COGNITIVE ASSESSMENT DURING LONG-DURATION SPACE FLIGHT

May 2011

Robert Kane, PhD, Washington DC
Kim Seaton, PhD, University of Texas Medical Branch
Walter Sipes, PhD, Johnson Space Center
I have no financial relationships to disclose.

I will not discuss off-label use and/or investigational use in my presentation.
Outline

- Background
- Implementation
- Defining off-nominal performance
- Overview of data from space
- Way forward
A recognized gap - no objective measure of cognitive performance was available in space operations.

Mir operations indicated a need for cognitive/performance assessment.

- work overload / underload
- on-board fire - contaminated air
- decompression - loss of module
- dangerous atmosphere (ethylene glycol and contaminant leaks)
- depression/withdrawal

A tool was needed that required minimal interaction with ground controllers due to communication limitations.
Therefore a tool was developed - Spaceflight Cognitive Assessment Tool for Windows (WinSCAT)

Purpose is to provide ISS flight surgeons with an objective clinical tool to monitor the astronauts’ cognitive status during long-duration space flight and allow immediate feedback to the astronaut

Developed for medical operations at NASA’s Johnson Space Center in Houston, Texas
Implementation
WinSCAT is a medical requirement for all long-duration missions

A computer-based, self-administered battery of five cognitive assessment subtests

In-flight tests scheduled monthly and compared against individual preflight baseline level of performance

Implemented with U.S.O.S. (United States Operational Segment) astronauts from one NASA/Mir mission and all expeditions to date (24) on the International Space Station (ISS).
1. Pre-Mission: Six one-half hour sessions for baseline data collection to:
   • Overcome learning effects
   • Develop individual baseline

2. During Mission: Medically required @ 30 days:
   • Maintain proficiency & update on-orbit baseline
   • May take as often as desired for self-assessment

3. Post-Mission: Minimum one session (R+30)
   • Check the re-adaptation to 1G before return to flight status (additional administrations as medically needed)
Viewing Results

• Test session is immediately scored, and all tests can be reviewed by the astronaut
• Off-nominal tests appear in red
• Monthly scores are downlinked to Behavioral Health & Performance Group
• Repeat testing may occur if off-nominal
• After a traumatic/medical event (0), WinSCAT testing would be indicated – contributes to medical decisions
Defining off-nominal performance
• Initial Scoring Rules
  – Based on validation study at the University of Northern Colorado
  – Change for accuracy and response time scores were based on a fixed baseline composed of last 3 preflight tests

• New Scoring Rules
  – Based on combining data from different studies incorporating initial data and data from studies using 6 baseline trials
  – Off-nominal based on atypical frequency of change value
  – Change for accuracy and throughput measured against moving baseline
  – Incorporates a summary score based on weighted throughput
Approaches To Assessing Change

I. Reliable Change Index

- **SEm:** Standard error of measurement
- **Sdiff:** Standard error of difference score
- **r:** Test-retest reliability
- **M2, M1:** Mean of T2 test score, Mean of T1 test score respectively
- **S1:** Standard deviation of control group, normal population
- **RC:** 90% confidence interval

\[
SE_m = s_1 \sqrt{1 - r}
\]

\[
S_{diff} = \sqrt{2(SE_m)^2}
\]

\[
RC = \pm 1.64 * S_{diff}
\]

\[
Mean\_Change = M_2 - M_1
\]

\[
Correction\_factor = Rounded(Mean\_Change)
\]

\[
Adjusted\_RC = Rounded(Correction\_factor + RC)
\]

