

Psychometric and Measurement Properties of Concussion Assessment Tools in Youth Sports

Tamara C. Valovich McLeod*; William B. Barr†; Michael McCrea‡;
Kevin M. Guskiewicz§

*Arizona School of Health Sciences, A.T. Still University, Mesa, AZ; †New York University School of Medicine, New York, NY; ‡Waukesha Memorial Hospital, Waukesha, WI, and Medical College of Wisconsin, Milwaukee, WI; §University of North Carolina at Chapel Hill, Chapel Hill, NC

Tamara C. Valovich McLeod, PhD, ATC, CSCS, contributed to conception and design; acquisition and analysis and interpretation of the data; and drafting, critical revision, and final approval of the manuscript. William B. Barr, PhD, and Michael McCrea, PhD, contributed to analysis and interpretation of the data and drafting, critical revision, and final approval of the manuscript. Kevin M. Guskiewicz, PhD, ATC, contributed to conception and design; analysis and interpretation of the data; and drafting, critical revision, and final approval of the manuscript.

Address correspondence to Tamara C. Valovich McLeod, PhD, ATC, CSCS, Athletic Training Program, Arizona School of Health Sciences, A.T. Still University, 5850 E Still Circle, Mesa, AZ 85206. Address e-mail to tmcloed@atsu.edu.

Context: Establishing psychometric and measurement properties of concussion assessments is important before these assessments are used by clinicians. To date, data have been limited regarding these issues with respect to neurocognitive and postural stability testing, especially in a younger athletic population.

Objective: To determine the test-retest reliability and reliable change indices of concussion assessments in athletes participating in youth sports. A secondary objective was to determine the relationship between the Standardized Assessment of Concussion (SAC) and neuropsychological assessments in young athletes.

Design: We used a repeated-measures design to evaluate the test-retest reliability of the concussion assessments in young athletes. Correlations were calculated to determine the relationship between the measures. All subjects underwent 2 test sessions 60 days apart.

Setting: Sports medicine laboratory and school or home environment.

Patients or Other Participants: Fifty healthy young athletes between the ages of 9 and 14 years.

Main Outcome Measure(s): Scores from the SAC, Balance Error Scoring System, Buschke Selective Reminding Test, Trail

Making Test B, and Coding and Symbol Search subsets of the Wechsler Intelligence Scale for Children were used in the analysis.

Results: Our test-retest indices for each of the 6 scores were poor to good, ranging from $r = .46$ to $.83$. Good reliability was found for the Coding and Symbol Search tests. The reliable change scores provided a way of determining a meaningful change in score for each assessment. We found a weak relationship ($r < .36$) between the SAC and each of the neuropsychological assessments; however, stronger relationships ($r > .70$) were found between certain neuropsychological measures.

Conclusions: We found moderate test-retest reliability on the cognitive tests that assessed attention, concentration, and visual processing and the Balance Error Scoring System. Our results demonstrated only a weak relationship between performance on the SAC and the selected neuropsychological tests, so it is likely that these tests assess somewhat different areas of cognitive function. Our correlational findings provide more evidence for using the SAC along with a more complex neuropsychological assessment battery in the evaluation of concussion in young athletes.

Key Words: neuropsychological testing, brain injury, athletic injuries, reliability

Sport-related concussion is a significant problem in all levels of athletic participation. Because of the potential complications and long-term consequences of returning to competition too early,¹⁻³ sports medicine professionals are using more objective tools to assess athletes after a concussive injury. Several mental status and neuropsychological tests have been commonly used in high school^{1,4-6} and collegiate⁶⁻¹⁰ athletes to assess various cognitive domains after injury. Additionally, measures of postural stability have made their way into many concussion assessment protocols.¹¹⁻¹⁴

An ideal sport concussion assessment battery should consist of tests that are objective, reliable, valid, easy to administer,

and time efficient.¹⁵ Part of the efficiency equation is finding tests that assess the areas susceptible to deficits after concussion yet do not overlap in their assessment domains. Therefore, an understanding of the relationships among the various tests will aid the clinician in creating an appropriate test battery. Additionally, before such assessments are used to evaluate sport-related concussion, they should be validated through establishment of test-retest reliability, sensitivity, validity, reliable change index (RCI) scores, and clinical utility.¹⁶

Test-retest reliability denotes the correlation between 2 test administrations and refers to the stability of the instrument. The test-retest reliability on some measures of cognitive func-

Table 1. Participants' Descriptive Data (Mean ± SD)

	Males		Females	
	9 to 11 y	12 to 14 y	9 to 11 y	12 to 14 y
n	12	11	9	17
Age, y	10.7 ± .96	13.24 ± .87	10.00 ± 1.12	13.29 ± .83
Height, cm	143.93 ± 13.15	164.47 ± 11.91	140.12 ± 11.61	156.51 ± 6.26
Mass, kg	34.75 ± 6.71	54.42 ± 11.45	33.84 ± 5.53	46.10 ± 8.85
Grade, y	4.67 ± .78	7.18 ± 1.25	4.00 ± 1.12	7.17 ± .88

tion has been studied in professional rugby players.¹⁷ Test-retest values were moderate to good for the Speed of Comprehension Test ($r = .78$), the Digit Symbol Test ($r = .74$), and the Symbol Digit Test ($r = .72$). At the high school level, Barr¹⁸ evaluated the test-retest effects, RCIs, and sex differences on the Wechsler Intelligence Scale for Children (WISC-III) Digit Span and Processing Speed Tests, Trail Making Test, Controlled Oral Word Association Test, and the Hopkins Verbal Learning Test. Test-retest reliabilities ranged from $r = .39$ to $.78$.

Test-retest reliability for various computerized neuropsychological platforms has also been reported, most often demonstrating moderate to good reliability. Iverson et al¹⁹ found moderate to good reliability on the composite scores for Verbal Memory ($r = .70$), Visual Memory ($r = .67$), Reaction Time ($r = .79$), and Processing Speed ($r = .86$) of ImpACT, version 2.0. Reliability coefficients of $.68$ to $.82$ have been reported on the HeadMinder Concussion Resolution Index,²⁰ and moderate to good reliability has been reported using CogSport for speed indices ($r = .69$ to $.82$); however, lower reliability coefficients were reported for the accuracy indices ($r = .31$ to $.51$).²¹

Interest has grown recently in defining the most clinically useful methods for detecting change in neuropsychological test scores when utilizing the test-retest paradigm. In interpreting postconcussion test scores, statistical methods involving a control group are not necessarily helpful for the clinician who needs to determine whether fluctuations on one athlete's assessment represent meaningful changes or normal variability in performance.²² Much attention has focused on the use of the RCI and standardized regression-based methods, techniques developed and refined in studies of outcomes from psychotherapy and surgical treatment. The RCI has been useful to help account for differences in the test-retest reliability and practice effects with serial testing and has been utilized to assess intraindividual differences over time.^{17,23-25} The RCI analysis includes adjustments for practice effects and can help with the predictive capabilities of a test instrument.²³ Reliable change scores have been published for both pencil-and-paper^{17,18} and computerized test batteries.¹⁹

