

Cross-Validation of Measures Used for Computer-Based Assessment of Concussion

Philip Schatz and Brendan O. Putz

Saint Joseph's University, Philadelphia, Pennsylvania, USA

The purpose of this study was to cross-validate subscales of computerized test batteries developed for the assessment and management of sports-related concussion, as well as to validate these subscales against select long-standing paper-based measures. In order to examine concurrent validity, we administered 3 such computerized measures, CogSport, ImPACT, and HeadMinder, along with more established paper-based measures, the Trail Making Tests and the Digit Symbol subtest of the Wechsler scale. Analysis of processing speed indices revealed significant but modest correlations between Trails B and the computer-based measures, ranging from $-.51$ (ImPACT) to $.60$ (HeadMinder), and for Digit Symbol ranging from $-.37$ (Headminder) to $.53$ (ImPACT). Analysis of complex reaction time (CRT) data revealed significant but modest correlations between ImPACT and CogSport ($.65$), and between ImPACT and HeadMinder ($.41$), but not between CogSport and HeadMinder. No intercorrelations were found between the memory indices from the 3 programs. Our results suggest that these tests share some common variance on constructs such as processing speed and reaction time, but not within the domain of memory. Clinicians obtaining baseline evaluations using 1 measure should not use the baseline as a basis for post-concussion assessment using another measure.

Key words: concussion testing, computerized neuropsychological assessment

Epidemiology of sports-related concussion has been well documented (Dick, 1999; Echemendia & Julian, 2001; Powell & Barber-Foss, 1999), and the assessment and management of sports-related concussions has received considerable attention in the literature. Since Barth and colleagues (1989) pioneered the use of neuropsychological tests at baseline and post-concussion time intervals in order to document recovery of cognitive deficit in a concussed athlete, researchers and clinicians have devoted considerable efforts towards understanding suitable schedules of post-concussion assessments and recovery trajectories, developing appropriate test batteries, and identifying the effects of concussion in youth, collegiate, and professional athletes (Collie, Maruff, Makdissi, et al., 2003; Collins, et al., 1999; Echemendia, Putukian, Mackin, Julian, & Shoss, 2001; Erlanger et al., 2003; Iverson, Lovell, & Collins, 2002; Lovell & Collins, 1998; Makdissi et al., 2001; Moser & Schatz, 2002).

Increased awareness of the effects of cerebral concussions, as well as the prevalence of mild traumatic

brain injury in organized sports, have combined to create a need for baseline and post-concussion neuropsychological testing programs (McKeever & Schatz, 2003). Baseline assessments of at-risk populations, coupled with on-site evaluations, may help answer the complex, dynamic, and critical questions concerning the athletes' ability to return to play (Echemendia & Cantu, 2003).

Computerized assessment has gained popularity in recent decades, and researchers have pointed to its advantages and disadvantages when compared to assessment using traditional paper-based measures. Standardized paper and pencil (or human-administered) procedures must be analyzed and administered by trained professional staff, often decreasing efficiency of test-taking procedures. In contrast, automated assessments are often administered or analyzed, or both, by computer, thereby decreasing time needed for professional participation and cutting costs while allowing more individuals to participate in examinations (Bartram & Bayliss, 1984; Wilson & McMillan, 1991).

Computerized assessment is not without its disadvantages. Critics have pointed to automated procedures as creating a false sense of ability such that anyone may be able to diagnose perceived neuropsychological defi-

Correspondence should be addressed to Philip Schatz, Saint Joseph's University, Department of Psychology, 222 Post Hall, Philadelphia, PA 19131, USA. E-mail: pschatz@sju.edu

cits with the aid of only a computer. For the administrator, computerized assessment can encourage a passive stance during clinical evaluation, rather than taking an active role in the testing process. Finally, computers are able to present visual stimuli that can be verbally encoded, such as words presented visually for recall or other cognitive comparisons, or even present auditory stimuli such as words presented as sounds. Even where test developers refer to such stimuli as “verbal” or “verbally-based,” or such subtests as “verbal memory,” computer-based measures cannot assess efferent language or verbal modalities. (For a more comprehensive review of computer-based assessment, see Schatz & Browndyke, 2002). The literature is not entirely in agreement with respect to the psychometric equivalence of computerized versus paper-based assessment measures. Bartram and Byliss (1984) noted that computerizing pencil and paper procedures requires not only a new set of norms, but also a correlation of that test against the selection criterion measures. To answer these criticisms, research has focused on whether or not there are significant differences between the two administrations, and various researchers have found equivalence between traditional and computer-based administrations (Epstein, Klinkenberg, Wiley, & McKinley, 2001; French & Beaumont, 1987; Vispoel, 2000).

