Comparing Methods to Evaluate Cognitive Deficits in Commercial Airline Pilots

Joshua Potocko, MD/MPH
UCSF Occupational and Environmental Medicine Residency
30 Min
Background
Philosophy
Scope
Comparisons
Evidence Review: Example
Questions for the Group
LCDR Joshua R. Potocko, MC (FS/FMF), USN

“The views expressed in this presentation reflect the results of research conducted by the author and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the United States Government.”
Potocko’s 6 Proclamations:

Certain baseline cognitive functions decline with age.

Within an individual, these declines are difficult to predict.

Between individuals, different types of decline are variable.

Understanding the following are critically important to aviation safety:

Age-related declines

Temporary disturbances in cognitive function,

Stable (or progressive) baseline disturbances due to injury, illness, disease, medication, and substance use.
<table>
<thead>
<tr>
<th>Cognitive Function</th>
<th>Typical, Normal, Adequate, Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Inefficiency</td>
<td>Circadian, Fatigue, Mood, Stress</td>
</tr>
<tr>
<td>Cognitive Deficiency</td>
<td>Injury, Illness, Meds, Substances</td>
</tr>
</tbody>
</table>
| Cognitive Disability            | Above plus regulatory decisions  
                                 | = requires safety factor          |
Age XX?
Disease?
Baseline
Progressive Decline
Permanently Unfit
PHILOSOPHY
Research Question: What is the “best” way to evaluate cognitive deficits in airline pilots?

Best: historical? expert opinion?...or evidence-based?

Evidence: “that which eliminates alternative explanations”
“Methodologies that eliminate the most bias are considered to be highest quality”
## Levels of Evidence: Oxford

<table>
<thead>
<tr>
<th>Diagnosis:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a: Systematic review (with homogeneity) of Level 1 diagnostic studies; or a clinical decision rule with 1b studies from different clinical centers.</td>
</tr>
<tr>
<td>1b: Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center</td>
</tr>
<tr>
<td>1c: Absolute SpPins And SnNouts (An Absolute SpPin is a diagnostic finding whose Specificity is so high that a Positive result rules-in the diagnosis. An Absolute SnNout is a diagnostic finding whose Sensitivity is so high that a Negative result rules-out the diagnosis).</td>
</tr>
<tr>
<td>2a: Systematic review (with homogeneity) of Level &gt;2 diagnostic studies</td>
</tr>
<tr>
<td>2b: Exploratory cohort study with good reference standards; clinical decision rule after derivation, or validated only on split-sample or databases</td>
</tr>
<tr>
<td>3a: Systematic review (with homogeneity) of 3b and better studies</td>
</tr>
<tr>
<td>3b: Non-consecutive study; or without consistently applied reference standards</td>
</tr>
<tr>
<td>4: Case-control study, poor or non-independent reference standard</td>
</tr>
<tr>
<td>5: Expert opinion without explicit critical appraisal, or based on physiology, bench research or &quot;first principles&quot;</td>
</tr>
</tbody>
</table>
Systematic Review: “A Study of Studies”

Clear Study Question (Definitions)
  Population, Intervention, Control, Outcome
Inclusion, Exclusion
Lit Search (include gray)
Selection (using criteria)
Data Extraction
Systematic Review: “A Study of Studies”

Quality or Bias Assessment
Heterogeneity
Meta-Analysis (if able)
Evaluate, Interpret Results (clinical relevance)
Publish (under peer review)
Define “cognitive”

Intelligence: Crystallized and Fluid
Visuospatial Abilities/Construction
Processing Speed
Attention
Memory
Language

Executive Functioning
Define “executive functioning”

“ability to self-monitor, plan, organize, reason, be mentally flexible, and problem-solve”

Define “cognitive” => aviation


*Intelligence: Crystallized and Fluid*

- Visuospatial Abilities/Construction
- Processing Speed
- Attention
- Memory
- Language
- Executive Functioning
- Spatial Abilities
- Perceptual Motor
- Attention
- Working Memory
- Processing Flexibility
- Planning or Sequencing
Individual Differences and Age-Related Performance Assessment in Naval Aviators
Part 1: Battery Development and Assessment

