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

BACKGROUND

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:

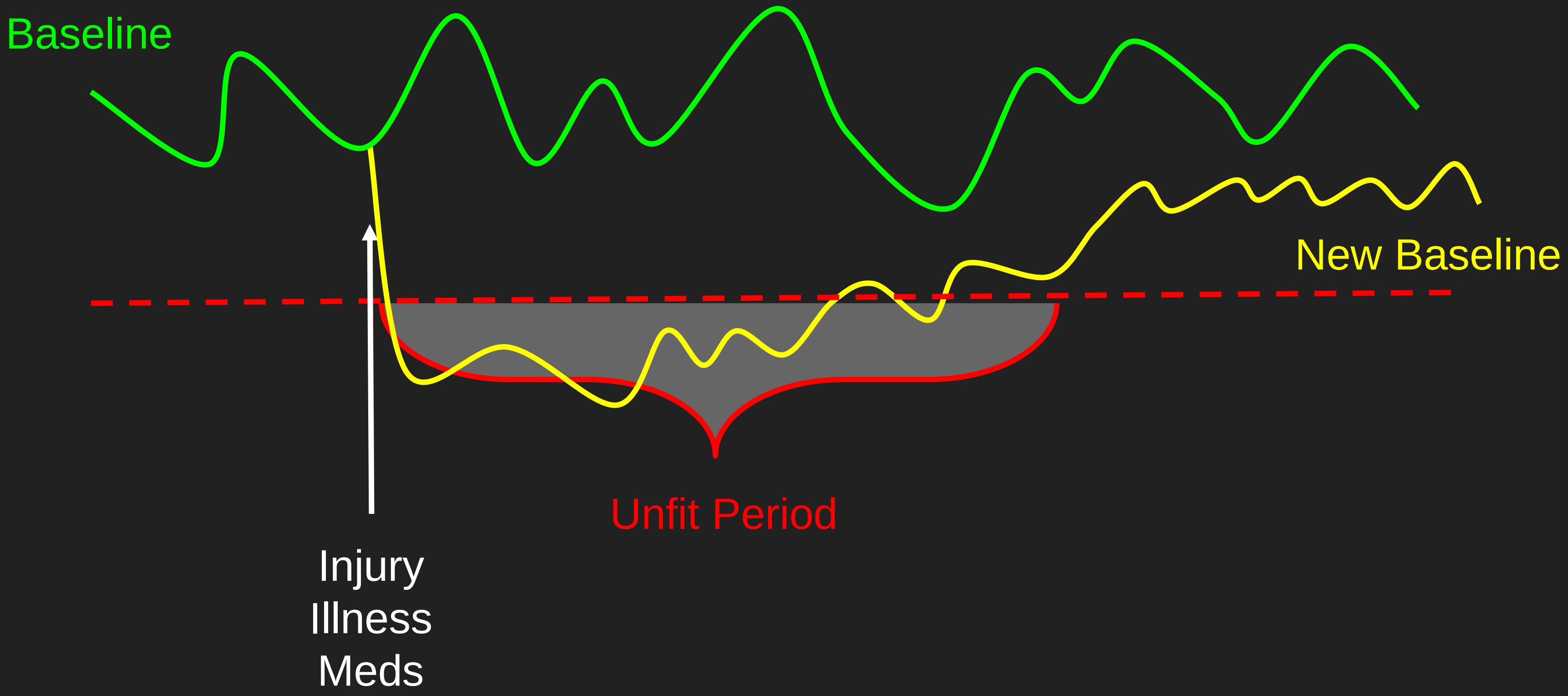
Age-related declines

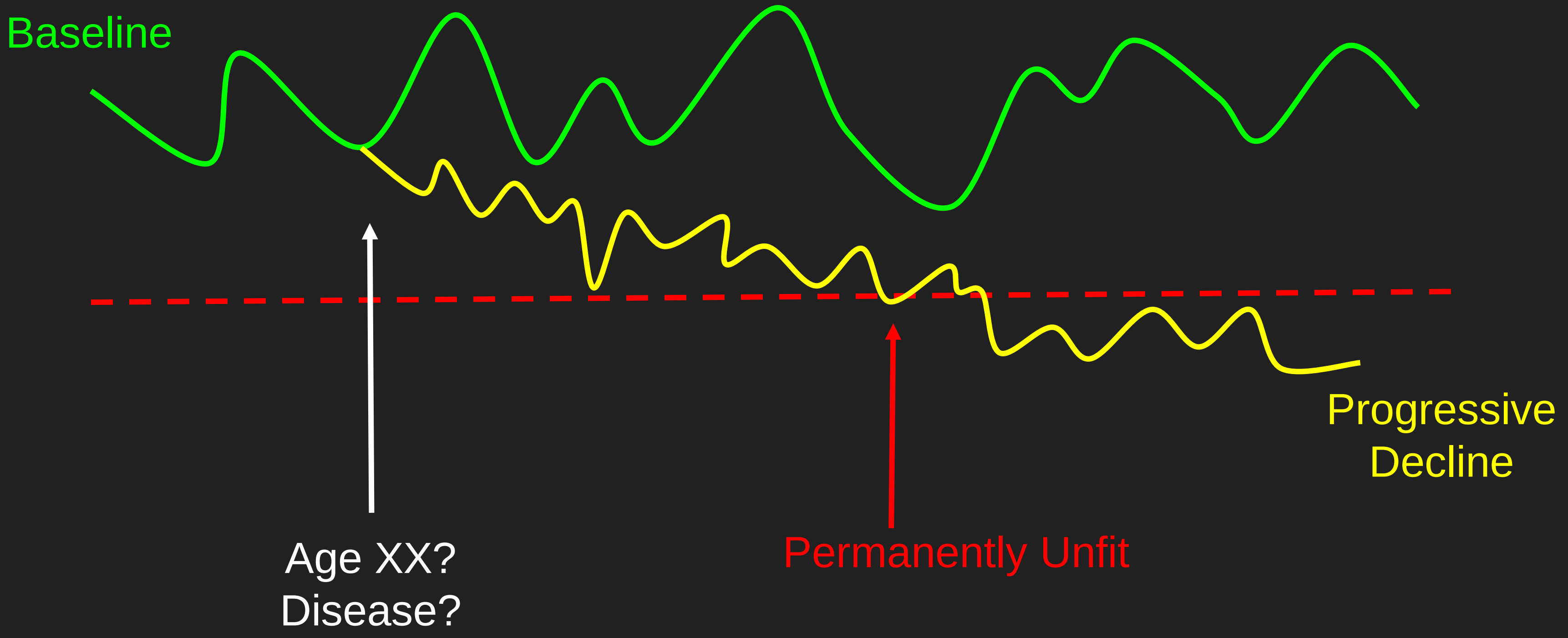
Temporary disturbances in cognitive function,

Stable (or progressive) baseline disturbances due to injury, illness, disease, medication, and substance use.

When does cognitive dysfunction become unsafe?

Cognitive Function	Typical, Normal, Adequate, Average
Cognitive Inefficiency	Circadian, Fatigue, Mood, Stress
Cognitive Deficiency	Injury, Illness, Meds, Substances
Cognitive Disability	Above plus regulatory decisions => *requires safety factor*


Baseline
New Baseline
Unfit Period
Injury
Illness
Meds


Baseline
Age XX?
Disease?
Permanently Unfit
Progressive
Decline

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"

Article

November 4, 1992

Evidence-Based Medicine

A New Approach to Teaching the Practice of Medicine

ORIGINAL ARTICLE

International Journal of Public Health Policy and Health Services Research

Evaluating the quality of medical evidence in real-world contexts

Andrew Jones MA Student[1] | Daniel Steel PhD, Associate Professor[2]

Levels of Evidence: Oxford

Diagnosis:

1a:	Systematic review (with homogeneity) of Level 1 diagnostic studies; or a clinical decision rule with 1b studies from different clinical centers.
1b:	Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center
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).
2a:	Systematic review (with homogeneity) of Level >2 diagnostic studies
2b:	Exploratory cohort study with good reference standards; clinical decision rule after derivation, or validated only on split-sample or databases
3a:	Systematic review (with homogeneity) of 3b and better studies
3b:	Non-consecutive study; or without consistently applied reference standards
4:	Case-control study, poor or non-independent reference standard
5:	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

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”

Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. Clin Geriatr Med. 2013

Intelligence: Crystallized and Fluid

Visuospatial Abilities/Construction

Processing Speed

Attention

Memory

Language

Executive Functioning

Define “executive functioning”

Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. Clin Geriatr Med. 2013

“ability to self-monitor, <u>plan</u>, organize, <u>reason</u>,

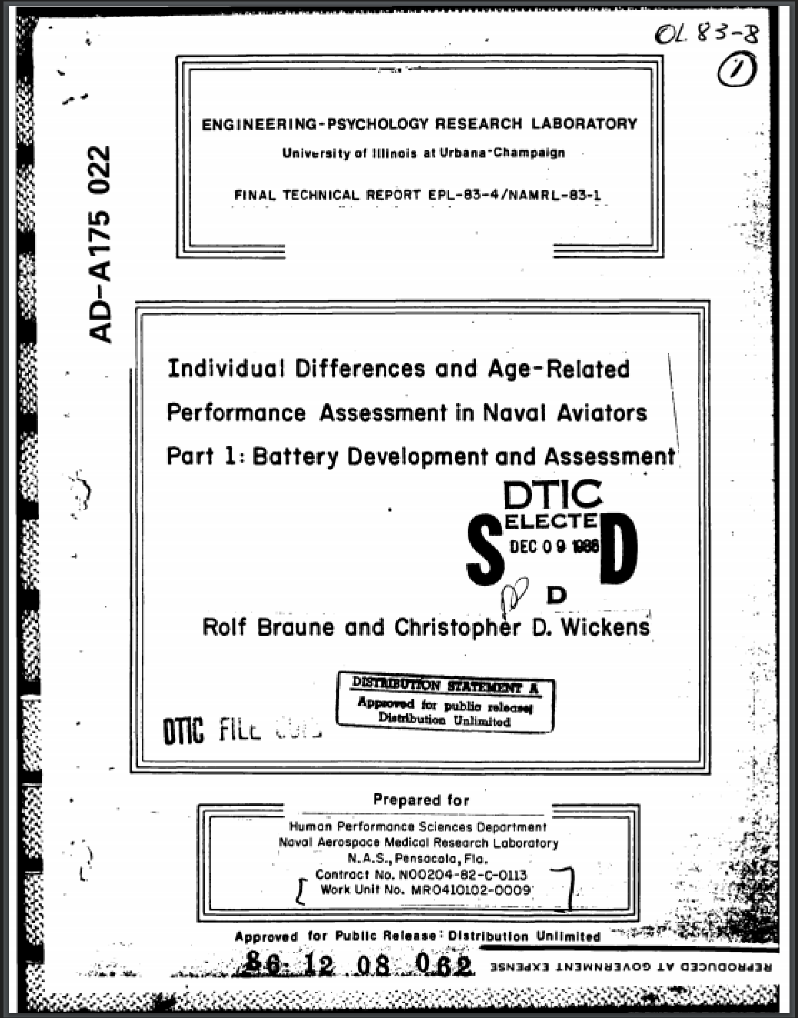
be <u>mentally flexible</u>, and <u>problem-solve</u>”

Define “cognitive” => aviation

Banich MT, Stokes A, Elledge VC. Neuropsychological screening of aviators: a review. Aviat Space Environ Med. 1989 Apr;60(4):361-6

Intelligence: Crystallized and Fluid	
Visuospatial Abilities/Construction	Spatial Abilities
Processing Speed	Perceptual Motor
Attention	Attention
Memory	Working Memory
Language	Processing Flexibility
Executive Functioning	Planning or Sequencing

Braune and Wickens
1983, NAMRL


OL 83-8

ENGINEERING-PSYCHOLOGY RESEARCH LABORATORY
University of Illinois at Urbana-Champaign

FINAL TECHNICAL REPORT EPL-83-4/NAMRL-83-1

AD-A175 022

Individual Differences and Age-Related Performance Assessment in Naval Aviators
Part 1: Battery Development and Assessment

DTIC ELECTE DEC 0 9 1986 S D D

Rolf Braune and Christopher D. Wickens

DISTRIBUTION STATEMENT A
Approved for public release; Distribution Unlimited

DTIC FILE COPY

Prepared for
Human Performance Sciences Department
Naval Aerospace Medical Research Laboratory
N.A.S., Pensacola, Fla.
Contract No. N00204-82-C-0113
Work Unit No. MR0410102-0009