While computer programs are able to analyze data, they are not capable of interpreting the clinical meaningfulness of those data. There is clear agreement among researchers and clinicians involved in the assessment and management of sports-related concussion that while physicians generally hold the ultimate responsibility for making return-to-play decisions, a clinical neuropsychologist should be involved in either baseline or post-concussive assessments as part of a larger concussion management program (Aubry, et al., 2002; Echemendia & Cantu, 2003).

In recent years, computer-based measures have been developed for the purpose of assessing sports-related concussions and providing data that could contribute to the management of return-to-play decisions. Concussion assessment and management programs have become commonplace, many incorporating computer-based software to assess athletes at the high school, collegiate, and professional levels. Given the prevalence of, and attention to, sports-related concussion, it is likely that athletes may soon transfer from one school or level to another having already completed baseline or post-concussion evaluations, or both, using one of the existing computer-based assessment measures. At present, there is no single universally adopted measure being used at the high school, collegiate, or profes-

sional level. Also, while each of the computerized batteries has been validated against at least one standardized paper-based measure, they have never been cross-validated. We sought to determine whether these computerized batteries are assessing similar or different characteristics, through cross-validation, as well as comparisons with select long-standing neuropsychological test measures of processing speed. Given that these computer-based tests are measures of processing speed, reaction time, and working memory, we expected that similar constructs measured by each test battery (i.e., memory, or processing speed) would be highly correlated.

METHOD

Participants

Participants were 30 individuals from Saint Joseph's University who volunteered in partial fulfillment of a course requirement. There were 14 men (47%) and 16 women (53%) ages 18–23, with an average age of 21 years. Participants were non-athletes who reported they had not experienced a concussion in the 6 months prior to testing.

Measures

Each participant completed each of the following test measures: ImPACT, HeadMinder's Concussion Resolution Index, and CogSport, as well as paper- and computer-based versions of the Trail Making Test A & B (TMT), and the Digit Symbol subtest of the Wechsler Adult Intelligence Scale–Revised (WAIS–R; these measures are described in detail next).

ImPACT (Lovell, Collins, Podell, Powell, & Maroon, 2000) is a computerized neuropsychological screening battery consisting of seven individual cognitive test modules, which form composite scores in the areas of memory, reaction time, and processing speed. Each of the composite scores is computed by standardized formulas derived from the results of the seven individual cognitive tasks. The verbal memory composite score represents the average percent correct for a word recognition task, a symbol–number match task, and a letter memory task with an accompanying interference task. The visual memory composite score is comprised of the average percent correct scores for an abstract figure recognition task, and memory of a series of illuminated Xs and Os after an intervening sequential number clicking task. The reaction time composite score represents the average response time (in milliseconds) on a

choice reaction time task, a go or no-go task, and the symbol–number match task. The processing speed composite score consists of the weighted average of three interference tasks during the visual and verbal memory paradigms. ImPACT also includes a post-concussion symptom scale comprised of the total score for 21 common symptoms of concussion. When completing this scale, participants rate the presence and severity of symptoms commonly associated with concussion, such as headache, sensitivity to light, or feeling slow. Iverson, Lovell, and Collins (2002) examined several validity measurements of ImPACT in a sample of 120 high school and college athletes. Concurrent validity was examined by examining the composite scores and their sensitivity to the acute effects of concussion. Concussed athletes reported significantly more symptoms, and performed worse on Memory and Reaction Time indices (Iverson et al., 2002). Performance on the Symbol Digit Modalities test (SDMT) has been shown to correlate significantly with ImPACT Processing Speed ($r = .70$) and Reaction Time ($r = -.60$) indices (Iverson, Lovell, & Collins, 2005). Post-concussive symptoms have been significantly related to decreased performance on ImPACT Reaction Time, Verbal Memory, and Processing Speed indices (Iverson, Gaetz, Lovell, & Collins, 2004), suggesting that ImPACT is sensitive to the acute effects of concussion. Divergent validity was examined through an intercorrelation matrix of composite scores at preseason and post-concussion. The nonsignificant correlations found between different test components (at preseason baseline testing) indicate they do not have much shared variance, and therefore appear to be measuring different constructs.