Although cognitive and postural stability assessments for sport-related concussion now are often performed on high school athletes,^{4,18} research into the use of these tools in athletes younger than high school age has been limited. Recently, Valovich McLeod et al²⁶ found the Standardized Assessment of Concussion (SAC) and the Balance Error Scoring System (BESS) to be appropriate tests for 9- to 14-year-old athletes. They also demonstrated a practice effect with serial BESS testing (subjects improved on their balance performance by the third time they performed the task)²⁶; however, little is known about the psychometric and measurement properties of other concussion assessments in this younger population. Therefore,

our purpose was to determine the test-retest reliability and RCIs of cognitive and balance tests in a youth sports population. Because many different assessment tools are available to the clinician and because little is known about the relationships among these tools, our secondary purpose was to determine the relationship between the SAC and several neuropsychological assessments in this age group.

METHODS

We used a quasi-experimental, repeated-measures design to evaluate the test-retest reliability of the concussion assessments in young athletes. All subjects underwent an initial test consisting of administration of the SAC, BESS, and a neuropsychological test battery designed for children between the ages of 9 and 14 years. All subjects returned approximately 60 days after their initial test for 1 follow-up test session (mean test-retest interval = 57.94 ± 4.15 days). This time interval was chosen to reflect an interval between baseline testing and the latter part of an athletic season¹⁸ and represented an adequate time frame for demonstrating the test-retest reliability of measures used to study sport-related concussion.¹⁶

Subjects

Fifty youth sport participants were recruited from the local community to participate in this study. Descriptive data are presented in Table 1. Male ($n = 24$) and female ($n = 26$) participants were selected based on the following general criteria: (1) participation in recreational or competitive athletics (baseball, softball, soccer, gymnastics, or swimming), (2) no lower extremity musculoskeletal injuries in the 6 months before testing, (3) no history of head injury, (4) no diagnosed visual, vestibular, or balance disorders, and (5) no diagnosis of attention deficit disorder or learning difficulty. All inclusion and exclusion criteria were determined from self-report and parental report. Before participation, each subject and his or her parent or guardian read and signed an informed consent form approved by the university's Institutional Review Board for the Protection of Human Subjects, which also approved the study. The test-retest data were excluded for a control subject who sustained a concussion before the follow-up test.

Instrumentation

Neuropsychological Test Battery. The neuropsychological test battery consisted of 4 tests with established norms for this age group: the Buschke Selective Reminding Test (SRT), Trail Making Test B (Trails B), and the Symbol Search and Coding subsets of the WISC-III (The Psychological Corp, San Antonio, TX). We chose each neuropsychological assessment be-

cause of the cognitive domain assessed as well as the similarity to the neuropsychological batteries commonly employed to assess older athletes after concussion.

The Buschke SRT was used to measure verbal learning and memory during a multiple-trial list-learning task.²⁷ The test involved reading the subject a list of words and then having the subject recall as many of the words as possible. Each subsequent reading of the list involves only the items that were not recalled on the immediately preceding trial. For the age group in our study, we presented 12 words for 8 trials or until the child recalled all 12 words on 2 consecutive trials. Scores were calculated for sum total (total number of words recalled), continuous long-term recall (CLTR) (number of words recalled continuously), and delayed recall (number of words recalled after 20 minutes). The 2 alternate forms of the SRT for children were counterbalanced between subjects and test sessions.²⁸

The Trails B from the Halstead Reitan Neuropsychological Test Battery (Reitan Neuropsychological Laboratory, Tucson, AZ) was used to test speed of attention, sequencing (a measure of mental organization and tracking with a set order of priority), mental flexibility, and visual scanning (measure of visual processing and target detection) and motor function.²⁹ The test was essentially a connect-the-dots activity, which required that the child alternate between the numeric and alphabetic sequencing systems (progressing from 1 to A to 2 to B to 3, etc). We used the children's, or intermediate, form for this study. We recorded the length of time it took the participant to complete the task and the number of errors committed.

The Symbol Search subset of the WISC-III assessed attention, visual perception, and concentration.³⁰ The subjects scanned 2 groups of symbols and indicated whether the target symbol appeared in the search group. They had 120 seconds to complete as many items as possible.

The Coding subset of the WISC-III assessed processing speed (rate of cognitive processing), concentration, and attention.³⁰ The testing tool consisted of rows of blank squares with numbers 1 to 9 written above each square. A key printed on the top of the test sheet paired each number with a symbol. The child's task was to fill in the blanks with the correct symbol as quickly as possible. The child was given 90 seconds to complete the coding task.

Standardized Assessment of Concussion. The SAC is an instrument designed to assess acute neurocognitive impairment on the sideline and includes measures of orientation, immediate memory, concentration, and delayed recall.^{6,31} The instrument requires 5 to 7 minutes to administer and was designed for use by individuals with no prior expertise in neurocognitive test administration, including certified athletic trainers. Alternate forms (A, B, and C) of the SAC were designed to minimize practice effects during follow-up testing. Previous researchers^{32,33} have demonstrated multiple form equivalence with no differences among the 3 forms. Orientation was assessed by asking the subject to provide the day of the week, date, month, year, and time. A 5-word list of unrelated terms was used to measure immediate memory. The list was read to the subject for immediate recall and the procedure repeated for a total of 3 trials. Concentration was assessed by having the subject repeat strings of numbers in the reverse order of their reading by the examiner and by reciting the months of the year in reverse order. Delayed recall of the 5-word list was also recorded. A composite score, with a max-

imum of 30 points, was derived. We used all 3 SAC forms and counterbalanced them among subjects and test sessions.

Balance Error Scoring System. The BESS consists of 6 separate 20-second balance tests that the subjects perform in different stances and on different surfaces.³⁴ A 16-in \times 16-in- (40.64-cm \times 40.64-cm-) piece of medium-density (60 kg/m³) foam (Exertools, Inc, Novato, CA) was used to create an unstable surface for the subjects. The test consisted of 3 stance conditions (double leg, single leg, and tandem) and 2 surfaces (firm and foam). Errors were recorded as the quantitative measurement of postural stability under different testing conditions. These errors included (1) opening the eyes, (2) stepping, stumbling, and falling out of the test position, (3) lifting the hands off the iliac crests, (4) lifting the toes or heels, (5) moving the leg into more than 30° of flexion or abduction, and (6) remaining out of the test position for more than 5 seconds.