Rolf Braune and Christopher D. Wickens

Prepared for
Human Performance Sciences Department
Naval Aerospace Medical Research Laboratory
N.A.S., Pensacola, F.L.
Contract No. N00014-82-C-0115
Work Unit No. MR-14-0322-0009

Approved for Public Release: Distribution Unlimited
SCOPE
Scope: 30 min, then questions

The larger “at-risk” population
Focus on Class 1

Population of interest:

Current airline transport pilots
FAA Numbers

U.S. Airports
2.6 million passengers / DAY

~ 165,000 airline transport certificates
~ 90,000 operational?
~ 300,000 AME exams per yr
~ 1.5% denial

ICAO
3.5+ BILLION passengers / year

---

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Total</th>
<th>Airline Transport</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>609,305</td>
<td>165,228</td>
</tr>
<tr>
<td>14-15</td>
<td>317</td>
<td>0</td>
</tr>
<tr>
<td>16-19</td>
<td>17,350</td>
<td>0</td>
</tr>
<tr>
<td>20-24</td>
<td>61,034</td>
<td>902</td>
</tr>
<tr>
<td>25-29</td>
<td>67,901</td>
<td>5,491</td>
</tr>
<tr>
<td>30-34</td>
<td>57,885</td>
<td>11,683</td>
</tr>
<tr>
<td>35-39</td>
<td>53,294</td>
<td>16,684</td>
</tr>
<tr>
<td>40-44</td>
<td>46,771</td>
<td>18,181</td>
</tr>
<tr>
<td>45-49</td>
<td>49,362</td>
<td>21,943</td>
</tr>
<tr>
<td>50-54</td>
<td>55,746</td>
<td>25,261</td>
</tr>
<tr>
<td>55-59</td>
<td>59,930</td>
<td>25,236</td>
</tr>
<tr>
<td>60-64</td>
<td>54,309</td>
<td>19,176</td>
</tr>
<tr>
<td>65-69</td>
<td>37,879</td>
<td>10,018</td>
</tr>
<tr>
<td>70-74</td>
<td>26,444</td>
<td>6,131</td>
</tr>
<tr>
<td>75-79</td>
<td>12,967</td>
<td>2,823</td>
</tr>
<tr>
<td>80 and over</td>
<td>8,116</td>
<td>1,699</td>
</tr>
</tbody>
</table>
# Flight Clearance: Operational Impacts

<table>
<thead>
<tr>
<th>Restrictive</th>
<th>Lenient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Too few pilots</td>
<td>Loss of public confidence</td>
</tr>
<tr>
<td>Loss of experience</td>
<td>Higher near-miss risk?</td>
</tr>
<tr>
<td>Recruitment &amp; Training</td>
<td>Higher mishap risk?</td>
</tr>
<tr>
<td>*Medical secrecy (increased risk?)</td>
<td>*Decreased stigma? (better MH care?)</td>
</tr>
</tbody>
</table>

Either way: bad press, political pressure, lawsuits...
AME Exams
  +/- Neuropsych Testing

Flight Syllabus
  Written, Simulator, Flight Tests
Two core aims of neuropsychological assessment are often to determine whether a patient (a) has cognitively declined from (or returned to) their premorbid status and (b) has cognitive difficulties that are significant enough to interfere with (or sufficient to support) real-world functional task performance.

These will herein be referred to as testing for impairment and deficiency, respectively.

*The main premise of this study is that detecting impairment and deficiency are distinct endeavors that require different interpretive methods.*
Scope: common neuropsych (NP) tests