HeadMinder's Concussion Resolution index (CRI; Erlanger, Feldman, & Kutner, 1999) assesses cognitive change to provide objective measurement of cognitive functions through online tests of attention and reaction time. The tests are constructed with statistical techniques to address many concerns regarding the feasibility of cognitive testing including test–retest effects, alternate forms, and ease of administration in order to administer reliable and valid examinations of cognitive ability. The CRI includes subtests designed to cumulatively assess cognitive functions typically associated with sports-related concussion in the following areas: reaction time, visual recognition, and speed of information processing, as well as a post-concussion symptom questionnaire. Reaction time measures responses to simple stimuli (i.e., speed of response to a geometric shape as well as a test of speed of response to a geometric shape following a cue) and choice reaction time (i.e.,

speed of decision making). The standardized mode of administration and testing follows the baseline and post-injury model of assessment. There is an automatic statistical adjustment to account for test–retest reliability and practice effects, which is unique to HeadMinder. Erlanger and colleagues (2001) validated HeadMinder software using several standardized neuropsychological pencil and paper measures, including the Digit Symbol subtest of the WAIS–III, TMT, and correlations between components of the CRI. In general, they found that correlations between CRI and neuropsychological measures were highest for tests that measured constructs similar to those measured by the CRI (Erlanger, et al., 2001, 2002, 2003). CRI Response Speed scores were found to correlate with Trails A (.73) and B (.74; (Erlanger, et al., 2002), and Processing Speed scores were shown to correlate with WAIS–III Digit Span (.53) and the SDMT (.66). In their 2003 study, CRI Simple Reaction Time (SRT) scores were shown to correlate with WAIS–III Digit Symbol (.45) and Trails A (.56), and Processing Speed scores were found to correlate with Trails B (.37) and SDMT (.66). Because psychomotor speed and processing speed are cognitive factors that are sensitive to the effects of mild concussion, it shows the sensitivity of the CRI to the effects of mild sports-related head trauma (Erlanger et al., 2003).

CogSport (CogState, 1999) is an assessment tool primarily concerned with assisting with return to play decisions. CogSport uses playing cards to create a game-like atmosphere and help motivate the athlete to perform the test. CogSport measures post-concussion results against baseline results in four categories to interpret level of impairment of the individual. Psychomotor, decision-making, problem-solving, and memory scales are related in terms of response speed, and ultimately change in response speed. In order to determine the reliability and validity of CogSport, a sample of 240 professional athletes in the Australian Football League (AFL), and 60 healthy youth volunteers, completed assessments prior to the AFL athletic season (CogState, 1999). Reliability was determined by calculating intra-class correlation (ICC) coefficients on serial data collected in the 60 youth volunteers at intervals of 1 hour and 1 week. Construct validity was determined by calculating ICC coefficients between CogSport outcome variables and performance on the Digit Symbol Substitution test (DSST) and the TMT in the 240 athletes tested at pre-season (Collie, Darbie, & Maruff, 2001). CogSport speed indices displayed high to very high (.69–.90) reliability. CogSport speed indices displayed significant, moderate to high correlations

with the DSST and the TMT (.23–.86). This data suggests that measures of response speed will be reliable indicators of cognitive change following concussion and CogSport tests measure the same cognitive functions as the DSST and TMT, two tests that are used commonly to monitor recovery from concussion.

The Digit Symbol subscale of the WAIS–R battery has been used to test attention primarily on those afflicted with mild or traumatic brain injury. This test falls in the Performance section of the test, and within a further subdivision that is rated to measure attention, monitoring, and evaluating task performance. This subdivision is known as the third factor, and designated the memory or freedom from distractibility section (Spreen & Strauss, 1991). The Digit Symbol subtest has been shown to have moderate correlations (.29–.60) with the Mini Mental Status Examination (Hopp et al., 1997), and moderate correlations (.43–.60) with the other WAIS subscales (Wechsler, 1955). Reliability coefficients range from .82 (Wechsler, 1981) to .92 (Wechsler, 1955). Although more recent versions of the Digit Symbol test were available (i.e., from the WAIS–III and –IV), we had previously computerized the version from the WAIS-R for a study validating (Schatz & Putz, 2002) the original Virginia football study (Barth, et al., 1989), and were also collecting ongoing validation data for these purposes (see below).

The TMT was designed to test for visual search, attention and motor function in correlation with speed. There are two forms, Part A and Part B, which are scored and completed under the same principal, however separately and using different stimuli. Part A uses only numbers, and Part B uses both letters and numbers, thereby making it possible for different reliability and validity scores to be found between parts. Lezak (1983) found significant practice effects on only Part A, and not Part B. In the same study, reliability was reported as .98 for Part A, and .67 for Part B (coefficient of concordance). Snow and colleagues (Snow, Tierney, Zorzitto, Fisher, & Reid, 1988) found a 1-year test–retest reliability of .64 for Part A, and .72 for Part B in respondents whose average age was 67. Goldstein and Watson (1989) found similar reliability coefficients (.69–.94 for Part A, .66–.86 for Part B) for different types of neurological groups, but not for schizophrenics (.36 for Part A, .63 for Part B). Charter and colleagues (Charter, Adkins, Alekoumbides, & Seacat, 1987) found alternate form reliability, by leaving the form intact and changing the letters and numbers, to be .89 for Part A, and .92 for Part B. Trails B is relatively free from education effects, with correlation coefficients between ed-

ucation and Trails A and B reported as .19 and .33 (Bornstein, 1985).

