Aviation Safety and the Value of Neuropsychological Assessment of Pilots

Chris M. Front, Psy.D., ABAP
Office of Aerospace Medicine
Federal Aviation Administration
Washington, DC

5th Annual Aerospace Psychology Seminar
Denver, Colorado, September 2017
Brief introduction

Why are neuropsychological assessments an essential tool in aerospace medicine?

The FAA core battery and exceptions

CogScreen-Aeromedical Edition (CogScreen-AE)

Why not use flight simulators?

What is the “gold standard” for the neurocognitive assessment of pilots?

A lot to cover; please save questions until end
Federal Air Surgeon’s Neurology Summit

In 2010, the FAS convened several of the top aerospace neurologists in the USA to review and refine FAA policies.

Wide agreement that brain imaging and neurological clinical examinations are not sufficient to detect a variety of aeromedically significant neurocognitive deficits.

Therefore, the FAA relies on neuropsychological assessment to identify or rule out aeromedically significant neurocognitive deficits.
Neurocognitive Skills Essential for Skilled Aviation Performance

- Decision Making / Reasoning
- Mental Flexibility
- Planning
- Workload Management
- Concentration
- Perceptual-Motor skills
- Memory
- Situational Awareness
- Verbal Comprehension
- Perceptual Speed
- Kinesthetic Sensitivity

- Reaction Time
- Spatial Perception
- Mental Calculations
- Alertness
- Coordination
- Rate Recognition & Control
- Application of Rules
- Sustained / Divided Attention
- Self-confidence
- Suppression of emotional responses under duress

Source: Robert W. Elliott, PhD, ABPP-CN  Aerospace Health Institute
Neurocognitive Abilities Essential for Pilots

- 6 Cognitive Skills Critical for Pilot Performance (task analysis by Banich, et al., 1989)
  - Perceptual-Motor Abilities
  - Spatial Abilities
  - Working Memory
  - Attentional Performance
  - Processing Flexibility
  - Planning/Sequencing Abilities

- “Taylor Factors” (Taylor, et al., 2000) consistent findings, and added:
  - Perceptual and Processing Speed (and Working Memory)
NASA Examination of Accident Root Causes

“The Limits of Expertise: Rethinking Pilot Error and the Causes of Airline Accidents” (Dismukes, Berman & Loukopoulos, 2007)

- Carefully examined 19 major accidents in US air carriers between 1991 and 2000 in which NTSB found crew error to play central role

- Found 6 common themes:
  1. Inadvertent slips & oversights while performing highly practiced tasks under normal conditions
  2. Inadequate execution of highly practiced normal procedures under challenging conditions
3. Inadequate execution of non-normal procedures under challenging conditions
4. Inadequate response to rare situations
5. Judgment in ambiguous situations that hindsight proves wrong
6. Deviation from explicit guidance or standard operating procedures

... In other words, failures related to “Executive Functions!”
NASA Examination of Accident Root Causes

→ “Cross-cutting factors” found by Dismukes, et al.

→ “Human cognitive vulnerability, task demands, environmental events, and social, cultural, and organizational factors interacted... in these accidents. Several specific patterns... appear repeatedly.” Most prominent:
  ➢ “Concurrent task management and workload issues”
  ➢ “Situations requiring very rapid response”
  ➢ “Plan continuation bias”

→ Again... in other words, failures related to “Executive Functions!”
FAA Mission: Safety + Efficiency

“Our continuing mission is to provide the safest, most efficient aerospace system in the world.”

- For Part 121 carriers (airlines) the rate in the United States is down to fewer than 2.5 fatalities per 100 million persons flown!
- Since the last U.S. carrier fatal accident in 2008, over 3 billion passengers have been safely transported!
- ASRS Review: Similar errors still occur; but Reason’s “swiss cheese” slices haven’t lined up again (yet)
- Part 121 redundancies have been effective!
Medical Certificate Required:

- **1st Class**
- **2nd Class**
- **3rd Class (or BM)**
- **None**

### Military Ops
- USN
- USMC
- USAF
- USA
- USCG

### Commercial Ops
- Airlines (Part 121)
- Corporate
- EMS/Medical
- Agricultural
- Air Show Performance
- Sightseeing Tours
- Flight Instruction
- Aerial Survey
- Charter (Part 135)
- Cargo
- Police
- Firefighting
- News/Traffic
- Pipeline Patrol
- Fish Spotting
- Banner Tow
- Skydiving
- Glider Tow
- Etc.