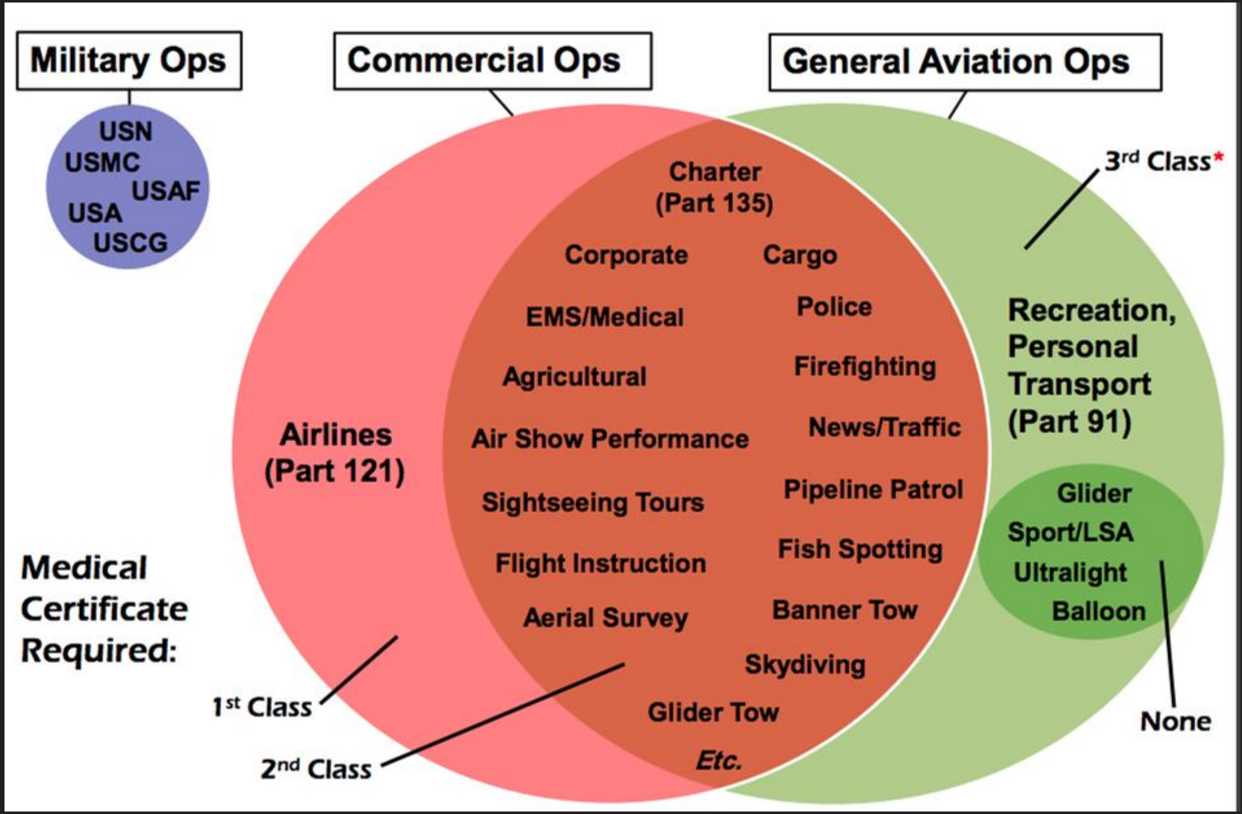
Approved for Public Release: Distribution Unlimited

86 12 08 062

REPRODUCED AT GOVERNMENT EXPENSE

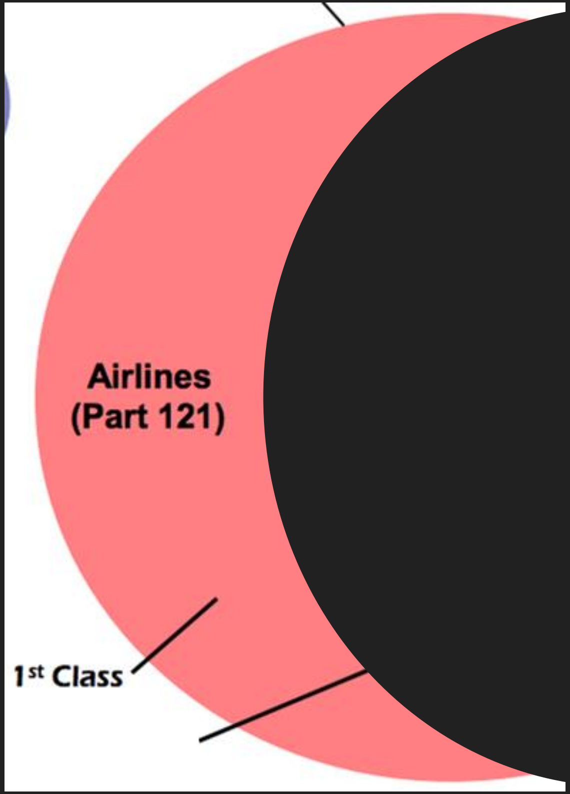
SCOPE

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

ESTIMATED ACTIVE PILOT CERTIFICATES HELD BY CATEGORY AND AGE GROUP OF HOLDER as of December 31, 2017

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

Flight Clearance: Operational Impacts

Restrictive	Lenient
Too few pilots	Loss of public confidence
Loss of experience	Higher near-miss risk?
Recruitment & Training	Higher mishap risk?
*Medical secrecy (*increased* risk?)	*Decreased stigma? (*better* MH care?)