Computer-generated versions of Trails A and B, the Digit Symbol test of the WAIS–R, and the d2 Test of Attention (Brickenkamp, 1981) were developed, modeled directly after the pencil and paper versions, and administered to all participants. However, as these were novel instruments with no previous psychometric comparison to the paper-based versions, the results from the computerized measures were not included in the analyses.

Procedure

This research was approved by the Saint Joseph's University Institutional Review Board. Participants signed up for three consecutive, independent testing sessions, which occurred at 48-hour intervals on Monday, Wednesday, and Friday. One experimenter directed each participant through approximately 40 min of testing each session, which took place in private test sessions, independent of other participants. The measures were assigned into three separate groupings:

Grouping A: ImPACT, d2 Test of Attention (computerized).

Grouping B: CRI, Trails A and B, and Digit Symbol (pencil and paper).

Grouping C: CogSport, Trails A and B, and Digit Symbol (computerized).

Assignment of participants to testing sessions was counter-balanced, with 5 participants completing each of the following orders of groupings: ABC, ACB, BAC, BCA, CAB, and CBA.

Analyses

SRT and Memory Indices of CogSport and HeadMinder were examined using Pearson's correlation coefficients. CRT components of ImPACT, CogSport and HeadMinder were compared with Pearson's correlation coefficients. Visual Memory components of ImPACT and CogSport were compared with Pearson's correlation coefficients. Finally, Processing Speed correlation matrix was formulated using indices from ImPACT, HeadMinder, Trails A & B, and Digit Symbol.

RESULTS

Simple Reaction Time (SRT)

Pearson’s correlation coefficients revealed no significant correlation between CogSport and HeadMinder SRT test indices, $r(30) = .29, p = .12$. HeadMinder SRT scores were significantly correlated with Trails A, $r(30) = .43, p = .018$, and Digit Symbol, $r(30) = .53, p = .003$, but not Trails B, $r(30) = .23, p = .22$. CogSport SRT scores were not correlated with any of these measures (see Table 1).

Complex Reaction Time (CRT)

Significant correlations were found between CRT indices on ImPACT and CogSport, $r(30) = .66, p = .001$, and ImPACT and HeadMinder, $r(30) = .41, p = .026$. However, there was no significant correlation found between CRT indices on CogSport and HeadMinder, $r(30) = .33, p = .07$ (see Table 2).

ImPACT reaction time composite scores were significantly correlated with Trails A, $r(30) = .64, p = .001$, Trails B, $r(30) = .44, p = .014$, and Digit Symbol, $r(30) = .46, p = .012$. HeadMinder CRT scores showed no significant correlations with these measures. CogSport complex time scores were significantly correlated with Trails A, $r(30) = .54, p = .002$, Trails B, $r(30) = .54, p = .002$, but not Digit Symbol, $r(30) = .28, p = .13$ (see Table 2).

Memory

Pearson’s correlation coefficients revealed significant correlations between the two different CogSport Memory indices, the One-back Working Memory Test and the Learning Memory Task, $r(30) = .72, p = .001$. However, no significant correlations were found between the ImPACT visual and verbal memory indices, or the ImPACT and CogSport indices (see Table 3).

Processing Speed

The processing speed indices yielded the highest and most consistent correlations among the indices. Trails B showed significant correlations with the HeadMinder Processing Speed raw score, $r(30) = .60, p = .001$, and the ImPACT Processing Speed index, $r(30) = -.51, p = .004$. Digit Symbol showed significant correlations with HeadMinder’s, $r(30) = -.37, p = .042$, ImPACT Processing Speed index, $r(30) = .53, p = .003$, as well as with Trails B, $r(30) = -.51, p = .004$ (note, higher scores on ImPACT Processing Speed index reflect better performance, resulting in negative correlations; see Table 4).