### General Aviation Ops
- Recreation, Personal Transport (Part 91)
- Glider
- Sport/LSA
- Ultralight
- Balloon

### Operations Categories
- **Military Ops**
- **Commercial Ops**
- **General Aviation Ops**

---

Aviation Safety and the Value of Neuropsychological Assessment
U.S. GA Accidents: Comparison of Segments of GA

Fatal Accident Rates per 100k Flight Hours

- All GA
- Corporate
- Business
- Instructional
- Personal

NTSB
Significance of Neurocognitive Status

Aviation safety experts agree that 70-80% of general aviation ("GA") accidents are rooted in aeronautical decision making.

- Pilots keep killing themselves and their passengers in pretty much the same ways they always have. Current NTSB & FAA focus:
  - “Loss of control in flight” (resulting from, e.g., VFR into IMC; low level maneuvering, stall/spin) → Proximal Cause

Root Cause: Aeronautical decision making

- Safe flying requires sound neurocognitive abilities, particularly with regard to Executive Functions
Significance of Neurocognitive Status

Aeronautical decision making depends, among other abilities, on “Executive Functions” – a group of cognitive processes that regulate, control, and manage other cognitive processes:

- Planning, judgment, situational awareness, prioritization, allocation of attentional resources, impulse control, etc.
- Includes the awareness of and taking conscious control of:
  - Personality Traits (e.g., thrill-seeking vs risk averse)
  - Normal human cognitive biases (e.g., plan continuation bias – referred to as “get-there-itis” by pilots)
Significance of Neurocognitive Status

Causse, et al. (2011) examined age, flight experience, and executive functions as predictors of flight performance and in-flight safety-related decision-making.

- Noted that largest number of accidents occur on or around airports; often related to poor decision making by pilots who do not sufficiently take into account external cues (e.g., landing distance available, crosswind speed, etc.)

- Simulator based flight scenario (45 mins):
  - Required periodic calculations, nav by radio aids. DG failure en route required use of magnetic compass for navigation
  - En route weather report: wind at destination indicated that x-wind exceeded aircraft’s limits by 6 knots (unsafe)
Significance of Neurocognitive Status

- Results: Executive functions (especially working memory and set shifting) as measured by neurocognitive testing were the best predictor of both in-flight performance (as measured by course deviations) and the decision to attempt an unsafe landing versus diverting to a safe airport.
  - Chronological age (M=47; S.D.=15.9) was not predictive!
  - Flight experience was eclipsed by executive functions status!

- Conclusion: “...the results of this study confirm that neuropsychological evaluation is a reliable means for predicting piloting and decision-making performance.”

Significance of Neurocognitive Status

- There is a critical nexus between aviation safety and neurocognitive abilities, especially executive functions.
- Neuropsychological assessment is the only procedure to reliably identify deficits in those abilities.

“Aviation is in itself not inherently dangerous. But to an even greater degree than the sea, it is terribly unforgiving of carelessness, incapacity or neglect.”

- Captain A. G. Lamplugh, British Aviation Insurance Group
All Psychological Assessments for the FAA

- Identifying information & Reason for referral
- Thorough review of records
- Thorough psychosocial history* and clinical interview
- Report(s) of collateral informants (if applicable)
- Mental Status Exam & behavioral observations
- Administration of psychological testing IAW Spec Sheet, with description of testing, results & interpretation
- Integration of all data sources/evidence
- Diagnostic impression (based on FAA standards)
  - Including conditions previously Dxd but now in remission or resolved
- Summary and Conclusions
  - Describe subclinical findings that may impact aeromedical status
  - Refrain from opining on aeromedical eligibility
Neuropsychological Assessment Core Battery