Procedures

Testing was conducted either in the Sports Medicine/Athletic Training Research Laboratory at the University of Virginia or at the child's school or home. For each subject, testing was performed in a quiet room at the same location for both test administrations. Approximately half of the subjects were tested at the University and half at their schools or homes. Testing consisted of 2 test sessions scheduled approximately 60 days apart. Participants were not restricted from sport participation or recreational activities in the time between the test sessions. A single investigator performed all test administrations to ensure optimal consistency of administration procedures. No other individuals were present in the room during the test sessions. Before data collection, the principal investigator completed training in administration of the neuropsychological assessments with a pediatric neuropsychologist and laboratory technician in the Neuropsychology Assessment Laboratory at the University of Virginia.

During each test session, the assessments were performed in the following order: SAC, BESS, Buschke SRT, Trails B, Coding, Symbol Search, and Delayed SRT. The alternative forms of the SAC and Buschke SRT were used and counterbalanced among subjects and test sessions. The neuropsychological test battery took approximately 15 minutes to administer and consisted of the Buschke SRT, Trails B (version for those 9 to 14 years of age), and the Coding and Symbol Search subsets of the WISC-III.

The BESS testing took approximately 10 minutes per subject, and all scores were recorded on a form by the primary investigator. The order of trials followed a format that progressively increased the demands placed on the sensory systems: double leg, single leg, and tandem on firm surface and then foam surface. To ensure consistency among subjects tested at different sites, all trials on the firm surface were performed on a thin carpet over a tile or linoleum floor. Subjects were asked to assume the required stance by placing their hands on their iliac crests, and once they closed their eyes, the test began. During the single-leg stances, the subjects were asked to maintain the contralateral limb in 20° of hip flexion and 40° of knee flexion. Additionally, subjects were asked to stand quietly and as motionless as possible in the stance position, keeping their hands on their iliac crests and their eyes closed. Subjects were told that upon losing their balance, they should make any necessary adjustments and return to the stance position as quickly as possible. Performance was as-

Table 2. Initial Test Session Scores for the Entire Sample

Test	Total Sample, (mean ± SD), n = 49	Males (mean ± SD), n = 23	Females (mean ± SD), n = 26	Younger (mean ± SD), n = 21	Older (mean ± SD), n = 28	Analysis of Variance Results		
						Main Effect for Sex (<i>P</i>)	Main Effect for Age Group (<i>P</i>)	Sex-by-Age Group Interaction (<i>P</i>)
Standard Assessment of Concussion, raw score	26.86 ± 2.07	26.22 ± 2.19	27.50 ± 1.79	26.95 ± 1.88	26.86 ± 2.24	.035*	.568	.476
Balance Error Scoring System, errors	15.54 ± 5.82	17.39 ± 5.19	13.54 ± 5.64	18.47 ± 4.74	13.00 ± 5.31	.061	.001*	.189
Buschke Selective Reminding Test Sum, total words	78.60 ± 7.93	76.83 ± 5.84	80.50 ± 9.14	78.76 ± 8.75	78.79 ± 7.39	.099	.784	.677
Continuous long-term recall, total words	64.24 ± 17.5	60.61 ± 15.65	67.88 ± 18.80	67.86 ± 18.04	61.92 ± 17.15	.088	.151	.683
Delayed recall, total words	10.52 ± 1.83	10.30 ± 1.74	10.65 ± 1.94	10.48 ± 1.44	10.50 ± 2.12	.527	.943	.971
Trail Making Test B, s	39.54 ± 21.74	39.49 ± 11.41	39.43 ± 28.48	47.63 ± 29.79	33.34 ± 10.44	.568	.022*	.206
Symbol Search, raw score	32.6 ± 5.89	32.17 ± 4.85	33.08 ± 6.83	30.76 ± 7.17	34.07 ± 4.45	.854	.067	.955
Coding, raw score	44.16 ± 9.56	40.91 ± 8.05	46.73 ± 10.16	38.19 ± 7.33	48.36 ± 8.83	.105	<.001*	.423

*Significant difference (*P* ≤ .05).

sessed by individual trial scores and by adding the error points for each of the 6 trials. Trials were considered incomplete if the subject could not sustain the stance position for longer than 5 seconds. In these instances, subjects were then assigned a standard maximum score of 10 for that stance,³⁴ a situation that occurred on only 3 occasions (once each in 3 subjects).

Statistical Analyses

Separate scores from the Buschke SRT trials were calculated for a sum total score (SRT Sum) and a CLTR, creating 6 neuropsychological test scores in our analyses. A 2 × 2 analysis of variance was used to evaluate sex and age group differences in the initial test data. We used intraclass correlation coefficients (ICC[2,1]) and the Pearson product moment correlation (*r*) to determine the test-retest reliability for each assessment. Pearson correlations were included because they have commonly been used in the test-retest literature regarding

concussion assessments¹⁷⁻¹⁹ and are needed to calculate the RCI scores. Coefficients of less than .50 were considered to indicate poor reliability; coefficients measuring from .50 to .75 indicated moderate reliability; and coefficients measuring greater than .75 indicated good reliability.³⁵ Separate paired-samples *t* tests were performed to analyze whether significant differences existed between the initial test and the retest. Reliable change index scores were calculated using the results from the Pearson correlations and the SD of the initial score, as recommended by Jacobson and Truax.³⁶ We corrected for practice effects by adding the mean change score to the confidence interval, as suggested by Chelune et al.³⁷ We also used Pearson product moment correlations to determine the relationship between the SAC and each of the 6 neuropsychological test scores. All analyses were performed using SPSS (version 12.0; SPSS Inc, Chicago, IL), and significance was set a priori at *P* ≤ .05.

RESULTS

Initial Testing

The means and SDs for the entire sample from the initial test session are presented in Table 2. Separate results are also listed for male and female athletes and for younger (9-to-11-year-old) and older (12-to-14-year-old) athletes, chosen to represent the younger and older 3-year spans in our sample. Significant differences were found between the sexes on the SAC, with females outperforming males. With respect to age, the older athletes scored better on the BESS, Trails B, and Coding tests; no significant differences were noted on the other tests.

Test-Retest Reliability

Our reliability values are presented in Table 3, and the test scores for the initial and retest sessions are listed in Table 4.