DISCUSSION

These results suggest that there is considerable variance with respect to the cross-validation of perfor-

Table 1. Simple Reaction Time Intercorrelations

	COG_SRT	TRA	TRB	DIGSYM
HD_SRT	.292 (.118)	.428* (.018)	.232 (.217)	-.526** (.003)
COG_SRT	—	.277 (.138)	.172 (.362)	-.076 (.689)
TRA		—	.613** (.001)	-.373* (.042)
TRB			—	-.381* (.038)

Note. HD = HeadMinder’s; COG = CogSport; SRT = Simple RT; TRA = Trails A; TRB = Trails B; DIGSYM = Digit Symbol subtest of the WAIS-R. * $p < .05$, ** $p < .01$. Significance levels in parenthesis.

Table 2. Complex Reaction Time Intercorrelations

	HD_CRT	COG_CRT	TRA	TRB	DIGSYM
IMPACT	.407* (.026)	.649** (.001)	.641** (.001)	.442* (.014)	-.455* (.012)
HD_CRT	—	.333 (.072)	.055 (.772)	.315 (.091)	-.303 (.104)
COG_CRT		—	.544** (.002)	.535** (.002)	-.281 (.132)
TRA			—	.613** (.001)	-.373* (.042)
TRB				—	-.381* (.038)

Note. IMPACT = Choice RT; HD = HeadMinder’s; COG = CogSport; CRT = Choice RT; TRA = Trails A; TRB = Trails B; DIGSYM = Digit Symbol subtest of the WAIS-R. * $p < .05$, ** $p < .01$. Significance levels in parenthesis.

Table 3. *Memory Index Intercorrelations*

	ImPACT Visual	CogSport Visual	CogSport Learning
ImPACT Verbal	.340	-.160	-.102
ImPACT Visual	—	-.012	-.058
CogSport Visual	—	—	.723*

Note. * $p < .05$.

Table 4. *Processing Speed Index Intercorrelations*

	Digit Symbol	ImPACT	HeadMinder
Trails B	-.381*	-.506**	.601***
Digit Symbol	—	.542**	-.610***
ImPACT	—	—	-.373*

Note. * $p < .05$, ** $p < .004$, *** $p < .001$.

mance by normal controls on tests designed to assess baseline and post-concussion performance. The four indices compared were delineated to commonly diagnosed areas of deficiencies resulting from concussion (Hinton-Bayre, Geffen, Geffen, McFarland, & Friis, 1999), including SRT, CRT, memory, and processing speed. While the computerized batteries used different methods to measure attention and concentration, each yielded raw data scores to some or all of the four indices.

The SRT indices from CogSport and HeadMinder produced low inter correlations. While the HeadMinder SRT was significantly correlated with both Trails A and the Digit Symbol test, the CogSport SRT index was not correlated with these measures. When looking at the components that make up these indices, the CogSport SRT is based on the most “traditional” SRT; pushing a button at the appearance of a single stimulus (i.e., when a card turned over, regardless of whether it was a numbered or face card). The HeadMinder CRI SRT index is comprised of pushing a button at the appearance of a white circle, but not when one of 5 other geometric shapes appeared. The HeadMinder SRT index appears to include cognitive processes shared by both the Trails A and Digit Symbol, beyond that of SRT (i.e., visual scanning, sustained attention), which may explain the intercorrelation with these measures for HeadMinder and not CogSport (which required the user to fixate on a single point looking for a global change in the stimulus).

When analyzing CRT indices, CogSport and HeadMinder indices were not intercorrelated, while the ImPACT index (which is a measure of CRT) did correlate with both the CogSport and HeadMinder indices. Each of the tests used a multi-step visual processing

technique to add complexity to simple reaction time variables. However, they employ different means of measuring CRT, with different types of stimuli and response requirements. ImPACT’s multiple tasks include several different examination techniques, which may explain its high correlation with both HeadMinder and CogSport; the Color Match test for ImPACT closely resembles the CogSport card recognition task, and the picture recognition task of HeadMinder. However, CogSport and HeadMinder tasks did not closely resemble one another. In addition, ImPACT’s CRT index showed significant correlations with all three external measures (TMT A & B, Digit Symbol), and CogSport CRT index shared significant and similar correlations with both TMT A & B (but not Digit Symbol). Differences in how these tests measure CRT, and on which indices specific sub-tasks load may help explain these differences. As ImPACT Symbol-Match test is similar to the Digit Symbol, shared correlation is likely. However, HeadMinder CRT, which includes its picture recognition with both the ImPACT Symbol-Match and the Digit Symbol, did not share the same correlations with Digit Symbol. Upon closer inspection, the HeadMinder CRT index is comprised of a type of One-Back CRT task in which users are required to push a button when a white circle appears after a black square. While this measure does share similar cognitive requirements to the ImPACT Color Match test, the HeadMinder Picture Recognition task (which shares similarities with the Digit Symbol), loads on the processing speed index, which may help explain the lack of shared correlations of HeadMinder with ImPACT and CogSport.