- Provides standardization for comparability across evaluations conducted by a variety of clinicians
- Ensures adequate assessment of critical neurocognitive domains, enabling assessment of a variety of potentially disqualifying conditions, from TBI to CVA to progressive illnesses like Parkinson's
- Tests within domains were chosen for their established validity, known sensitivity & specificity, acceptance and availability, and (when possible) availability of pilot normative data
- Additional tests IAW clinical judgment of examiner
Neuropsychological Assessment Core Battery

**What is required in the core test battery?** The core test battery listed below provides a standardized basis for the FAA's review of cases, and must include:

- CogScreen-Aeromedical Edition (CogScreen-AE)
- The complete Wechsler Adult Intelligence Scales (Processing Speed and Working Memory Indexes must be scored)
- Trail Making Test, Parts A and B (Reitan Trails A & B should be used since aviation norms are available for the original Reitan Trails A & B, but not for similar tests [e.g., Color Trails; Trails from Kaplan-Delis Executive Function, etc.])
- Executive function tests to include:
  1. Category Test or Wisconsin Card Sorting Test, and
  2. Stroop Color-Word Test
- Paced Auditory Serial Addition Test (PASAT).
Neuropsychological Assessment Core Battery

- A continuous performance test (i.e., Test of Variables of Attention [TOVA], or Conners' Continuous Performance Test [CPT-II], or Integrated Visual and Auditory Continuous Performance Test [IVA+]), or Gordon Diagnostic System [GDS].
- Test of verbal memory (WMS-IV subtests, Rey Auditory Verbal Learning Test, or California Verbal Learning Test-II)
- Test of visual memory (WMS-IV subtests, Brief Visuospatial Memory Test-Revised, or Rey Complex Figure Test)
- Tests of Language including Boston Naming Test and Verbal Fluency (COWAT and a semantic fluency task)
- Psychomotor testing including Finger Tapping and Grooved Pegboard or Purdue Pegboard
- Personality testing, to include the Minnesota Multiphasic Personality Inventory (MMPI-2). (The MMPI-2-RF is not an approved substitute. All scales, subscales, content, and supplementary scales must be scored and provided. Computer scoring is required. Abbreviated administrations are not acceptable.)
The Value of CogScreen-AE

- Inclusion of CogScreen-Aeromedical Edition in the core battery provides multiple benefits
- Developed under contract for FAA, with focus on being able to detect changes in brain function that could interfere with flight operations
  - FAA contract was prompted by AMA criticism that the routine Aviation Medical Examination “has never included tests that would detect a diminution in cognitive function, which left unnoticed, may result in poor pilot judgment or slow reaction time in critical operational situations.” (JAMA, 1986)
The Value of CogScreen-AE

- Computer administered battery assessing:
  - Attention
  - Memory
  - Visual perceptual functions
  - Sequencing functions
  - Logical problem solving
  - Psychomotor speed & coordination
  - Simultaneous information processing abilities
  - Executive functions
The Value of CogScreen-AE

Normed on pilots

- 584 pilots from Part 121 Major carriers (aged 24-67)
  - 403 from large carriers
  - 181 from medium & small carriers
- 76 pilots from Part 121 Regional carriers (aged 24-62)

GA Norming Project in 2013

- Collected GA norms at the 2 major GA fly-in events: “Sun & Fun” and “Airventure/Oshkosh”
- 207 GA pilots age range 17-86

Consistent with data on aging GA population, the age distribution of GA sample was skewed; more seniors
The Value of CogScreen-AE

Result: CogScreen-AE utilizes pilot norms for all levels of pilot certification, types of operations, and age ranges:

- Private Pilot to ATP
- GA personal/recreational ops to Part 121 airline ops
- Pilots aged 17 to 86 years
The Value of CogScreen-AE

- Excellent sensitivity and specificity
  - Identification of known brain injury during validation
  - LRPV Score
  - Sensitivity = 73%
  - Specificity = 90%

*Figure 14. Estimated probability of brain dysfunction for normals and neurologic patients with confirmed brain damage.*
The Value of CogScreen-AE

- Excellent sensitivity and specificity, cont’d
- At 95% **Specificity**, CogScreen-AE Base rate analysis achieved the following CogScreen-AE **Sensitivity**:
  - Alcohol referrals 55%
  - Aviation performance referrals 82%
  - Psychiatric referrals 50%
  - Confirmed neurologic patients 80%
  - Suspected neurologic patients 68%
The Value of CogScreen-AE