Table 3. Reliability of Assessments Across the Entire Sample

	Intraclass Correlation Coefficient		Pearson <i>r</i>
	(2,1)	SEM	
Standard Assessment of Concussion, raw score	.46	1.5	.46
Balance Error Scoring System, errors	.70	3.3	.70
Buschke Selective Reminding Test Sum, total words	.63	6.2	.65
Continuous long-term recall, total words	.52	14.2	.52
Delayed recall, total words	.51	1.4	.51
Trail Making Test B, s	.65	13.2	.71
Symbol Search, raw score	.83	2.8	.83
Coding, raw score	.82	4.3	.82

Table 4. Test Scores From the Initial and Retest Sessions for the Entire Sample

Test	Initial Session			Retest Session			Retest-Initial	Initial, Retest		
	Mean	SD	Range	Mean	SD	Range	Mean	<i>r</i>	SEM	<i>S</i> _{diff}
Standard Assessment of Concussion, raw score	26.86	2.07	20.0–30.0	27.14	1.93	23.0–30.0	+0.28	.46	1.52	2.15
Balance Error Scoring System, errors	15.54	5.82	5.0–26.0	13.47	5.64	2.0–27.0	–2.07	.70	3.17	4.48
Buschke Selective Reminding Test Sum, total words	78.60	7.93	57.0–92.0	80.71	10.00	51.0–96.0	+2.11	.65	4.70	6.64
Continuous long-term recall, total words	64.24	17.5	28.0–91.0	66.86	20.37	12.0–96.0	+2.62	.52	12.12	17.15
Delayed recall, total words	10.52	1.83	4.0–12.0	10.78	2.04	3.0–12.0	0.26	.51	1.28	1.81
Trail Making Test B, s	39.54	21.74	19.71–169.15	36.14	14.01	16.03–93.95	–3.40	.71	11.67	16.50
Symbol Search, raw score	32.6	5.89	12.0–45.0	34.43	6.16	12.0–45.0	+1.83	.83	2.44	3.46
Coding, raw score	44.16	9.56	22.0–70.0	45.61	9.82	23.0–68.0	+1.45	.82	4.05	5.72

Our test-retest indices for each of the 8 assessments ranged from poor (ICC = .46, *r* = .46) to good (ICC = .83, *r* = .83). We found significant decreases in BESS errors ($t_{48} = 3.010, P = .004$) and the time to complete the Trails B ($t_{48} = -3.496, P = .001$) at the retest session compared with the initial test (Table 5). For both of these assessments, a significant decrease demonstrated improved performance. No other assessments were significantly different between test sessions (Table 5). We did find some observable differences in the test-retest reliability between the sexes and between the 9-to-11-year-old and the 12-to-14-year-old subjects (Table 6).

Reliable Change Indices

The RCIs for the assessments are listed in Table 7 for the 90%, 80%, and 70% confidence intervals (CIs) for both the raw score and the whole-number units. Based on the 70% CI, which is the most conservative index, a decrease in 2 SAC points, 5 Buschke SRT words, 16 CLTR words, 2 Coding points, 2 Symbol points, and 2 SRT delayed recalls, as well as an increase of 3 BESS errors and 14 seconds in the Trails B time, would indicate a change in performance consistent with impairment on these measures.

Relationship Between Standard Assessment of Concussion and Neuropsychological Tests

The correlation matrix describing the relationship between the SAC and the neuropsychological test scores for all 50 sub-

jects during their initial test is found in Table 8. We noted significant ($P < .05$) positive correlations between the SAC and 4 of the 6 neuropsychological scores: SAC and Symbol Search ($r = .32$), SRT Sum ($r = .28$), and SRT Delayed Recall ($r = .36$). The SAC and Trails B were negatively correlated ($r = -.29$). Higher scores on the SAC tended to be associated with higher scores on the Symbol Search, SRT Sum, and SRT Delayed Recall and with faster times on the Trails B.

We found some higher correlations among the neuropsychological assessments, with the Trails B negatively correlated with the Coding ($r = -.53$) and Symbol Search ($r = -.73$). Positive correlations were found between the Symbol Search and Coding ($r = .65$), the SRT Sum and CLTR ($r = .92$), and the SRT Sum and SRT Delayed Recall ($r = .54$) tests.

DISCUSSION

Our findings provide some insight into the psychometric and measurement properties of various concussion assessment tools that could be used to evaluate concussion in young athletes. Although more evidence exists on the use of various assessments in professional and collegiate athletes and although high school athletes are increasingly being studied, our investigation is one of the first to research the measurement properties of neuropsychological and balance tests in a youth sports population.

Initial Testing

Although both age and sex effects on neuropsychological test performance have been studied in high school athletes,¹⁸ published data on these variables in professional, collegiate, or younger athletes are limited. We did find significant sex differences on performance of the SAC, with females scoring higher than males. Although previous authors³¹ have indicated a trend toward females achieving slightly higher scores, these findings did not reach statistical significance. Our lack of significant differences with respect to the verbal memory tests (Buschke SRT Sum, CLTR, Delayed Recall) is surprising considering that both age and sex affect performance on the Buschke SRT.³⁸ However, the SRT is also moderately related to IQ, which was not measured in this investigation. Our findings indicate that separate norms for males and females should be used on the SAC for younger athletes, as reported previously for high school athletes.¹⁸

Table 5. Paired-Samples *t* Tests Between the Initial and Retest Sessions

Test	<i>t</i>	<i>P</i>
Standard Assessment of Concussion, raw score	–.8225	.413
Balance Error Scoring System, errors	3.010	.004*
Buschke Selective Reminding Test Sum, total words	1.503	.139
Continuous long-term recall, total words	–1.943	.058
Delayed recall, total words	–1.039	.304
Trail Making Test B, s	–3.496	.001*
Symbol Search, raw score	–.891	.377
Coding, raw score	–1.754	.086

*Indicates significantly improved performance on the retest ($P < .05$).

Table 6. Test-Retest Reliability by Sex and Age Group (Intraclass Correlation Coefficient [2,1])

	Males (n = 23)	Females (n = 26)	Younger (n = 21)	Older (n = 28)
Standard Assessment of Concussion, raw score	.41	.48	.43	.49
Balance Error Scoring System, errors	.75	.61	.56	.68
Buschke Selective Reminding Test				
Sum, total words*	.38	.77	.66	.61
Continuous long-term recall, total words*†	.40	.57	.60	.48
Delayed recall, total words*†	.17	.74	.59	.47
Trail Making Test B, s*	.45	.70	.61	.55
Symbol Search, raw score*	.68	.90	.83	.78
Coding, raw score*	.70	.86	.82	.74

*Indicates observable difference (>.10) between the sexes.
 †Indicates observable difference (>.10) between the age groups.