Ammons Quick Test
Beck Depression Inventory, Anxiety Inventory, and Hopelessness Scale
Bender Visual Motor Gestalt (BVMG) Test
Boston Diagnostic Aphasia Examination
Boston Naming Test
California Verbal Learning Test
CANTAB (Cambridge Neuropsychological Test Automated Battery)
CDR Computerized Assessment System
Clinical Dementia Rating
CNS Vital Signs
Cognitive Assessment Screening Instrument (CASI)
Cognitive Function Scanner (CFS)
Cognitive Symptom Checklists
Comprehensive Aphasia Test (CAT)
Cognistat (The Neurobehavioral Cognitive Status Examination)
CogScreen: Aeromedical Edition
Controlled Oral Word Association Test (COWAT or FAS)
Continuous Performance Task (CPT)
d2 Test of Attention
Dean-Woodcock Neuropsychology Assessment System (DWNAS)
Delis-Kaplan Executive Function System (D-KEFS)
Dementia Rating Scale
Digit Vigilance Test
Figural Fluency Test
Finger Tapping (Oscillation) Test
General Practitioner Assessment Of Cognition (GPCOG)
Grooved Pegboard
Halstead Category Test
Halstead-Reitan Neuropsychological Battery
Hayling and Brixton tests
Hooper Visual Organization Test
Iowa gambling task
Kaplan Baycrest Neurocognitive Assessment
Kaufman Functional Academic Skills Test
Kaufman Short Neuropsychological Assessment
Lexical decision task
Luria-Nebraska Neuropsychological battery
Minnesota Multiphasic Personality Inventory
MCI Screen
Memory Assessment Scales
MicroCog
Millon Clinical Multiaxial Inventory (MCMI)
Mini mental state examination (MMSE)
Mooney Problem Checklist
Multilingual Aphasia Examination
NEPSY
North American Reading Test
Paced Auditory Serial Addition Test (PASAT)
Pediatric Attention Disorders Diagnostic Screener (PADDs)
Paulhus Deception Scales
Personality Adjective Checklist
Repeatable Battery for the Assessment of Neuropsychological Status
Quick Neurological Screening Test
Rey Auditory Verbal Learning Test
Rey-Osterrieth Complex Figure
Rivermead Behavioural memory Test
Rogers Criminal Responsibility Scale
Rorschach test
Ruff Figural Fluency Test
Sensory Screening Test
SCL-90 (Symptom Checklist 90)
Shipley Institute of Living Scale
Stroop Task
Symbol Digit Modalities Test
Tactual Performance Test
Test of Memory Malingering
Test of Memory and Learning (TOMAL)
Test of Variables of Attention (T.O.V.A.)
The Tower of London Test
Trail-Making Test (TMT) or Trails A & B
Validity Indicator Profile
Verbal fluency tests
Wechsler Adult Intelligence Scale (WAIS)
Wechsler Intelligence Scale for Children (WISC-IV IQ test)
Wechsler Memory Scale (WMS)
Wechsler Test of Adult Reading
Wide Range Achievement Test (WRAT-4)
Wisconsin card sorting task (WCST)
Wonderlic Personnel Test
Word Memory Test
Ammons Quick Test
Beck Depression Inventory, Anxiety Inventory, and Hopelessness Scale
Bender Visual Motor Gestalt (BVMG) Test
Boston Diagnostic Aphasia Examination
Boston Naming Test
California Verbal Learning Test
CANTAB (Cambridge Neuropsychological Test Automated Battery)
CDR Computerized Assessment System
Clinical Dementia Rating
CNS Vital Signs
Cognitive Assessment Screening Instrument (CASI)
Cognitive Function Scanner (CFS)
Cognitive Symptom Checklists
Comprehensive Aphasia Test (CAT)
Cognistat (The Neurobehavioral Cognitive Status Examination)
CogScreen: Aeromedical Edition
Controlled Oral Word Association Test (COWAT or FAS)
Continuous Performance Task (CPT)
d2 Test of Attention
Dean-Woodcock Neuropsychology Assessment System (DWNAS)
Delis-Kaplan Executive Function System (D-KEFS)
Dementia Rating Scale
Digit Vigilance Test
Figural Fluency Test
Finger Tapping (Oscillation) Test
General Practitioner Assessment Of Cognition (GPCOG)
Grooved Pegboard
Halstead Category Test
Halstead-Reitan Neuropsychological Battery
Hayling and Brixton tests
Hooper Visual Organization Test
Iowa gambling task
Kaplan Baycrest Neurocognitive Assessment
Kaufman Functional Academic Skills Test
Kaufman Short Neuropsychological Assessment
Lexical decision task
Luria-Nebraska Neuropsychological battery
Minnesota Multiphasic Personality Inventory
MC1 Screen
Memory Assessment Scales
MiniCog
Millon Clinical Multiaxial Inventory (MCMI)
Mini mental state examination (MMSE)
Mooney Problem Checklist
Multilingual Aphasia Examination
NEPSY
North American Reading Test
Paced Auditory Serial Addition Test (PASAT)
Pediatric Attention Disorders Diagnostic Screener (PADDs)
Paulhus Deception Scales
Personality Adjective Checklist
Repeatable Battery for the Assessment of Neuropsychological Status
Quick Neurological Screening Test
Rey Auditory Verbal Learning Test
Rey-Osterrieth Complex Figure
Rivermead Behavioural memory Test
Rogers Criminal Responsibility Scale
Rorschach test
Ruff Figural Fluency Test
Sensory Screening Test
SCL-90 (Symptom Checklist 90)
Shipley Institute of Living Scale
Stroop Task
Symbol Digit Modalities Test
Tactual Performance Test
Test of Memory Malingering
Test of Memory and Learning (TOMAL)
Test of Variables of Attention (T.O.V.A.)
Tower of London Test
Trail-Making Test (TMT) or Trails A & B
validity Indicator Profile
Verbal fluency tests
Wechsler Adult Intelligence Scale (WAIS)
Wechsler Intelligence Scale for Children (WISC-IV IQ test)
Wechsler Memory Scale (WMS)
Wechsler Test of Adult Reading
Wide Range Achievement Test (WRAT-4)
Wisconsin card sorting task (WCST)
Wonderlic Personnel Test
Word Memory Test
COMPARISONS
Comparing Safety-Sensitive Positions

