# Test-Retest Reliability and Effort on the CNS-Vital Signs Assessment

Ryan G. Jones, MA
Joseph P. Francis, Psy.D
Linda J. Baum, Ph.D.
Scott W. Sautter, Ph.D., ABN

# Abstract


- **Objective:** Computerized neuropsychological assessment programs are commonly used to aid in concussion diagnosis and management. Although test developers of these programs report moderate to good test-retest reliability, research has shown that some of these programs have low to moderate test-retest reliability when using clinically relevant time frames. This study examined the reliability of a computerized cognitive battery (CNS-Vital Signs) using a clinically relevant test-retest interval and controlling for effort. **Participants and Methods:** Thirty-one healthy college graduate adults completed select subtests of the CNSVS for the baseline and were retested approximately 30 days later. Subtests were selected to assess constructs similar to those measured by other computerized neuropsychological assessments: verbal memory, processing speed, executive functioning, and reaction time. Each participant also completed Green's Medical Symptom Validity Test and the Reliable Digit Span to evaluate effort. **Results:** Intraclass correlation coefficients were calculated on the computerized output scores to estimate test-retest reliability. Intraclass correlation coefficient estimates from baseline to retest ranged from .63 to .91 on the CNSVS. All participants demonstrated adequate levels of effort according to Medical Symptom Validity Test and Reliable Digit Span interpretive guidelines. **Conclusions:** This data demonstrated moderate to good test-retest reliability in a nonclinical sample performing with sufficient effort, especially when compared to the test-retest reliability of other computerized neuropsychological assessment programs using pragmatic time intervals.

# Current Literature

- The reliability of automated neuropsychological screening assessments has ranged widely.
- The developers of three widely used sports cognitive assessments report acceptable to good reliability of these tests (.67-.90). However, testing intervals were brief (7-14 days) and based on less sensitive statistical procedures.
- Independent research has shown the test-retest reliability of these assessments to be generally lower (.23-.65) than the acceptable minimal standard of .60 using appropriate statistical procedures and a more pragmatic interval between testing (45 days) (Broglio et al., 2007)

# Reliability of Computerized Concussion Assessments

| | Developers Report (7-14 days between testing periods) | Broglio et al., 2007 (45 days between testing periods) |
|---|---|---|
| **ImPACT** | **Pearson *r*** | **ICC** |
| Verbal Memory | .70 | .23 |
| Visual Memory | .67 | .32 |
| Visual-Motor Speed | .86 | .38 |
| Reaction Time | .79 | .39 |
| **Concussion Sentinel** | **ICC** | **ICC** |
| Reaction Time | .78 | .60 |
| Decision Making | .81 | .56 |
| Matching | .84 | .23 |
| Attention | .56 | .43 |
| Working Memory | .81 | .65 |
| **Concussion Resolution Index** | **Pearson *r*** | **ICC** |
| Simple Reaction Time | .73 | .65 |
| Complex Reaction Time | .72 | .43 |
| Processing Speed Index | .90 | .66 |

# Current Literature

- If the test-retest reliability is inflated then the Reliable Change Index (RCI) used to make return to play decisions may not be accurate.
- (Broglio et al., 2007) found that 19-38% of the participants exceeded the RCI cutoff score on the three computerized cognitive assessment when none of the participants had experienced a concussion
- Variability in cognitive performance can be attributed to multiple factors, such as fatigue, effort, and medication (Gualtieri & Johnson, 2006).
- Effort
  - Influences test performance and validity
  - Poor effort is associated with poor test performance. (Lezak, Howieson, & Loring, 2004; Griffin, Normington, & Glassmire, 1996; Gorp et al., 1999)
  - Unaccountable variations are difficult to interpret when effort is poor on baseline assessment. (Green et al., 2001)

# CNS-Vital Signs

- The CNS-Vital Signs (CNS-VS) is a computerized neuropsychological battery that was developed as a screening instrument and is comprised of seven neuropsychological tests that are well known and widely used.
- Specific tasks were adapted from the: Rey Auditory Verbal Learning Test; Rey Visual Design Test; Finger Tapping Test; Symbol Digit Modality Test; Stroop Test; Shifting Attention Test; and the Continuous Performance Test.
- Pearson correlation coefficients for the following domains were: .73 for Memory; .80 for Reaction Time; .87 for Psychomotor Speed; .74 for Cognitive Flexibility; and .64 for Complex Attention. (Gualtieri & Johnson, 2008).

# Hypotheses

- Purpose: To examine the test-retest reliability of CNS-VS with a pragmatic interval of time between pre and post tests while controlling for effort.
- Hypothesis 1: CNS-VS test-retest interclass correlation coefficients will be adequate relative to CNS-VS's reported test-retest coefficients on all domains of the assessment.
- Hypothesis 2: Subjects demonstrating suboptimal effort on the MSVT (defined by performance at or below 85 on the IR, DR, and/or CNS scales) and/or the RDS (defined by a RDS at or below 7) will achieve lower scores on CNS-VS subtests than those demonstrating adequate effort on measures of effort.