We also found that older athletes (aged 12 to 14 years) performed better on the BESS, Coding, and Trails B than did the younger athletes in our sample. Although our groups are close in age from a developmental standpoint, the finding that the older group performed better on the Trails B and Coding is substantiated by improvement in the age-appropriate norms cited in the literature for those tests.^{30,38} The age-appropriate norms for the Trails B indicate that performance improves with each year's increase in age,³⁸ whereas the norms for the WISC-III Coding improve with each 3-month increase in age.³⁰ With respect to the BESS, previous authors³⁹ have demonstrated improved performance on postural stability tests with increasing age. In addition, investigators who were specifically using the BESS as a measure of postural stability have found lower BESS scores in healthy college⁴⁰ and high school athletes,⁴¹ compared with our findings.

Test-Retest Reliability

Test-retest reliability is important in all measures as a means to identify practice effects, a factor that could influence the test result. During serial testing, assessments with low reliability can cause profound variability in the scores of individuals with no alterations in cognitive function or deficits in

Table 8. Correlations Between the Standard Assessment of Concussion and Neuropsychological Tests

	Selective Reminding Test					
	Trail Making Test B	Symbol Coding	Sum Recall	Continuous Long-Term Recall	Delayed Recall	
Total Standard Assessment of Concussion	-.29*	.27	.32*	.28*	.22	.36†
Trail Making Test B	—	-.53‡	-.73‡	-.45‡	-.33*	-.37†
Coding	—	—	.65‡	.22	.11	.21
Symbol Search	—	—	—	.27*	.27	.16
Buschke Selective Reminding Test						
Sum recall	—	—	—	—	.92‡	.54‡
Continuous long-term recall	—	—	—	—	—	.41†

*P ≤ .05.
 †P ≤ .01.
 ‡P ≤ .001.

balance ability. Test-retest reliability is the first step in the process to validate cognitive batteries for the assessment of concussion.¹⁶ In the sport-related concussion literature, reports have been published on the test-retest reliability of paper-and-pencil neuropsychological assessments^{17,18} computerized test batteries,¹⁹⁻²¹ and the SAC in older populations.²³ However, to date, no authors have reported the test-retest reliability of the BESS or neuropsychological assessments specific to the pediatric athletic population.

With respect to the cognitive assessments, we found poor to good reliability coefficients, ranging from .46 to .83. We noted lower test-retest reliability (ICC = .51 to .65, *r* = .51 to .65) in the tests that assessed verbal learning and memory and the SAC (ICC = .46, *r* = .46). We did show differences in the reliability of the verbal learning tests between male and female subjects, with females exhibiting better test-retest reliability on the Buschke SRT, CLTR, and Delayed Recall. Further investigation of these differences revealed an outlier in the males, which likely led to lower reliability in the male subjects. Slight practice effects were observed on most of the tests. Significant changes at time 2, indicating practice effects, were observed on only 2 of the measures (Trails B, BESS).

Table 7. Adjusted Reliable Change Indices Calculated for 90%, 80%, and 70% Confidence Intervals

Test	90% Confidence Interval		80% Confidence Interval		70% Confidence Interval	
	Raw Scores	Test Units	Raw Scores	Test Units	Raw Scores	Test Units
Standard Assessment of Concussion, raw score	-3.2, +3.8	±4	-2.5, +3.1	-3, +4	-2.0, +2.5	-2, +3
Balance Error Scoring System, errors	-9.4, +5.3	-10, +6	-7.9, +3.8	-8, +4	-6.8, +2.6	-7, +3
Buschke Selective Reminding Test						
Sum, raw score	-8.8, +13.0	-9, +13	-6.5, +10.8	-7, +11	-4.9, +9.1	-5, +10
Continuous long-term recall, raw score	-25.5, +30.7	-26, +31	-19.7, +24.9	-20, +25	-15.4, +20.6	-16, +21
Delayed recall, raw score	-2.7, +3.2	-3, +4	-2.1, +2.6	±3	±1.90	±2
Trail Making test B, s	-30.4, +23.7	-31, +24	-24.9, +18.1	-25, +19	-20.7, +13.9	-21, +14
Symbol Search, raw score	-3.8, +7.5	-4, +8	-2.7, +6.3	-3, +7	-1.6, +2.2	-2, +3
Coding, raw score	-7.9, +10.8	-8, +11	-6.0, +8.9	-6, +9	-1.8, +5.5	-2, +6

These findings are consistent with those of previous authors who were testing both uninjured high school athletes¹⁸ and healthy adults.²⁵ In a sample of high school athletes, Barr¹⁸ reported reliability coefficients of $r = .54$ for the Hopkins Verbal Learning Test total score and $r = .56$ for the Hopkins Verbal Learning Test Delayed Recall score. Similarly, the test-retest reliability of adults demonstrated poor to moderate reliability on the Buschke SRT total ($r = .62$), CLTR ($r = .54$), and SRT Delayed Recall ($r = .46$)²⁵ tests and on the California Verbal Learning Test ($r = .29$ to $.67$).⁴²

We did find moderate to good test-retest indices (ICC = $.65$ to $.83$, $r = .71$ to $.83$) for the tests that assessed attention, concentration, and visual processing, including the Trails B, Coding, and Symbol Search tests. Two groups studying adult populations reported higher reliability coefficients than we found; however, this could be due to the increased variability in performance often seen on these tests in younger subjects.³⁸ Dikmen et al²⁵ noted good test-retest reliability for the Trails B ($r = .89$) and the Digit Symbol Test ($r = .89$) in subjects with a mean age of 43.6 years after a test-retest interval that ranged from 2 to 12 months, whereas Hinton-Bayre et al⁴³ also found good reliability coefficients in professional rugby players (19.4 ± 2.1 years of age) on the Speed of Comprehension Test ($r = .78$), Digit Symbol ($r = .74$), and Symbol Digit ($r = .72$) assessments after a 1- to 2-week test-retest interval. In contrast, Barr¹⁸ found lower test-retest reliability on the Symbol Search ($r = .58$), Trails B ($r = .65$), and Digit Symbol ($r = .73$) tests in high school athletes with a mean age of 15.9 ± 0.98 years, with a test-retest interval of 60 days.

Test-retest reliability of composite scores for various computerized neuropsychological platforms has also been reported. Using a 2-week test-retest interval on the HeadMinder Concussion Resolution Index (HeadMinder Inc, New York, NY), good reliability was found on indices for Simple Reaction Time ($r = .70$) and Processing Speed ($r = .82$), whereas moderate reliability was noted for Complex Reaction Time ($r = .68$).²⁰ Similarly, Iverson et al¹⁹ reported moderate to good reliability on the composite scores for Verbal Memory ($r = .70$), Visual Memory ($r = .67$), Reaction Time ($r = .79$), and Processing Speed ($r = .86$) of ImPACT (version 2.0; ImPACT Applications Inc, Pittsburgh, PA) after a mean test-retest interval of 5.8 days. Using a 1-week test-retest interval, moderate to good reliability has been reported using CogSport for speed indices ($r = .69$ to $.82$); however, lower reliability coefficients were noted for the accuracy indices ($r = .31$ to $.51$).²¹ Although most of the measures reported from the computerized assessments demonstrate moderate to good reliability, all of the authors mentioned above studied test-retest reliability over a shorter period of time (1 to 2 weeks) rather than the 60 days used in our investigation. It should be noted that both the age of the participants and the test-retest interval may affect the test-retest reliability coefficients reported across the various studies.