Correlations showed CogSport to have strong internal validity between its two memory tasks, One-Back Working Memory and the Learning Task. These tasks required participants to remember previously presented stimuli in order. Although difficulty between the two tasks varied, their shared stimuli and similar construct may explain the strong correlation between these tasks. However, the tasks that compose the Memory index for CogSport and ImPACT varied considerably with regard to the type of stimuli used. CogSport uses a visual memory paradigm with universal playing cards, while

ImPACT uses words, letters, and common symbol or number pairs. The simple visual memory tasks used by CogSport just do not correlate with the multi-step verbal-oriented ImPACT tasks. ImPACT does have a Visual Memory index, however the developers state that the measurement is currently undergoing field-testing and does not recommend that clinical decisions be made based upon this composite score. To this end, the ImPACT Visual Memory index was not correlated with either of the CogSport Memory tasks.

The Processing Speed indices or measures, including scores from ImPACT, HeadMinder, TMT B, and Digit Symbol, correlated consistently. Both ImPACT and TMT B version correlated significantly with all of the Processing Speed index tests, and the Digit Symbol test correlated significantly with ImPACT and HeadMinder. The overall medium sized correlation between each of the processing speed subtest scores suggests they are sharing some common variance and, to some degree, measuring similar constructs.

There is no immediate predecessor to this study, and therefore the novelty of the data collected should be considered pertinent to further understanding of how computerized assessment may be compared or incorporated into future assessment techniques. One aim of this study was to determine whether different assessment measures are measuring the same constructs. Although some indices shared similar theoretical applications and correlated scores, results show that CogSport, ImPACT, and HeadMinder cannot be integrated with one another for purpose of precise clinical measurement.

This study is not without its limitations. While we were able to obtain an adequate sample of 30 volunteers, a larger sample would have been preferable and may have yielded more generalizable results. However, that we were able to obtain significant findings with this small sample size speaks to the strength of those relationships observed. This study primarily focused on cross-validation of computer-based assessment measures and external validation with long-standing measures of processing speed. The exclusion of external criterion measures within the domain of verbal or visual memory was, in hindsight, unfortunate, and would have strengthened the utility of the results. Future investigations validating concussion measures may benefit from a larger population size and would more likely yield stronger correlations between similarly constructed test items. Such investigations should be conducted with a broader array of external criterion measures, with particular emphasis on measures in the domains of verbal and visual memory, as well as verbal processing.

Each of the computerized test batteries we examined were developed for comparisons between baseline and post-concussion performance, and our sample was comprised of non-clinical non-concussed controls. It would be difficult to justify exposing recently concussed athletes to a long battery of different computer-based measures with similar overlapping constructs for the purpose of research. Yet, it would be helpful to collectively administer these measures pre- and post-concussion, in order to analyze change-from-baseline performance, as well as cross-validate these measures in the context of their ability to distinguish concussed athletes from non-concussed controls, and sensitivity to the effects of concussion within the same sample. Finally, the participants in this study completed computer-based assessments in private test sessions. Group testing of teams of athletes has become the norm, especially when conducting computer-based baseline assessments, and private assessments are often reserved for those athletes who are post-concussion.

In summary, our results showed shared correlations between all the computer-based tests in the domain of processing speed, and between select tests in the domains of simple and choice reaction time. Little shared variance was seen in the domain of memory, although external criterion measures were lacking in this area. Of the test measures used, ImPACT shared the most consistent correlations with the other two computer-based measures, as well as with all external criterion, save for internal correlations in the domain of memory. However, given the limited sample size, lack of clinical population, as well as the other limitations of the study, it would be premature to claim that one computer-based measure is superior to another, simply on the basis of intercorrelations and correlations with a subset of external measures. To be sure, each of these measures has strengths and weaknesses, many of which are documented in the literature, and each appear to measure aspects of attention, reaction time, processing speed, and working memory.