II. Regression Equation: allows for adding different sources of variation

III. Frequency (normative) based
<table>
<thead>
<tr>
<th>Measure</th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
<th>Trial 5</th>
<th>Trial 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Code Substitution (N = 129)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>Mean(Sd) 50.30(10.18)</td>
<td>51.50(10.36)</td>
<td>54.51(11.44)</td>
<td>56.00(11.53)</td>
<td>57.48(11.76)</td>
<td>57.14(11.54)</td>
</tr>
<tr>
<td>RT</td>
<td>Mean(Sd) 1207(278)</td>
<td>1186(280)</td>
<td>1125(274)</td>
<td>1094(332)</td>
<td>1070(311)</td>
<td>1064(298)</td>
</tr>
<tr>
<td>Acc</td>
<td>Mean(Sd) 96.79(2.94)</td>
<td>97.31(2.34)</td>
<td>97.27(9.62)</td>
<td>96.67(2.96)</td>
<td>96.99(2.77)</td>
<td>96.08(3.52)</td>
</tr>
<tr>
<td><strong>Running Memory CPT (N=115)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>Mean(Sd)  97.29(20.85)</td>
<td>110.83(23.38)</td>
<td>115.17(22.39)</td>
<td>118.62(22.54)</td>
<td>120.38(23.26)</td>
<td>122.58(23.24)</td>
</tr>
<tr>
<td>RT</td>
<td>Mean(Sd)  567(112)</td>
<td>529(112)</td>
<td>511(103)</td>
<td>496(95)</td>
<td>492(98)</td>
<td>480(93)</td>
</tr>
<tr>
<td>Acc</td>
<td>Mean(Sd)  86.69(7.79)</td>
<td>92.66(5.76)</td>
<td>93.62(5.07)</td>
<td>94.09(4.30)</td>
<td>94.69(4.37)</td>
<td>94.31(4.96)</td>
</tr>
<tr>
<td><strong>Mathematical Processing (N =109)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>Mean(Sd)  25.89(6.58)</td>
<td>27.64(7.98)</td>
<td>27.00(7.22)</td>
<td>28.52(7.17)</td>
<td>31.51(7.60)</td>
<td>31.57(7.11)</td>
</tr>
<tr>
<td>RT</td>
<td>Mean(Sd)  2135(423)</td>
<td>2073(481)</td>
<td>2092(480)</td>
<td>1976(453)</td>
<td>1845(426)</td>
<td>1821(410)</td>
</tr>
<tr>
<td>Acc</td>
<td>Mean(Sd)  88.89(8.71)</td>
<td>90.59(8.63)</td>
<td>89.91(8.31)</td>
<td>89.76(7.59)</td>
<td>92.52(7.06)</td>
<td>92.16(7.31)</td>
</tr>
<tr>
<td><strong>Matching to Sample (N=128)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>Mean(Sd)  36.42(10.65)</td>
<td>38.57(11.77)</td>
<td>39.77(13.65)</td>
<td>37.87(11.63)</td>
<td>39.02(14.10)</td>
<td>39.29(13.36)</td>
</tr>
<tr>
<td>RT</td>
<td>Mean(Sd)  1608(430)</td>
<td>1559(436)</td>
<td>1557(454)</td>
<td>1607(478)</td>
<td>1568(496)</td>
<td>1555(517)</td>
</tr>
<tr>
<td>Acc</td>
<td>Mean(Sd)  92.97(6.86)</td>
<td>94.43(6.68)</td>
<td>95.31(6.11)</td>
<td>94.23(7.02)</td>
<td>93.80(6.16)</td>
<td>94.06(6.91)</td>
</tr>
<tr>
<td><strong>Code Substitution Recognition (N =127)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>Mean(Sd)  43.72(12.78)</td>
<td>39.84(11.40)</td>
<td>47.42(12.21)</td>
<td>50.32(12.90)</td>
<td>51.85(13.24)</td>
<td>44.72(12.84)</td>
</tr>
<tr>
<td>RT</td>
<td>Mean(Sd)  1276(400)</td>
<td>1361(407)</td>
<td>1178(303)</td>
<td>1122(274)</td>
<td>1123(281)</td>
<td>1191(310)</td>
</tr>
<tr>
<td>Acc</td>
<td>Mean(Sd)  88.06(11.98)</td>
<td>86.74(11.46)</td>
<td>90.33(10.24)</td>
<td>90.99(9.34)</td>
<td>92.73(7.81)</td>
<td>86.48(11.93)</td>
</tr>
</tbody>
</table>
### Table 1. Passing criteria (cut scores)

<table>
<thead>
<tr>
<th>Test</th>
<th>Throughput (change from baseline mean)</th>
<th>Accuracy (% correct responses)</th>
<th>Lapses (trials that time out with no responses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code sub (learning)</td>
<td>&gt; -7</td>
<td>&gt; 94</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Code sub (recall)</td>
<td>&gt; -15</td>
<td>&gt; 80</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Running Memory CPT (repeating numbers)</td>
<td>&gt; -8</td>
<td>&gt; 88</td>
<td>&lt;5</td>
</tr>
<tr>
<td>Matching to Sample</td>
<td>&gt; -8.5</td>
<td>&gt; 85</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Mathematical Processing</td>
<td>&gt; -5</td>
<td>&gt; 80</td>
<td>&lt;2</td>
</tr>
</tbody>
</table>

ICE score must be no less than 36 less than mean of baseline sessions
Moving Baseline for all baseline comparisons

### Table 2. Overall session score

<table>
<thead>
<tr>
<th>Score</th>
<th>Descriptor</th>
<th>Criteria</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Nominal</td>
<td>All scores passed AND ICE within expected range</td>
<td>green</td>
</tr>
<tr>
<td>1</td>
<td>Marginal</td>
<td>One score failed AND ICE within expected range</td>
<td>yellow</td>
</tr>
<tr>
<td>2</td>
<td>Off-Nominal</td>
<td>Two or more scores failed OR ICE below expected range</td>
<td>red</td>
</tr>
</tbody>
</table>
Overview of data from space
Analysis

• ISS performance data were assessed to compare initial to modified interpretation rules for detecting potential changes in cognitive functioning in space flight.

• Applying the newly derived rules to ISS data (172 total test sessions) resulted in 158 nominal test sessions (92%) and 14 off-nominal performances at various times during flight.

• Performance data from ISS missions do not indicate significant cognitive decrements due to microgravity/space flight alone but has shown decrements.
Explanations for off-nominal performances have included fatigue, post slam shifting, distraction of life events on Earth, distraction during testing, or minimal effort on the cognitive testing (obtained from post-mission debriefings).

Correlation to actual events is needed.
Way Forward
Refine Decision Rules

- Correlation to actual events is needed
- Challenge studies to further refine change score analysis
- Continue assessment using space flight data with incorporation of additional change metrics
WinSCAT has been updated to add network capability to support a 6-person crew on the station support computers.

WinSCAT Version 2.0.28 has increased difficulty of items in Mathematics, increased number of items in Match-To-Sample, incorporates a moving rather than fixed baseline, and implements stricter interpretation rules.