One contributing factor that may explain our low test-retest reliability on some assessments is the large SEM value. For example, our scores for the CLTR ranged from 28 to 91 on the initial test and from 12 to 96 on the retest, with an SEM of 12.12 and an S_{diff} of 17.15. Similarly, we had large SEM (11.67) and S_{diff} (16.50) values with the Trails B. However, these findings are less surprising for the latter test, given the results of other sport-related concussion studies in which the Trails B was administered without prior administration of the Trail Making Test A. Guskiewicz et al⁴⁴ reported SDs ranging

from 14.09 to 18.23 for the Trails B in a collegiate population, and McCrea et al⁹ noted SDs of 18.69 and 22.12 in collegiate control athletes and those with concussions, respectively.

Another explanation for the lower reliability of some measures could be related to our subjects' scores within a truncated range on tests with a restricted range of scores available.²⁵ Having subjects who score within a truncated range has been shown to produce lower reliability correlations.⁴⁵⁻⁴⁷ For example, the SAC is scored between 0 and 30 points. In our specific population, the range of scores for the SAC was 20 to 30 on the initial test and 23 to 30 on the retest. Our subjects represented a more homogeneous group for this test and presented less variability and, thus, a lower test-retest reliability coefficient.⁴⁶ On the other hand, assessments without a limited scoring range, such as the Trails B, usually provide higher test-retest reliabilities.⁴⁵⁻⁴⁷

Although several groups have investigated the test-retest reliability of various cognitive assessments in athletes,¹⁷⁻²¹ the literature is somewhat lacking with respect to the test-retest reliability of postural stability assessments in athletes. Previous authors have found that eyes-closed balancing tasks are novel for most children; therefore, the possibility of a learning effect exists, which can influence the retest scores.⁴⁸ Such a finding has been noted with the BESS: a learning effect has been reported upon serial testing.^{26,40,41} We found acceptable ($r = -.70$) reliability of BESS performance in our subjects, but the significant improvement during the retest session likely affected our reliability. Additionally, increased variability in children with regard to measures of balance is well documented until children reach adult-like postural stability, near the age of 11 years.³⁹ In 2 studies testing younger children than we used in our investigation, the test-retest reliabilities on an eyes-closed, single-leg stance ranged from $.59$ to $.77$, and on a tiltboard test of balance, for both eyes-open and eyes-closed conditions, reliabilities were low ($r = .45$).⁴⁸ Similarly, Westcott et al,⁴⁹ using the Pediatric Clinical Test of Sensory Interaction and Balance, found reliability coefficients for combined sensory condition scores ranging from $.45$ to $.69$ with the feet together and from $.44$ to $.83$ during a heel-to-toe (tandem) stance in children between 4 and 9 years of age. It is important to note that no authors have compared the BESS to the aforementioned balance tests; therefore, one explanation for the different reliabilities could be that the tests assess balance differently. Additionally, our subjects were older than those reported in the above studies; it is plausible that they exhibited better balance ability and, therefore, improved reliability.

Reliable Change Indices

Although use of the RCI has become more popular in determining change in cognitive function after concussion,¹⁷⁻¹⁹ no method has yet been accepted for determining how many "points" on a particular assessment indicates a cognitive or balance deficit. Some authors^{17,23,50} have identified the change score as clinically meaningful if it lies outside of the 90% CI. However, the use of the 90% CI might be too conservative for sport-related concussion, because the impairments are often subtle and resolve rather quickly.²³ Yet this recommendation was based on the results of a study of high school and collegiate athletes. Recently, it has been suggested that sports medicine clinicians should be more conservative when making decisions regarding a young athlete who has sustained a

concussion¹⁵; therefore, our 70% CI scores may not be too conservative to use with our younger population. It should also be noted that the RCI method we used included a correction for practice effects, as suggested by Chelune et al.³⁷

Clinicians should not use a single RCI as the sole determinant in return-to-play decision making after a concussion. The RCI values are intended to help the clinician decide what constitutes a meaningful change in an athlete's score and should be interpreted along with the individual's clinical examination, concussion history, presenting symptoms, and other assessment data. Additionally, the limitations of using the RCI as a means of determining change include the need to understand the statistical procedure and alternate methods of detecting change (ie, standard regression) that may provide better results.

Relationship Between the Standard Assessment of Concussion and Neuropsychological Tests

We found that the SAC was significantly correlated with 4 of 6 neuropsychological measures; however, these correlations demonstrated a weak relationship, accounting for only 8.2% to 13.3% of the variance in SAC scores. Two possible reasons for the weak relationships are the restricted range of scores and the domains tested. The SAC is known to have a restricted range of scores in a normative sample,²⁵ as was the case in our study. This factor may help explain why we found only weak relationships between it and the neuropsychological tests.

Another possible explanation for the weak relationships is that these measures likely assess somewhat different domains of cognitive function. The gap here may lie between the measurement of global cognitive functioning and specific cognitive abilities. The correlations between total SAC score (ie, as an indicator of overall cognitive functioning) and scores on measures of specific cognitive functions (eg, memory) were weak. The SAC is advocated as a sideline mental status assessment tool useful in the first 48 hours after injury and is a valid and moderately reliable gross measure of cognitive functioning in a mixed sample of high school and collegiate athletes during this acute phase of injury.^{6,9,23,51} In contrast, more advanced neuropsychological testing is often useful in detecting subtle cognitive abnormalities in specific cognitive domains, in determining prolonged deficits in cognitive function, or in aiding in return-to-play decision making once the athlete is asymptomatic.^{7,15,44,52-54} Cognitive areas, including verbal learning and memory (Buschke SRT), attention and concentration (Trails B, Coding, Symbol Search), visual-motor function (Trails B, Coding, Symbol Search), sequencing (Trails B), and processing speed (Trails B, Coding, Symbol Search), were assessed with the neuropsychological battery.³⁸ Although the SAC assesses the domains of memory and concentration, the brevity of those sections as well as the entire SAC may prevent an extensive assessment of those cognitive areas further out from the acute injury. These results support the notion that both the SAC and more complex neuropsychological assessments should be used in the evaluation of an athlete after concussion.