This research area remains open, and provides unique opportunities for future research designs. There is a paucity of external, independent validation of computer-based measures, and this may serve as a call to action for future research projects. Such projects should look at the cross-validation of these batteries in a larger normative sample, include wider representation of external criterion measures across more behavioral domains, investigate the test-retest reliability, clinical validity, and specificity and sensitivity of these computerized measures in detecting and tracking sport-related concussion, and include qualitative data in

the form of both subjectively reported symptoms as well as objective measures of personality and emotional adjustment. Such research may strengthen our current understanding of computer-based assessment, and can direct the future development of computer-based assessment of sports-related concussion. From the present research, we recommend that test developers consider the following factors when creating or revising measures for the assessment of sports-related concussion: ability of the measure to discriminate between concussed and non-concussed athletes, sensitivity of test data with respect to return-to-play decisions, external validation of timing accuracy, external validation with previously validated neuropsychological test measures, inclusion of verbal modalities within the computer-based measures, requirements for repeat baseline assessments, effects of group versus individual assessment, and the ability to administer sub-scales of test measures, with normative data for these sub-scales.

REFERENCES

- Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Johnston, K., Kelly, J., et al. (2002). Summary and agreement statement of the First International Conference on Concussion in Sport, Vienna 2001: Recommendations for the improvement of safety and health of athletes who may suffer concussive injuries. *British Journal of Sports Medicine*, *36*, 6–10.
- Barth, J. T., Alves, W. M., Ryan, T. V., Macciocchi, S. N., Rimel, R. W., Jane, J. A., et al. (1989). Mild head injury in sports: Neuropsychological sequelae and recovery of function. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), *Mild head injury* (pp. 257–275). New York: Oxford University Press.
- Bartram, D., & Bayliss, R. (1984). Automated testing: Past, present and future. *Journal of Occupational Psychology*, *57*, 221–237.
- Borstein, R. A. (1985). Normative data on selected neurological measures from a nonclinical sample. *Journal of Clinical Psychology*, *41*, 651–659.
- Brickenkamp, R. (1981). *Test d2: Aufmerksamkeits-Belastungs-Test (Handanweisung, 7th ed.)* [Test d2: Concentration-Endurance Test]. Goettingen: Hogrefe.
- Charter, R. A., Adkins, T. G., Alekoumbides, A., & Seacat, G. F. (1987). Reliability of the WAIS, WMS and Reitan Battery: Raw scores and standardization scores corrected for age and education. *International Journal of Clinical Neuropsychology*, *9*, 28–32.
- CogState. (1999). CogSport [Computer software]. Parkville, Victoria, Australia: CogState, Ltd.
- Collie, A., Darby, D., & Maruff, P. (2001). Computerized cognitive assessment of athletes with sports related head injury. *British Journal of Sports Medicine*, *35*, 297–302.
- Collie, A., Maruff, P., Makdissi, M., McCrory, P., McStephen, M., & Darby, D. (2003). CogSport: Reliability and correlation with conventional cognitive tests used in postconcussion medical evaluations. *Clinical Journal of Sports Medicine*, *13*(1), 28–32.
- Collins, M. W., Grindel, S. H., Lovell, M. R., Dede, D. E., Moser, D. J., Phalin, B. R., et al. (1999). Relationship between concussion and neuropsychological performance in college football players. *Journal of the American Medical Association*, *282*(10), 964–970.
- Dick, R. W. (1999). *NCAA Injury Surveillance System: 1984–91*. Overland Park, KS: National Collegiate Athletics Association.
- Echemendia, R. J., & Cantu, R. C. (2003). Return to play following sports-related mild traumatic brain injury: The role for neuropsychology. *Applied Neuropsychology*, *10*(1), 48–55.
- Echemendia, R. J., & Julian, L. J. (2001). Mild traumatic brain injury in sports: Neuropsychology's contribution to a developing field. *Neuropsychology Review*, *11*(2), 69–88.
- Echemendia, R. J., Putukian, M., Mackin, R. S., Julian, L., & Shoss, N. (2001). Neuropsychological test performance prior to and following sports-related mild traumatic brain injury. *Clinical Journal of Sports Medicine*, *11*(1), 23–31.
- Epstein, J., Klinkenberg, D., Wiley, D., & McKinley, L. (2001). Insuring sample equivalence across internet and paper and pencil assessments. *Computers in Human Behavior*, *17*, 339–346.
- Erlanger, D., Feldman, D., & Kutner, K. (1999). *Concussion Resolution Index*. New York: HeadMinder, Inc.
- Erlanger, D., Feldman, D., Kutner, K., Kaushik, T., Kroger, H., Festa, J., et al. (2003). Development and validation of a Web-based neuropsychological test protocol for sports-related return-to-play decision-making. *Archives of Clinical Neuropsychology*, *18*(3), 293–316.
- Erlanger, D., Kaushik, T., Broshek, D., Freeman, J., Feldman, D., & Festa, J. (2002). Development and validation of a Web-based screening tool for monitoring cognitive status. *Journal of Head Trauma Rehabilitation*, *17*(5), 458–476.
- Erlanger, D., Saliba, E., Barth, J., Almquist, J., Webright, W., & Freeman, J. (2001). Monitoring resolution of postconcussion symptoms in athletes: Preliminary results of a Web-based neuropsychological test protocol. *Journal of Athletic Training*, *36*(3), 280–287.
- French, C. C. & Beumont, J. G. (1987). The reaction of psychiatric patients to computerized assessment. *British Journal of Clinical Psychology*, *26*, 267–278.
- Goldstein, G., & Watson, J. R. (1989). Test-retest reliability of the Halstead-Reitan battery and the WAIS in a neuropsychiatric population. *Clinical Neuropsychologist*, *3*, 265–273.
- Hinton-Bayre, A. D., Geffen, G. M., Geffen, L. B., McFarland, K. A., & Friis, P. (1999). Concussion in contact sports: Reliable change indices of impairment and recovery. *Journal of Clinical Experimental Neuropsychology*, *21*(1), 70–86.
- Hopp, G. A., Dixon, R. A., Grut, M., & Backman, L. (1997). Longitudinal and psychometric profiles of two cognitive status tests in very old adults. *Journal of Clinical Psychology*, *53*(7), 673–686.
- Iverson, G. L., Gaetz, M., Lovell, M. R., and Collins, M. W. (2004). Relation between subjective foginess and neuropsychological testing following concussion. *Journal of the International Neuropsychology Society*, *10*, 904–906.
- Iverson, G. L., Lovell, M. R., & Collins, M. W. (2002). Validity of IMPACT for measuring the effects of sports-related concussion [Abstract]. *Archives of Clinical Neuropsychology*, *17*(8), 769.
- Iverson, G. L., Lovell, M. R., & Collins, M. W. (2005). Validity of ImPACT for measuring processing speed following sports-