1. Truck Drivers
2. Railroads
3. Maritime
4. DoD, DHS
5. Nuclear Power
6. Physicians
Identification of Evidence Bases Used in Evidence Report
In developing the evidence report titled, Stroke and Commercial Motor Vehicle Driver Safety, a comprehensive systematic literature search was undertaken accessing several electronic databases: MEDLINE, PubMed (PreMEDLINE), EMBASE, PsycINFO, CINAHL, TRIS, the Cochrane library (through January 10, 2008). Abstracts of identified studies were examined to determine which articles would be retrieved, before they could be included in each evidence base. Hand searches of the “gray literature” were also performed.

Medical Expert Panel Members
Abiodun Akinwuntan, PhD
Philip Gorelick, MD
Meheroz Rabadi, MD
Development and Findings of Evidence Report

The three key questions asked in the evidence report were as follows:

Key Question 1: Among individuals who have experienced a TIA (transient ischemic event), what is the risk of experiencing a future stroke?

Key Question 2: Are individuals who have experienced a stroke at an increased risk for a motor vehicle crash (crash risk or driving performance)?

Key Question 3: If so, can neuropsychological testing of individuals who have experienced a stroke predict crash risk?
Key Question 3: If so, can neuropsychological testing of individuals who have experienced a stroke predict crash risk?

Summary: Certain neuropsychological tests can predict the outcome of driving performance measured by a road test or in-clinic driving evaluation (Strength of Conclusion: Moderate).

*Whether neuropsychological tests can predict actual crash risk cannot be determined as no such currently available evidence exists.*
Definition: clinical outcome of interest? Ability to handle complex emergency on any given flight?
**Sample Article: 2011 Mentioned in AsMA 2017, FAA Neuropsych talk**

**Cognitive aging and flight performances in general aviation pilots**

Mickaël Causse, Frédéric Dehais, Mahé Arexis & Josette Pastor

<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Title</th>
<th>Journal</th>
<th>Public?</th>
<th>Location</th>
<th>Pop</th>
</tr>
</thead>
</table>

