, the lack of understanding of the automation due to poor training and/or poor design in combination with interruption or unexpected events within the cockpit has hindered optimal performance. For the proposed discussion panel, past and current approaches that attempt to address the problems associated with automated cockpits will be

discussed. In addition, the collaborative research effort between the University of Central Florida and George Mason University will be addressed. In particular, this effort has focused on all the aforementioned factors that may influence a pilot's performance. The main goal of the collaboration is to integrate past and present research findings from cognitive modeling, training, and assessment in order to develop tools/methods that can be used to capture, assess, and enhance the skills needed to operate in an automated environment.

Program Update and Prospects for In-Flight Simulation Upset Recovery Training
Janeen A. Kochan, University of Central Florida; James E. Priest, Calspan Corporation

The Flight Research Training Center, established in 2002 in cooperation with the Federal Aviation Administration, focuses on improving the safety of commercial air transportation through the reduction of the loss-of-control events, which continue to be the leading cause of fatal commercial air carrier accidents (Boeing Commercial Airplanes Group, 2004). The primary research purpose of this program is the optimization of in-flight simulation-based upset recovery training. The goal of the training is to have a beneficial impact on the loss-of-control accident and incident rate. The program is designed to collect research data through an extensive training program offered to commercial airline pilots. To date, more than 235 commercial pilots have completed the integrated two-day program which includes classroom, aerobatic aircraft, and advanced in-flight simulation aircraft training on how to best respond to a variety of upset situations. This paper presents the results of the data collection and analysis effort for the FAA-Upset Recovery Training (URT) program for the twenty-four month period from August 8, 2002 through July 30, 2004.

Evidence against Crew Resource Management as a Cognitive Skill. Stacey Hendrickson, David Trumpower, and Timothy E Goldsmith. University of New Mexico.

In recent years, airlines have begun to train and assess crew resource management (CRM) tasks similarly to technical skills. CRM tasks occur as part of line-oriented training scenarios and also as performance elements within line-oriented evaluations (LOEs). In order for individual CRM categories (e.g., workload management, communication, etc.) to be thought of as skills, performance of a particular CRM category should transfer to different situations. Technical skills (e.g., proficient in use of flight management system), are assumed to transfer across flight contexts. In this study, we examined evidence that CRM behaviors generalize across different flight contexts.

We analyzed pilot assessment data from several major airlines. For each set of data, pilot (crew) performance was assessed by trained evaluators in either a full-flight simulator or an actual commercial flight. Performance was graded along a numeric scale with four or five levels of performance. Many different behaviors were evaluated throughout a grading session. Some of these behaviors were previously classified as belonging to a particular CRM category (e.g., workload management) and behaviors also occurred within a particular context (phase of flight). The basic statistical analysis compared how pilots' grades correlated within vs. between CRM categories and within vs. between phases of flight. A regression analysis showed that less than one percent of the total variance in grades was due to CRM categories; in contrast, phase of flight accounted for roughly 10% of the total variance in grades. Thus, pilots performed more

consistently within a phase of flight (regardless of the CRM skill category) than within a specific CRM category.

We discuss several caveats and limitations associated with these findings. However, the findings do suggest that CRM should be viewed differently than technical skills. One implication of these results is that pilot training may be more effectively focused around contexts rather than around CRM skills.

Automatic Text Analysis of ASAP Narrative Reports. Timothy E Goldsmith. University of New Mexico.

ASAP reports routinely contain a narrative description of a reported incident or accident. Although airlines store these narrative descriptions within their ASAP database, there appears to be little effort to systematically read and analyze these narratives, primarily due to limited personnel on the ASAP team. We have begun to investigate the use of latent semantic analysis (LSA) to analyze ASAP narrative reports. The expected results from these analyses would include: (a) helping to classifying new reports into an appropriate category; (b) uncovering an empirically-derived structure associated with ASAP incidents; and (c) suggesting to an event review committee what course of action might be best suited for an incident.

We view the use of LSA to analyze ASAP reports as aiding the ASAP team, rather than replacing functions normally performed by a team member. During the panel discussion, we will discuss some of the general issues involved in applying LSA to a collection of narratives. Additionally, we will describe the initial results from an LSA analysis of a set of narratives obtained from a major carrier.

How High Is High Enough? Quantifying the Impact of Air Traffic Control Tower Observation Height on Distance Perception. William K. Krebs, Glen Hewitt, ATO-P R&D: Steven R. Murrill, Army Research Laboratory; Ronald G. Driggers, US Army RDECOM CERDEC NVESD.

Each year the Federal Aviation Administration (FAA) builds approximately seven air traffic control towers in the National Airspace System. Each airport has unique surface and airspace characteristics, but all airports must determine the location and height of the new air traffic control tower (ATCT). These two factors impact cost and safety; therefore the FAA must develop a quantitative means in measuring what improvement in ATCT visibility can be gained by increasing tower height at different locations on the airport surface. Two metrics were developed (Object Discrimination, Line of Sight Angle of Incidence) to assess the impact of tower height on distance perception. The two metrics are robust and easy to use to assess the impact of tower height on air traffic control tower specialist distance perception.