- related concussion. *Journal of Clinical and Experimental Neuropsychology*, 27, 683–689.
- Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.
- Lovell, M. R., & Collins, M. W. (1998). Neuropsychological assessment of the college football player. *Journal of Head Trauma Rehabilitation*, 13(2), 9–26.
- Lovell, M. R., Collins, M. W., Podell, K., Powell, J., & Maroon, J. (2000). ImPACT: Immediate post-concussion assessment and cognitive testing. [Computer software]. Pittsburgh, PA: NeuroHealth Systems, LLC.
- Makdissi, M., Collie, A., Maruff, P., Darby, D. G., Bush, A., McCrory, P., et al. (2001). Computerised cognitive assessment of concussed Australian Rules footballers. *British Journal of Sports Medicine*, 35(5), 354–360.
- McKeever, C., & Schatz, P. (2003). Current issues in the identification, assessment, and management of concussions in sports-related injuries. *Applied Neuropsychology*, 10(1), 4–11.
- Moser, R. S., & Schatz, P. (2002). Enduring effects of concussion in youth athletes. *Archives of Clinical Neuropsychology*, 17(1), 81–90.
- Powell, J. W., & Barber-Foss, K. D. (1999). Traumatic brain injury in high school athletes. *Journal of the American Medical Association*, 282(10), 958–963.
- Schatz, P., & Browndyke, J. (2002). Applications of computer-based neuropsychological assessment. *Journal of Head Trauma Rehabilitation*, 17(5), 395–410.
- Schatz, P., & Putz, B. (2002). Use of reliable change indices, correlation coefficients, and dependent-samples *t*-tests to document stability of computerized assessment measures over a one-year period [Abstract]. *Archives of Clinical Neuropsychology*, 17(8), 851.
- Snow, W. G., Tierney, M. C., Zorzitto, M. L., Fisher, R. H., & Reid, D. W. (1988, February). One year test–retest reliability of selected tests in older adults. Paper presented at the meeting of the International Neuropsychological Society, New Orleans, LA.
- Spren, O., & Strauss, E. (1991). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press.
- Vispoel, W. P. (2000). Computerized versus paper-and-pencil assessment of self-concept: Score comparability and respondent preferences. *Measurement & Evaluation in Counseling & Development*, 33(3), 130–143.
- Wechsler, D. (1955). *Wechsler Adult Intelligence Scale manual*. New York: The Psychological Corporation.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale@150>Revised manual*. New York: The Psychological Corporation.
- Wilson, S. L., & McMillan, T. M. (1991). Microcomputers in psychometric and neuropsychological assessment. In A. Ager (Ed.), *Microcomputers and clinical psychology: Issues, applications and future developments* (pp. 79–94). Chichester, England: Wiley.

Copyright of Applied Neuropsychology is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.