(up to date) (financial, academic bias) (confirmation bias) (journal impact factor) (peer-review, publication/reporting bias, article impact factor) (publication bias) (cross-cultural bias) (relevance)
<table>
<thead>
<tr>
<th>N</th>
<th>Evidence</th>
<th>Study Type</th>
<th>Hypothesis</th>
<th>Recruitment</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>3b</td>
<td>Exploratory Cohort, compared to poorly defined ref standard</td>
<td>chronological <strong>age is not a sufficient criterion</strong> to predict piloting performance and decision-making relevance and that <strong>cognitive performance is a much more relevant criterion</strong></td>
<td>No information</td>
</tr>
<tr>
<td>(power)</td>
<td>(evidence hierarchy)</td>
<td>(design bias)</td>
<td>(clinically significant question?) <em>(a priori hypothesis testing? or exploratory correlations?)</em></td>
<td>(selection bias)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incl</th>
<th>Excl</th>
<th># of Evaluators, Blinded?</th>
<th>Participants Blinded?</th>
<th>Predictor Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>M, RHD, French, College+</td>
<td>Logicians, <strong>airlines</strong>, sens/neuro/psych deficits Emotional deficits (BIS/STAI) CNS-affecting substance use</td>
<td>1</td>
<td>32</td>
<td>Chrono Age Chrono Flight Time NP Test Battery (FAA Core?): Target Hitting Test (No) 2-Back Test (No) WCST (Yes) Spatial Stroop Test (Yes)</td>
</tr>
<tr>
<td>(selection bias - convenience)</td>
<td>(selection bias - omission)</td>
<td>(observer bias)</td>
<td>(placebo effect)</td>
<td>(instrument bias)</td>
</tr>
<tr>
<td>Outcome Variables</td>
<td>Statistics</td>
<td>Confounding Age-Flight Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>--------------</td>
<td>-----------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim Flight Perf:</td>
<td></td>
<td>&quot;No significant correlation&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Flight Path Deviations</strong></td>
<td>Regression</td>
<td>(p=.117, r=0.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(angular deviation in the horizontal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>axis from the ideal flight path)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Crosswind &quot;no-land&quot; decision</strong></td>
<td>1-way ANOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(incorrect if inappropriate with 6-knot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CW tolerance)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(response bias, procedural bias)</td>
<td>(statistical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>assumptions)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(statistical assumptions)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are these the <strong>outcomes of interest</strong></td>
<td>Assumes linear</td>
<td>Assumes linear relationship</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in aviation safety, as demonstrated</td>
<td>relationship</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>by crashes, near-misses, incident</td>
<td>Why?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reports, safety studies?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Does incorrect = unsafe?</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Eye Test: does this look linear?
Case Example: AsMA/CAMA 2017

Significance of Neurocognitive Status

- Results: Executive functions (especially working memory and set shifting) as measured by neurocognitive testing were the best predictor of in-flight performance (as measured by course deviations) and the decision to attempt an unsafe landing versus divert based on CW.
  - 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.”

Case Example:

AsMA/CAMA 2017 Misconstrues findings, and fails to mention:

“In contradiction with our expectations, analysis showed that age was correlated with piloting performance.”

Significance of Neurocognitive Status

➢ Results: Executive functions (especially working memory and set shifting) as measured by neurocognitive testing were the best predictor of in-flight performance (as measured by course deviations) and the decision to attempt an unsafe landing versus divert based on CW:
  • 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 certain neuropsychological test performance may correlate with certain piloting and decision-making abilities in a small group of general aviation pilots

Opportunities for evidence-based assessment (FAA)

1. Medical Re-Certification Process
   AME Designee <-> Specialists
   => FAA Reviewers, determination

2. Simulator Syllabus (“functional evaluation”)
   Level D, Full Motion, Airline operated

3. Check rides (simulator or flight)
   Designated Pilot Examiners, supervised
AME Exams
+/- Neuropsych Testing

Flight Syllabus
Written, Simulator, Flight Tests

Because NP testing precedes simulator or flight test as a type of “gatekeeper,” this is the logical place to conduct a medical evidence based systematic review.
Thank you!

Ansa Jordaan (ICAO), Immanuel Barshi (NASA Ames)
Michael Berry (FAA/FAS), Randy Georgemiller (FAA/NP)
John Hastings (Neuro/AME)
Gary Kay (CogScreen/NP), Nicolle Ionascu (HIMS/NP)
Steven Porter (Navy NP), Ed Park (Navy Neuro)
Comparing Methods to Evaluate Cognitive Deficits in Commercial Airline Pilots

Joshua Potocko, MD/MPH
UCSF Occupational and Environmental Medicine Residency