→ Includes measures of multi-tasking that do not exist in any other neurocognitive tests

➢ E.g., One of the measures of multi-tasking requires monitoring of a trend indicator (not unlike a glide slope deviation indicator) and constantly inputting fine-tuned corrections necessary to maintain it within parameters while simultaneously attending to a working memory task (not unlike maintaining SA with ATC directions while flying a precision approach)
The Value of CogScreen-AE

Computer administration yields multiple benefits

- Improved standardization of administration
- Precise timing, including on multi-tasking measures that human examiner could not possibly accomplish
- Accurate recording of examinee responses
- Automated generation of correct scoring algorithms
- Automated databasing for follow-up testing (& research)
- Permits serial readministration without practice effects
  - Most conventional neurocognitive measures (e.g., Wisconsin Card Sort Test) are not designed for repeated measures testing
The Value of CogScreen-AE

- Impressive empirically demonstrated criterion / ecological validity regarding *flight performance*!
- Initial validation during test development utilized data from airline Flight Data Recorders (FDR)
  - Captains tested, then flew line ops for 3-year period
  - 64 different flight parameters recorded by FDRs
    - Deviations from course, glide slope, speed, G meter on nose gear for hard landings, etc.
- RESULTS: Flight performance was significantly correlated \((p < .01)\) with 11 CogScreen variables
The Value of CogScreen-AE

- Aeroflot TU-154
- Regression analysis revealed CogScreen predicted quality of flight performance (obtained from cockpit flight data recorders for 3 years)
  - Course, glide slope, speed
  - G meter on nose gear, etc.
- $R^2 = 0.32$ (i.e., CogScreen-AE scores explained 32% of the variance!)

The Value of CogScreen-AE

→ Aeroflot IL-86

→ Regression analysis revealed CogScreen predicted quality of flight performance (obtained from cockpit flight data recorders for 3 years)
  ➢ Course, glide slope, speed
  ➢ G meter on nose gear, etc.

→ $R^2 = 0.46$ (i.e., CogScreen-AE scores explained 46% of the variance!)

The Value of CogScreen-AE

- Taylor et al. at Stanford University
- Frasca model 141 flight simulator
- 100 pilots, aged 50-69 (m = 57)
- Flight parameters measured:
  - Staying on course
  - Correct frequency changes per “ATC”
  - Avoiding conflicting traffic
  - Accurate instrument scan
  - Executing instrument approach
  - Flight Summary Score
The Value of CogScreen-AE

Results:

- 4 CogScreen-AE variables accounted for 45% of the variance in Flight Summary Scores
  - Speed/Working Memory
  - Visual Associative Memory
  - Motor Coordination
  - Tracking

- Speed/Working Memory had the highest correlation with Flight Summary Score ($r = .57$)
The Value of CogScreen-AE

🔗 Data collected by Part 121 airlines over several decades demonstrate CogScreen-AE predictive ability

🔗 Airline data are generally proprietary, unavailable to scientific community

➤ Exception: Lehenbauer, Bettes & Wolbrink (presentation at AsMA, 2007) findings consistent with other airline findings:

➤ CogScreen *predicts training performance* – program completion versus problems (e.g., flight simulator session failures, remedial simulator sessions)

➤ CogScreen *predicts actual line operations performance* – success versus problems (e.g., failed check rides, pilots who are later released or resign)
CONCLUSION: CogScreen-AE assesses the domains of cognitive functioning that are critical to flight performance, and does so effectively enough to be predictive of pilot performance, both in flight simulators and in real-world operations.
The Value of CogScreen-AE

Why administer a “screening test” (CogScreen) if you are administering a complete neurocognitive battery?