Either way: bad press, political pressure, lawsuits...

AME Exams

+/- Neuropsych Testing

Flight Syllabus

Written, Simulator, Flight Tests

Impairment *versus* deficiency in neuropsychological assessment: Implications for ecological validity

NOAH D. SILVERBERG[1] AND SCOTT R. MILLIS[2]
[1]G.F. Strong Rehab Centre, Vancouver, British Columbia, Canada
[2]Department of Physical Medicine and Rehabilitation, Wayne State University School of Medicine, Detroit, Michigan

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.)
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

Focus on FAA Core Battery

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.)
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

FMCSA MEP: Truck Drivers and Stroke (2009)

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

FMCSA MEP: Truck Drivers and Stroke (2009)

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?

FMCSA MEP: Truck Drivers and Stroke (2009)

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 <u>no such currently available evidence exists</u>.

Definition: clinical outcome of interest?
Ability to handle complex emergency on any given flight?

FORMER AIR FORCE

FORMER NAVY

EVIDENCE REVIEW

Example

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

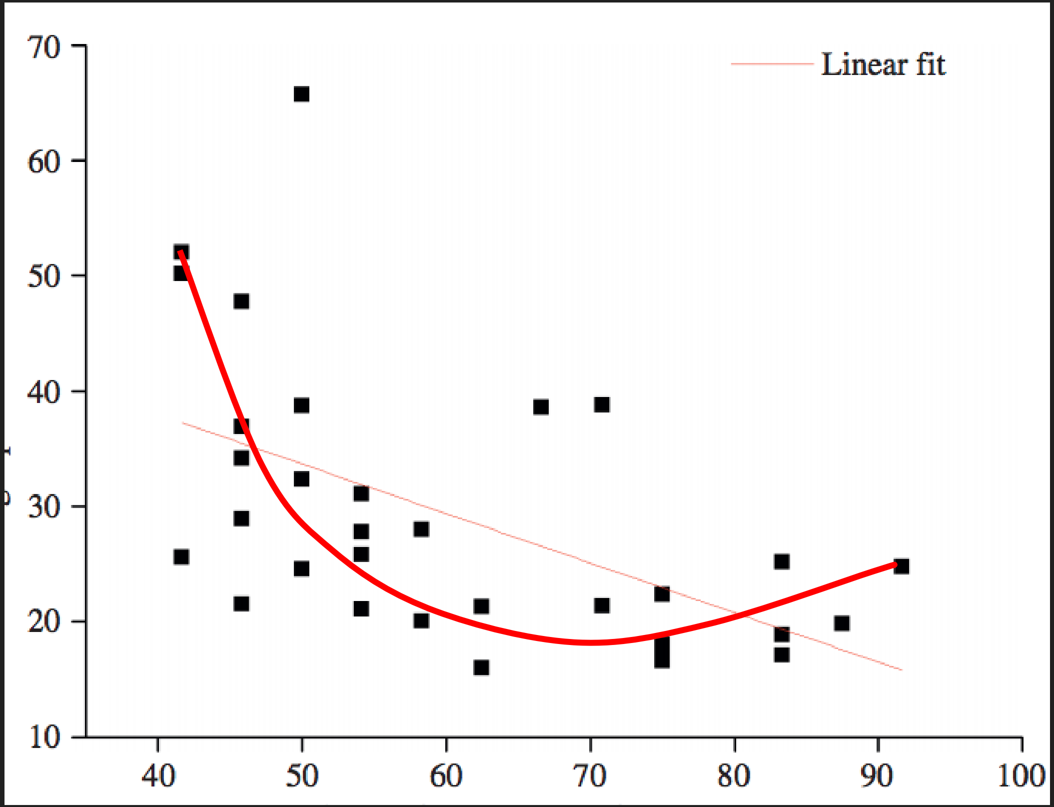
Year	Authors	Title	Journal	Public?	Location	Pop
2011	Causse, Dehais, Arexis & Pastor	Cognitive Aging and Flight Performance in General Aviation Pilots	Aging, Neuropsychology & Cognition	Y	France	General Aviation
(up to date)	(financial, academic bias) (confirmation bias)	(journal impact factor)	(peer-review, publication/reporting bias, article impact factor)	(publication bias)	(cross-cultural bias)	(relevance)