In an attempt to better isolate various cognitive domains, we performed a post hoc analysis of the relationship between each of the 4 SAC domains and the 6 neuropsychological measures. No meaningful relationships were revealed; thus, it is

likely that these assessments do serve different purposes in a concussion assessment protocol.

Also of clinical importance are the high correlations found among some of the neuropsychological assessments, specifically the Trails B, which was highly correlated with 2 of the other neuropsychological measures. Those clinicians using a neuropsychological test battery should select tests that will assess the various cognitive domains typically affected by concussion and that can be administered in a short amount of time. Whereas a standard neuropsychological examination consists of multiple cognitive domains and requires 3 to 6 hours to administer, baseline and follow-up examinations for sport-related concussion typically last 20 to 30 minutes and target neurocognitive functioning most sensitive to impairments after a concussive injury.⁵⁵ We feel that the tests used in our study could be added to a battery for prospective injury investigations. These tests can be administered in 20 minutes and assess the domains typically affected by concussion: memory (Buschke SRT), attention (Coding, Symbol Search, Trails B), and speed and flexibility of cognitive processing (Coding, Trails B).⁵⁵ The inclusion of additional tests related to reaction time, visual memory, and complex attention would strengthen the battery and provide a more global measure of cognitive function. Future authors should address the psychometric properties of additional assessments.

Although we determined that we cannot predict SAC scores very well based on the neuropsychological test scores, we found that improved performance on the neuropsychological tests correlated with improved performance on the SAC. Our results indicate that psychometrically the SAC behaves similarly in young athletes, adolescents, and adults, in whom the instrument has demonstrated reliability, validity, clinical sensitivity, and specificity in detecting neurocognitive impairment after concussion. However, further research in injured subjects is required to determine the sensitivity and specificity in detecting cognitive dysfunction during the acute period after concussion in younger athletes.

Limitations

We acknowledge that certain limitations exist in this investigation. We identified a neuropsychological test battery that assessed domains similar to those reported in studies of high school and collegiate populations; however, some neuropsychological domains (verbal memory, reaction time, complex attention) were not included in our battery. In addition, our study included a small sample size and contained more females in the older age group.

Clinical Relevance

In conclusion, healthy young athletes performed comparably with older athletes and adults with respect to test-retest reliability, with the exception of the lower reliability we found on the SAC. These findings begin to establish a group of cognitive tests appropriate for use in young sports participants and "set the table" for clinical studies to evaluate the validity of these measures in a younger population. Future authors should continue to refine the test battery with additional neuropsychological assessments and begin to collect preinjury and post-injury data. The weak relationship found between the SAC and the selected neuropsychological test measures provides further evidence that both the SAC and traditional neuropsychological

chological tests should be used in a concussion assessment battery when evaluating young athletes. Clinicians should also understand that the results from these assessment tools should be used in conjunction with the clinical examination and the athlete's concussion history when making return-to-play decisions. It is also important to recognize that neuropsychological assessments are designed by and for the trained neuropsychologist and should only be interpreted by qualified personnel.

ACKNOWLEDGMENTS

We thank Robert Diamond, PhD, and the Division of Neuropsychology, Department of Psychiatric Medicine at the University of Virginia for their assistance with this study. We thank Curt Bay, PhD, for his assistance with the manuscript revisions. Sections of this manuscript were presented in abstract form at the Fourth Annual Neurotrauma and Sports Medicine Review in Phoenix, AZ (February 2004), and at the Rocky Mountain Athletic Trainers' Association District Meeting in Broomfield, CO (March 2004).