A Systematic Approach to Addressing Human Factors Considerations in the Design of Flight Deck Components. Beth Lyall, Greg Harron, Research Integrations, Inc.

The FAA Aircraft Certification Job Aid for Flight Deck Human Factors is a decision-support tool for addressing human factors considerations during the flight deck design portion of the

aircraft certification process. The current version presents decision-support information related to the review of flight deck displays, controls, and systems. This tool provides a systematic approach for assessing human factors considerations related to the design of flight deck components.

Analysis Methods for Determining the Safety Consequences of Mixed-Fleet Flying. Beth Lyall, Research Integrations, Inc.; Mike New, Delta Air Lines

This paper describes the development and implementation of analysis methods for identifying human factors safety vulnerabilities associated with the mixed-fleet flying of the Boeing 767-400 and the Boeing 777. The results of the analysis were two sets of vulnerabilities: those with potentially critical pilot performance consequences that should be further tested, and those with minor consequences that should be considered when developing procedures and training for the mixed-fleet flying program. A longitudinal study was conducted that included data collection to address the potentially critical vulnerabilities. Examples of how vulnerabilities were addressed with video data from the study are presented and conclusions from the analysis are described.

Human Factors Design of Electronic Documents. Barbara G. Kanki, NASA-Ames Research Center; Thomas L. Seamster, Cognitive & Human Factors.

As the aviation industry continues the transition from paper to electronic documentation, the opportunity for standard and more efficient information exchange and reuse has been recognized, but the emphasis has often been on technical and engineering solutions without full consideration of user needs. This study provides an opportunity to understand operator needs in an area of documentation that encompasses ground and flight organizations within airlines. The Federal Aviation Administration (FAA), working with the Master Minimum Equipment List (MMEL) Industry Group, is investigating options for a new MMEL electronic format. The MMEL refers to a series of documents controlled by the FAA that lists equipment that may be inoperative under certain conditions that allow the aircraft to be airworthy. Each aircraft model has an MMEL, and operators must work with that master document to determine the relief items for their specific aircraft. The resulting Minimum Equipment List (MEL) for an operator's aircraft is used by both ground personnel and pilots to determine the procedures for maintaining airworthiness. Currently, the MMEL is available in text format, and the industry needs a format that is more efficient and that incorporates key aspects of the future data exchange standards. This study identifies key operator needs that should direct the development of the new MMEL format. Members of the MMEL Industry Group as well as other industry representatives were surveyed to determine the main user needs and human factors considerations for the development and evaluation of the MMEL electronic format. The results are analyzed in the context of authoring, revising and reusing aviation information in more standard and efficient ways across ground and flight operations

Special Session: Maintenance Training & Communication Technologies. **Barbara Kanki, NASA-Ames Research Center; Anand Gramopadhye, Clemson University**

When maintenance error risks are identified, they must be managed, and proactive solutions are sought that provide both corrective action solutions as well as effective process improvements. In short, cost effective interventions are needed to resolve existing human error problems. This panel focuses on tools and techniques that help identify interventions solutions and clarify various consequences of task and team changes as well as the organizational and procedural enhancements that support them. Panel participants also explore the use of new training and communication technologies that can be effectively incorporated into problem solutions. For example, virtual reality technologies offer a new way of providing accelerated training in what has been traditionally obtained through long years of on-the-job experience. The proliferation of information technologies has offered new opportunities for solving old problems as well. However, as panelist will discuss, their implementation requires thoughtful consideration and testing of the numerous technical, operational and social factors that surround their practical use.

Special Session: Maintenance Risk Management. Barbara Kanki, Alan Hobbs. NASA-Ames Research Center.

Part of the challenge of identifying, understanding and preventing maintenance error has been due to the lack of human factors data in this aviation domain. However, over the last five years, maintenance error databases have started to gain support, both at the system-wide level and within company safety organizations. Although types of maintenance errors cover a wide range of diversity, the systematic collection and analysis of data can help to reveal problem areas. Although reactive in nature, incident event data is one source of information and, in addition to the typical focus on specific corrective actions, a focus on analysis for trends and the proactive identification of risk priorities has become an expanded interest. Traditional databases that have focused on quality and human reliability have also provided opportunities to re-assess underlying causes and risk vulnerability. This panel explores several effective ways in which maintenance error risks can be managed. At the industry level, the NASA Aviation Safety Reporting System has long been respected as a system-level source of data ,and some of the discussion will focus on substantive findings, maintenance error classification techniques and data analysis strategies. At the organization-level, implications to company databases will be discussed as well as tools and techniques for improving the collection of existing reliability data.

Improving Novice Flight Performance Using a Functional Avionics Display. Carl F. Smith and Deborah A. Boehm-Davis. George Mason University

Supporters of functional interface design argue that direct interaction with the essential functional relationships of a system may aid in the acquisition of domain-specific skill. To evaluate the potential use of a functional display in assisting in the development of piloting skill, twenty novices were trained on either a conventional display or an alternative display that displays the functional relationship of power and airspeed (the Oz display). Novices trained on the functional display showed greater control of power and less deviation from a flight profile

over multiple maneuvers. Implications for future research and potential uses in training are discussed.