N	Evidence	Study Type	Hypothesis	Recruitment
32	3b	Exploratory Cohort, compared to poorly defined ref standard	chronological **age is not a sufficient criterion** to predict piloting performance and decision-making relevance and that **cognitive performance is a much more relevant criterion**	No information
(power)	(evidence hierarchy)	(design bias)	(clinically significant question?) (*a priori* hypothesis testing? or exporatory correlations?)	(selection bias)

Incl	Excl	# of Evaluators, Blinded?	Participants Blinded?	Predictor Variables
M, RHD, French, College+	Logicians, **airlines**, sens/neuro/psych deficits Emotional deficits (BIS/STAI) CNS-affecting substance use	1 No	32 No	**Chrono Age** **Chrono Flight Time** <u>NP Test Battery (FAA Core?):</u> **Target Hitting Test (No)** **2-Back Test (No)** **WCST (Yes)** **Spatial Stroop Test (Yes)**
(selection bias - convenience)	(selection bias - omission)	(observer bias)	(placebo effect)	(instrument bias)

Outcome Variables	Statistics	Confounding Age-Flight TIme
<u>Sim Flight Perf:</u> **Flight Path Deviations** (angular deviation in the horizontal axis from the ideal flight path) **Crosswind "no-land" decision** (<u>incorrect if inappropriate</u> with 6-knot CW tolerance)	Regression 1-way ANOVA	"No significant correlation" (p=.117, r=0.28)
(response bias, procedural bias)	(statistical assumptions)	(statistical assumptions)
Are these the **<u>outcomes of interest</u>** in aviation safety, as demonstrated by crashes, near-misses, incident reports, safety studies? **Does incorrect = unsafe?**	Assumes linear relationship **Why?**	Assumes linear relationship

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."

Causse, Dehais, Arexis & Pastor (2011). Cognitive Aging and Flight Performance in General Aviation Pilots. *Aging, Neuropsychology & Cognition, 18(5),* 544-561.

Case Example:

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

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

Significance of Neurocognitive Status

- Results: ~~Executive functions~~ (~~especially~~ working memory and set shifting) as measured by neurocognitive testing were ~~the best~~ a weak predictor of ~~in-flight~~ simulator 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~~ suggest that certain neuropsychological test performance may correlate with certain piloting and decision-making abilities in a small group of general aviation pilots

Causse, Dehais, Arexis & Pastor (2011). Cognitive Aging and Flight Performance in General Aviation Pilots. *Aging, Neuropsychology & Cognition, 18(5)*, 544-561.

Opportunities for evidence-based assessment (FAA)

1. Medical Re-Certification Process

   AME Designee <-> Specialists

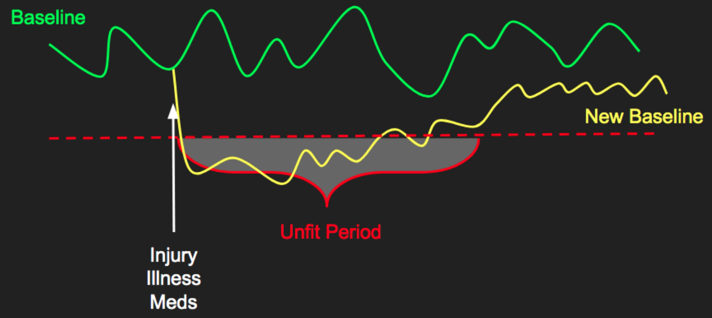
   => FAA Reviewers, determination

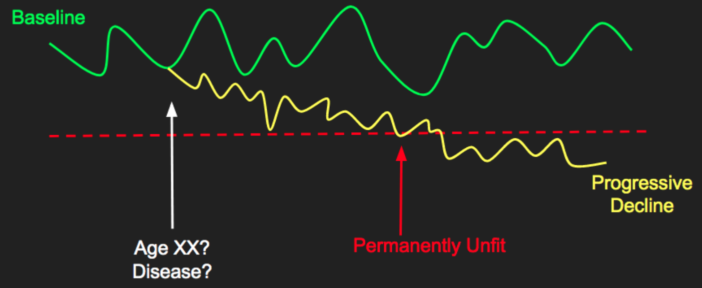
2. Simulator Syllabus (“functional evaluation”)

   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