REFERENCES

- Collins MW, Lovell MR, Iverson GL, Cantu RC, Maroon JC, Field M. Cumulative effects of concussion in high school athletes. *Neurosurgery*. 2002;51:1175–1181.
- Guskiewicz KM, McCrea M, Marshall SW, et al. Cumulative effects associated with recurrent concussion in collegiate football players: the NCAA Concussion Study. *JAMA*. 2003;290:2549–2555.
- Macciocchi SN, Barth JT, Littlefield L, Cantu RC. Multiple concussions and neuropsychological functioning in collegiate football players. *J Athl Train*. 2001;36:303–306.
- Field M, Collins MW, Lovell MR, Maroon J. Does age play a role in recovery from sports-related concussion? A comparison of high school and collegiate athletes. *J Pediatr*. 2003;142:546–553.
- Lovell MR, Collins MW, Iverson GL, et al. Recovery from mild concussion in high school athletes. *J Neurosurg*. 2003;98:296–301.
- McCrea M, Kelly JP, Randolph C, et al. Standardized assessment of concussion (SAC): on-site mental status evaluation of the athlete. *J Head Trauma Rehabil*. 1998;13:27–35.
- Guskiewicz KM, Ross SE, Marshall SW. Postural stability and neuropsychological deficits after concussion in collegiate athletes. *J Athl Train*. 2001;36:263–273.
- Barth JT, Macciocchi SN, Giordani B, Rimel R, Jane JA, Boll TJ. Neuropsychological sequelae of minor head injury. *Neurosurgery*. 1983;13:529–533.
- McCrea M, Guskiewicz KM, Marshall SW, et al. Acute effects and recovery time following concussion in collegiate football players: the NCAA Concussion Study. *JAMA*. 2003;290:2556–2563.
- Macciocchi SN, Barth JT, Alves W, Rimel RW, Jane JA. Neuropsychological functioning and recovery after mild head injury in collegiate athletes. *Neurosurgery*. 1996;39:510–514.
- Guskiewicz KM, Perrin DH, Gansneder BM. Effect of mild head injury on postural stability in athletes. *J Athl Train*. 1996;31:300–306.
- Guskiewicz KM, Riemann BL, Perrin DH, Nashner LM. Alternative approaches to the assessment of mild head injury in athletes. *Med Sci Sports Exerc*. 1997;29(suppl 7):S213–S221.
- Peterson CL, Ferrara MS, Mrazik M, Piland S, Elliott R. Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. *Clin J Sport Med*. 2003;13:230–237.
- Riemann BL, Guskiewicz KM. Effects of mild head injury on postural stability as measured through clinical balance testing. *J Athl Train*. 2000;35:19–25.
- Guskiewicz KM, Bruce SL, Cantu RC, et al. National Athletic Trainers' Association position statement: management of sport-related concussion. *J Athl Train*. 2004;39:280–297.
- Randolph C, McCrea M, Barr WB. Is neuropsychological testing useful in the management of sport-related concussion? *J Athl Train*. 2005;40:139–152.
- Hinton-Bayre AD, Geffen GM, Geffen LB, McFarland KA, Friis P. Concussion in contact sports: reliable change indices of impairment and recovery. *J Clin Exp Neuropsychol*. 1999;21:70–86.
- Barr WB. Neuropsychological testing of high school athletes: preliminary norms and test-retest indices. *Arch Clin Neuropsychol*. 2003;18:91–101.
- Iverson GL, Lovell MR, Collins MW. Interpreting change in ImPACT following sport concussion. *Clin Neuropsychol*. 2003;17:460–467.
- Erlanger D, Feldman D, Kutner K, et al. Development and validation of a Web-based neuropsychological test protocol for sports-related return-to-play decision-making. *Arch Clin Neuropsychol*. 2003;18:293–316.
- Collie A, Maruff P, Makdissi M, McCrory PR, McStephen M, Darby D. CogSport: reliability and correlation with conventional cognitive tests used in postconcussion medical evaluations. *Clin J Sport Med*. 2003;13:28–32.
- Heaton RK, Temkin N, Dikmen S, et al. Detecting change: a comparison of three neuropsychological measures, using normal and clinical samples. *Arch Clin Neuropsychol*. 2001;16:75–91.
- Barr WB, McCrea M. Sensitivity and specificity of standardized neurocognitive testing immediately following sports concussion. *J Int Neuropsychol Soc*. 2001;7:693–702.
- Erlanger DM, Saliba E, Barth J, Almquist J, Webright W, Freeman J. Monitoring resolution of postconcussion symptoms in athletes: preliminary results of a Web-based neuropsychological test protocol. *J Athl Train*. 2001;36:280–287.
- Dikmen SS, Heaton RK, Grant I, Temkin NR. Test-retest reliability and practice effects of Expanded Halstead-Reitan Neuropsychological Test Battery. *J Int Neuropsychol Soc*. 1999;5:346–356.
- Valovich McLeod TC, Perrin DH, Guskiewicz KM, Diamond R, Gansneder BM, Shultz SJ. Serial administration of clinical concussion assessments and learning effects in healthy youth sports participants. *Clin J Sport Med*. 2004;14:287–295.
- Buschke H, Fuld PA. Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology*. 1974;24:1019–1025.
- Clodfelter CJ, Dickson AL, Newton Wilkes C, Johnson RB. Alternate forms of the selective reminding for children. *Clin Neuropsychol*. 1987;1:243–249.
- Reitan RM, Wolfson D. *The Halstead-Reitan Neuropsychological Test Battery*. Tucson, AZ: Neuropsychology Press; 1985.
- Wechsler D. *Wechsler Intelligence Scale for Children*. 3rd ed. San Antonio, TX: The Psychological Corp, Harcourt Brace Jovanovich Inc; 1991.
- McCrea M, Kelly JP, Randolph C. *The Standardized Assessment of Concussion: Manual for Administration, Scoring, and Interpretation*. 2nd ed. Waukesha, WI: CNS Inc; 2000.
- McCrea M, Kelly JP, Kluge J, Ackley B, Randolph C. Standardized assessment of concussion in football players. *Neurology*. 1997;48:586–588.
- McCrea M, Kelly JP, Randolph C, Cisler R, Berger L. Immediate neurocognitive effects of concussion. *Neurosurgery*. 2002;50:1032–1042.
- Riemann BL, Guskiewicz KM, Shields EW. Relationship between clinical and forceplate measures of postural stability. *J Sport Rehabil*. 1999;8:71–82.
- Portney LG, Watkins MP. *Foundations of Clinical Research: Application to Practice*. Upper Saddle River, NJ: Prentice Hall; 2000.
- Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. 1991;59:12–19.
- Chelune GJ, Naugle RI, Luders H, Sedlack J, Awad IA. Individual change after epilepsy surgery: practice effects and base-rate information. *Neuropsychology*. 1993;7:41–52.
- Spreen O, Strauss E. *A Compendium of Neuropsychological Tests*. New York, NY: Oxford University Press; 1998.
- Riach CL, Hayes KC. Maturation of postural sway in young children. *Dev Med Child Neurol*. 1987;29:650–658.
- Mancuso JJ, Guskiewicz KM, Onate JA, Ross SE. An investigation of the learning effect for the Balance Error Scoring System and its clinical implications [abstract]. *J Athl Train*. 2002;37(suppl):S-10.

41. Valovich TC, Perrin DH, Gansneder BM. Repeat administration elicits a practice effect with the Balance Error Scoring System but not with the Standardized Assessment of Concussion in high school athletes. *J Athl Train.* 2003;38:51–56.
42. Duff P, Westervelt HJ, McCaffrey RJ, Haase RF. Practice effects, test-retest stability, and dual baseline assessments with the California Verbal Learning Test in an HIV sample. *Arch Clin Neuropsychol.* 2001;16:461–476.
43. Hinton-Bayre AD, Geffen GM, McFarland KA. Mild head injury and speed of information processing: a prospective study of professional rugby league players. *J Clin Exp Neuropsychol.* 1997;19:275–289.
44. Guskiewicz KM, Marshall SW, Broglio SP, Cantu RC, Kirkendall DT. No evidence of impaired neurocognitive performance in collegiate soccer players. *Am J Sports Med.* 2002;30:157–162.
45. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. *Psychol Assess.* 1995;7:309–319.
46. Garrett HE. *Statistics in Psychology and Education.* New York, NY: David McKay Co; 1962.
47. Nunnally JC. *Psychometric Theory.* New York, NY: McGraw-Hill Book Co; 1978.
48. Atwater SW, Crowe TK, Deitz JC, Richardson PK. Interrater and test-retest reliability of two pediatric balance tests. *Phys Ther.* 1991;70:79–87.
49. Westcott SL, Crowe TK, Deitz JC, Richardson P. Test-retest reliability of the Pediatric Clinical Test of Sensory Interaction for Balance. *Phys Occup Ther Pediatr.* 1994;14:1–22.
50. McSweeney AJ, Naugle RI, Chelune GJ, Luders H. “T scores for change”: an illustration of a regression approach to depicting change in clinical neuropsychology. *Clin Neuropsychol.* 1993;7:300–312.
51. McCrea M. Standardized mental status testing on the sideline after sport-related concussion. *J Athl Train.* 2001;36:274–279.
52. Bleiberg J, Halpern EL, Reeves D, Daniel JC. Future directions for the neuropsychological assessment of sports concussion. *J Head Trauma Rehabil.* 1998;13:36–44.
53. Collins MW, Grindel SH, Lovell MR, et al. Relationship between concussion and neuropsychological performance in college football players. *JAMA.* 1999;282:964–970.
54. McCrory P, Johnston K, Meeuwisse W, et al. Summary and agreement statement of the 2nd International Conference on Concussion in Sport, Prague 2004. *Clin J Sport Med.* 2005;15:48–55.
55. Randolph C. Implementation of neuropsychological testing models for the high school, collegiate, and professional sport settings. *J Athl Train.* 2001;36:288–296.