- Relationship between conventional measures in the test battery and aviation are unknown. CogScreen is a valid predictor of both flight training and flight performance.
- Conventional measures are weak with respect to assessment of multi-tasking ability. CogScreen provides measures of multi-tasking ability that are otherwise unavailable.
- CogScreen is designed for repeated measures testing. Most conventional measures are not designed for repeated measures testing and are vulnerable to “practice effects.”
- Extensive normative data on pilots are available for CogScreen, whereas only limited normative data are available for pilots on a limited number of conventional measures.
The Value of CogScreen-AE

- Although developed for the FAA, CogScreen-AE is acknowledged internationally as an indispensable test battery in aviation
  - U.S. Navy: primary neuropsych assessment component
  - Many major airlines: selection tool with extraordinary predictive utility
  - Utilized by numerous other countries (in air forces and/or in civil aviation), including:
    - United Kingdom; Germany; Canada; Sweden; Norway; Israel; Mexico; Japan; Russia; Australia; South Africa; India; Chile; Columbia; Singapore; Kazakhstan; Hungary; U.A.E.
FAA Specification Sheets

Guide for Aviation Medical Examiners (AMEs), detailing neuropsychological evaluation requirements for:

- Potential Neurocognitive Impairment (TBI, CVA, MS, etc.)
- ADHD/ADD
- HIV+ Status
- Substances of Dependence/Abuse
- Depression/SSRI Medication
Example of Modified Spec Sheet: SSRI Policy

April 2010: FAA announced “SSRI Policy” allowing FAA airman SI medical certification for certain SSRIs

Recognition that:

- Conditions (e.g., depression, anxiety) Txd with SSRIs may result in impairment in cognitive functioning.
- Selective Serotonin-Reuptake Inhibitor (SSRI) medications can potentially impair cognitive functioning

Hence, one of the requirements for SI was a neuropsychological evaluation to rule out deficits
Example of Modified Spec Sheet: SSRI Policy

Initially, the Core Battery described above was required for the first 24 months of the program.

In order to reduce burden on airmen, the FAS directed a study of the initial cases, to investigate the feasibility of utilizing CogScreen-AE as a screen.

Based on the final determinations made after reviews of the entire NP protocols:

• If only CogScreen-AE had been utilized, would there have been any false negatives?
• Identify the sensitivity, specificity, positive predictive value and negative predictive value of CogScreen-AE.
Example of Modified Spec Sheet: SSRI Policy

Examined 98 cases submitted/reviewed APR 2010 – JUN 2012:

<table>
<thead>
<tr>
<th>CogScreen-AE: Weak/Deficient Performance</th>
<th>Cognitive Deficiency on Full NP Test Battery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
<tr>
<td>No</td>
<td>76</td>
</tr>
</tbody>
</table>

| Total NP Determinations                  | 11 (Disqualified)                           | 87 (Qualified) | 98 |

Examined 98 cases submitted/reviewed APR 2010 – JUN 2012:
Example of Modified Spec Sheet: SSRI Policy

CogScreen-AE performance:

Sensitivity .................. 100.0%
Specificity .................... 87.4%
PPV .......................... 50.0%
NPV ........................... 100.0%
for LRPV score >0.80 . . . O.R. = 3.95
Example of Modified Spec Sheet: SSRI Policy

→ CogScreen-AE performed well as a screening test by identifying 100% of airmen with significant cognitive deficiencies evidenced by testing with the full neuropsychological test battery.

→ The positive predictive value (PPV) of 50% indicates that aviators performing poorly on CogScreen have a 50/50 chance of passing on the comprehensive conventional neuropsychological test battery.
Example of Modified Spec Sheet: SSRI Policy

- Based on the findings of this study, 77.6% of airmen tested by CogScreen-AE may be spared full core battery
- These results were the basis for the FAA revising the SSRI Program requirement, as follows:
  - If the airman can produce WNL performance on CogScreen-AE, then they no longer must complete the core battery
  - A significant savings in time and expense for airman, as well as more efficient review processing by FAA, with very low risk of a false negative finding
Why Not Use Flight Simulators?

- There is a difference between evaluating flight skills and evaluating medical conditions
  - 14 CFR Part 61 ≠ 14 CFR Part 67
  - We do not use flight simulators to evaluate the flight risks associated with other medical conditions (epilepsy, cardiac conditions, diabetes, etc.)
- Even if one were to accept the notion that flight simulators could be a valid measure of aeromedical suitability, there are a number of problems with operationalizing that notion
Reasons Not To Use Flight Simulators

Among expert sources considered:


- Provides a list of reasons that simulator checks are not the most reliable method of detecting cognitive impairment

Carlos Porges, Psy.D., ABPP. Clinical neuropsychologist and ATP, First Officer in Boeing 767 for a major airline.