# Method

- Participants (N=31)
  - Adult student and non-student volunteers recruited through a graduate university population.
  - Exclusion Criteria
    - Diagnosed concussive injury within six months or during the study.
    - Suboptimal effort will be excluded from the test-retest reliability data set
- Measures
  - Medical Symptom Validity Test (MSVT)
  - Reliable Digit Span (RDS)
  - CNS-Vital Signs (CNSVS)
    - Verbal Memory, Symbol-Digit Coding, and Shifting Attention
  - Neurobehavioral Symptom Inventory

# Procedure

- Participants were administered portions of the CNSVS and effort measures at baseline and approximately 30 days later (m=30.9 ± 4.2 days).
- Questionnaire used to document demographics, concussion history, ADHD/LD history, drug/alcohol use, anxiety/depression, and engagement in sports/exercise.
- Order of test administration was identical in both test sessions.
  - The Verbal Memory, Shifting Attention, and Symbol Digit Coding subtests of the CNS-VS was administered to the participants.
  - These subtests were chosen to reduce administration time while still assessing for memory, attention, processing speed, and reaction time.
- Participants were evaluated individually in a quiet setting with one examiner.
- Administration time was approximately 30 minutes.

# Results: Demographics

| | Number | Percentage |
|---|---|---|
| **Gender** | | |
| Male | 16 | 53.3% |
| Female | 14 | 46.7% |
| **Age** | | |
| 23-26 | 18 | 60% |
| 27-29 | 9 | 30% |
| 30-36 | 3 | 10% |
| **Education** | | |
| College Degree | 5 | 16.7% |
| Graduate school; no masters degree | 9 | 30% |
| Masters Degree | 16 | 53.3% |
| **Race** | | |
| Caucasian | 24 | 80% |
| African American | 4 | 13.3% |
| Hispanic | 1 | 3.3% |
| Asian/Pacific Islander | 1 | 3.3% |

# Results: H1: Test-Retest Reliability of CNS-VS

- No significant problems with: normality, linearity, or homoscedasticity
- Similar results to reported reliability (Pearson *r)* by Gualtieri and Johnson (2006).
- A 2-way random effects analysis of variance ICC (2,1) was calculated to estimate reliability from baseline to retest.

| CNS-VS | ICC | (2006) *r* |
|---|---|---|
| Verbal Memory | .635 | .611 |
| Psychomotor Speed | .752 | .869** |
| Symbol Digit Coding | .747 | .840 |
| Cognitive Flexibility | .883* | .744** |
| Shifting Attention, Correct | .915* | .773 |
| Shifting Attention, Errors | .669 | .697 |
| Shifting Attention, Reaction Time | .913* | .803 |

* Indicates significantly improved performance on the retest (p.<.05)
** Additional subtests, not administered to our sample, were used to create this domain score

# Results: H2: Comparison of High/Low Effort group

- A score of pass/fail was determined using the following cutoff scores:
  - Cutoff score of MSVT was ≤ 85 on any one of the 3 effort/consistency scales imbedded in the test (IR, DR, and CNS) (Green, 2004)
  - Cutoff score of RDS was >7 digits forwards and backwards (Greiffenstein et al., 1994)
- No participant failed either of the effort tests at baseline or retest, therefore a comparison of performance could not be made

*% Correct on MSVT at Baseline and Day 30 (Mean ± SD)*

| MSVT | IR | DR | CNS | PA | FR |
|---|---|---|---|---|---|
| Baseline | 100 ± 0 | 99.3 ± 2.1 | 99.3 ± 2.1 | 100 ± 0 | 84.1 ± 9.1 |
| Retest | 99.5 ± 1.5 | 99 ± 2.4 | 98.5 ± 2.9 | 100 ± 0 | 84.6 ± 12.1 |

# Post Hoc Analysis

- A bivariate correlation was conducted to examine if NSI score correlated with performance on the CNS-VS.
- There were no significant correlations on baseline testing.
- At retest, NSI significantly correlated with Cognitive Flexibility ($r$=-0.37, $p<0.05$) and Shifting Attention errors ($r$=0.68, $p<0.01$)

# Discussion

- This study is unique in using an intraclass correlation coefficient to estimate test-retest reliability of several subtests of the CNS-VS while also assessing for effort.
- Research has shown considerable variability on test-retest reliability on other computerized assessments used for evaluation of sports-related concussions.
- Implications
  - This research project showed that selected subtests of the CNS-VS had acceptable to good test-retest reliability, which is consistent with prior research.
  - The findings indicate better reliability than other computerized sports concussion assessments when using clinically pragmatic time periods.
  - Effort testing aided in controlling for inadequate effort that could affect variable performance.
  - Practice effects were evident on several subtests.

# Discussion

- Limitations
  - Small sample size.
  - Sample was a well educated, motivated, high functioning, low-diversity group and results may not generalize to other groups.
  - Administration was 1:1 with participants whereas other research incorporated group administration.
  - Potential for interference from the MSVT on the Verbal Memory subtest of the CNS-VS.
- Directions for Future Research:
  - More research is needed to assess the reliability of the CNS-VS in group administrations, as this is the likely platform for baseline assessments of athletes.
  - While this project showed adequate reliability, future research will need to examine the validity of the CNS-VS at detecting impairment after a concussion when compared to baseline.
  - More research is needed in evaluating the relationship between effort and performance on concussion assessments for athletes.