- Has been completing Part 121 simulator checks as a pilot for over 15 years
Reasons Not To Use Flight Simulators

Any simulator tasks used for evaluation of neurocognitive status would require rigorous standardization and the development of appropriate normative data.

- Simulator & aircraft types
- Airman age range
- Lack of Availability
- Airman cert. types
- Need Ns for each!
- High Cost
Reasons Not To Use Flight Simulators

→ Adequate standardization is currently impractical

➢ Dr./FO Porges observed: “Different simulators of the same aircraft -- from different manufacturers -- fly differently. They are sufficiently close to actual flight characteristics of the aircraft that training can be conducted in them, but they certainly do not fly the same (a) to the actual aircraft and (b) to each other.” Also, “Simulators made by the same manufacturer, simulating the same aircraft, fly differently. Everyone knows this at the training center [as evidenced by comments such as]: "That simulator’s visuals are off,” or "That one taxis like it was on soap,” etc.

➢ The above pertains to the full-motion simulators used in Part 121 ops – the most sophisticated available!
Reasons Not To Use Flight Simulators

- Airline simulator sessions follow a set format; pilots can predict many of the tasks they will be asked to perform; Tasks are highly developed, testing “overlearned” skills (i.e., practiced to the point of mastery)

- Overlearned skills tend to be maintained under stress and may be retained during the early stages of cognitive decline – unlike cognitive abilities such as novel reasoning, problem-solving and decision-making

- A pilot with deficits in executive functions (e.g., problem-solving, strategy selection, flexibility of thinking) would require novel simulator flight scenarios in order to test those abilities
Reasons Not To Use Flight Simulators

NASA (Loukopoulos, Dismukes & Barshi, 2009), studied multi-tasking & prospective memory failures in actual line operations

“Although modern flight simulators are so similar to actual aircraft cockpits in layout of instruments and controls and use such realistic visual displays and motion platforms that they are used for both training and certification of pilots, simulation scenarios typically do not capture the full range of operational demands occurring in line operations. This is especially true of the unscripted aspects of operations that are the focus of our study.” [i.e., “real world” perturbations]
Reasons Not To Use Flight Simulators

- Novel simulator flight scenarios would need to be developed, standardized, and data gathered on normative performance parameters – an enormous, costly undertaking.

- Novel scenarios would not remain novel!
  - Dr./FO Porges points out that pilots routinely share “gouge” such that, as soon as a novel task was introduced with one airman, it would become common knowledge and others would prepare for it, eliminating its value as a novel task.

- Instructor/examiner decides whether a pilot has passed or failed the sim test, but examiners’ ratings may vary due to feelings toward the examinee and differing interpretations of the rating system – *No established inter-rater reliability!*
Reasons Not To Use Flight Simulators

» Finally, simulator examiners/instructors are not trained to detect cognitive impairment or mental health issues in pilots – a set-up for erroneous assessment (*Part 61 ≠ Part 67*)

» This same problem arises when relying on CFIs to opine on a pilot’s neurocognitive status as evidenced by their performance in training (*Part 61 ≠ Part 67*)
 Reasons Not To Use Flight Simulators

→ **CONCLUSION**: In evaluating aeromedically significant neurocognitive deficits, *simulator testing cannot provide the quality of aeromedical data for risk-based decision making that is available through the use of standardized neurocognitive measures*, particularly those normed on pilots and validated with pilot flight performance.
So What Is the FAA’s “Gold Standard?”

Reliance on standardized procedures with appropriate normative data can reveal subtle, yet aeromedically significant neurocognitive issues that are not detectable by brain imaging, neurological clinical examination, or clinical interview alone and, for the reasons noted, would not reliably be detectable in flight simulator performance.

Therefore: *The gold standard is neuropsychological assessment by an aeromedically trained clinical neuropsychologist utilizing CogScreen-AE* alone or in combination with the FAA core neuropsychological battery (and/or other tests as indicated by the particular condition).
Questions?

Boeing 767-400 Full Motion Flight Simulator