2-2020

Convergent Validity of Baseline Concussion Measures

Shannon N. Dugan

Convergent Validity of Baseline Concussion Measures

by

Shannon N. Dugan, M.S.

A Doctoral Research Project

Presented to

The College of Psychology and Liberal Arts

School of Psychology

Florida Institute of Technology

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Psychology

Melbourne, Florida

February 2020

We the undersigned committee, having examined the submitted doctoral research project, "Convergent Validity of Baseline Concussion Measures" by Shannon Dugan, M.S. hereby indicates its unanimous approval.

_____

Frank M. Webbe, Ph.D.
Professor, School of Psychology
Major Advisor

_____

Anthony P. LoGalbo, Ph.D., ABPP-CN
Associate Professor, School of Psychology

_____

Natalie M. Dorfeld, Ph.D.
Associate Professor of English, School of Arts and Communication

_____

Lisa A. Steelman, Ph.D.
Professor and Dean
College of Psychology and Liberal Arts

ABSTRACT

Convergent Validity of Baseline Concussion Measures

Shannon N. Dugan, M.S.

Frank M. Webbe, Ph.D., Major Advisor

The aim of this study was to examine convergent validity of similar domains in two commonly used neurocognitive assessments for concussion, the Sports Concussion Assessment Tool-5 (SCAT-5) and the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT). The SCAT-5 and ImPACT will be described in great detail in the background review and only noted in the methods section of this paper. Baseline performance on these measures was analyzed from a sample of approximately 747 college athletes. Individual performance in the same neurocognitive domain was compared for consistency across SCAT-5 and ImPACT tests. Domains included immediate and delayed verbal memory, as well as concentration and attention. Additionally, group performance based on order of administration was compared to rule out the possibility of extraneous influences on performance. Results of this study confirm that these widely used concussion measures demonstrate convergent validity within the tested domains by showing stability within individual athlete's performance across measures; however, the size of

the effects ranged from small to medium. Findings also support the literature on the use of multiple cognitive performance measures in concussion management.

Table of Contents

Literature Review

A concussion is a mild traumatic brain injury (mTBI) characterized by a transient disruption in the metabolic functioning of the brain (Hubertus, Marklund, & Vajkoczy, 2019). Concussions occurring during sports, also known as sports-related concussions (SRCs), are common due to the high level of physical movement and the increased chances of contact between players. It is estimated that between 1.6 to 3.8 million SRCs occur each year (Mullally, 2017). Sports are the second most common cause of head injury among individuals between ages 15 and 24 years old (Mullally, 2017). Concussions can result from a variety of mechanisms, including a direct trauma to the head, such as that which occurs during a fall or when objects hit one's head. Another common mechanism of sports related concussion occurs from collisions, such as head-to-head collisions. A common misconception is that the head must be directly hit in order for a concussion to occur. However, a sudden violent motion causing an acceleration of the brain against the skull, such as "whiplash," can also cause a concussion (Hubertus, Marklund, & Vajkoczy, 2019). Even if an athlete's head does not hit the ground, an object, or another athlete's body part, a concussion can still result if an abrupt stop in movement occurs. As a concussion does not typically involve an opening of the head or fracture of the skull, it is considered to be a closed head injury.

Symptoms of concussion can involve physical, emotional, and cognitive functioning. While the recovery period from concussion varies by individual, they generally resolve in a period of 7-10 days (Giza & Hovda, 2001). During this time, the brain is in a vulnerable state due to metabolic disruption and altered blood flow, and neural matter is especially susceptible to further injury (Giza & Hovda, 2014). Therefore, accurate detection and diagnosis of concussion is crucial to optimize recovery and functioning of concussed individuals. This paper will focus on two commonly used assessment tools in sports-related concussion, the Sports Concussion Assessment Tool-5 (SCAT-5) and the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT), a computerized battery of tests designed to detect changes in neurocognitive functioning following sports-related concussion.

Test Retest Paradigm

In order to diagnose sports related concussions, a test-retest paradigm is typically utilized in which an individual's neurocognitive performance is measured over time. This method begins by capturing an individual's baseline level of performance before an injury occurs. Following a suspected injury, the athlete will then be retested using the same testing battery. Previous research has identified attention, processing speed, and working memory as the most sensitive to change after a mild traumatic brain injury (mTBI) (Barr &

McCrea, 2001). Extraneous factors that can affect performance on neurocognitive testing as they relate to sports-related concussion will be discussed.

The Use of Testing Batteries and Order of Administration

When conducting neuropsychological testing, a battery of tests may be employed. It is recommended that a minimum of two tests that measure the same domain be used to obtain valid results (Larrabee, Greiffenstein, Greven, and Bianchini, 2007). Along with the testing battery arises many possible orders of presentation for each test. In determining the order of presentation, factors including difficulty, domain, and timing requirements are taken into consideration. For example, if a test contains a list of verbal information, examiners should avoid presenting it immediately after a test containing a list of different verbal information, in order to minimize interference of information. In addition, if a delayed trial is included in a test, this measure would need to be placed in an order within the battery to allow for enough time to pass between initial and delayed presentation. Furthermore, some psychometrists may choose to place more difficult or lengthy tests at the end of a battery, so as to decrease the impact on effort or motivation early on in testing. However, these tests may also be placed early on in the testing battery so that fatigue does not interfere with performance and affect interpretation. The battery used in this protocol includes a combination of performance

validity tests (both verbal and visual) as well as a computerized neurocognitive test and a paper and pencil sports concussion assessment test. While the assessments are measuring similar domains, the tests contain different stimulus material, and are relatively brief in nature. In addition, when using separate modalities, such as pencil and paper versus computerized administration, it may be possible that exposure to stimulus material in one modality affects exposure in a second modality. Potential areas of influence can include a preference for verbal information presented auditorily or visually. The likelihood of interference contributing significantly to an athlete's performance is likely minimal because the battery and word lists utilized are relatively brief.

Furthermore, having been exposed to a computerized test should not affect performance on an in-person administration and vice versa. However, possible areas of interference in this battery include word lists from SCAT-5 and ImPACT interfering with one another. In this case, an appropriate amount of time is likely to have passed between presentation of different word lists. While both SCAT-5 and ImPACT assess for several of the same cognitive domains as well as symptom presentation, it is necessary to include both assessments in the battery, as the SCAT-5 contains balance information and questions about orientation (e.g. date, time, place), which is absent on ImPACT. Also, if an individual's performance on the same cognitive domain

differs between assessments, then having a back-up measure will give a more accurate picture of their functioning.

Motivation, Practice Effect, Fatigue

Bailey, Echemendia, and Arnett (2006) examined groups of athletes with varying levels of motivation at baseline concussion neurocognitive testing. Groups based on levels of motivation (high versus suspect baseline motivation) were created by identifying individuals who scored one or more standard deviations above or below the mean of all athletes for each baseline measure, respectively. Results confirmed that those with lower levels of motivation at baseline demonstrated increased improvement during post injury assessments, thus skewing their testing data to appear as if larger improvements in cognitive functioning had been made between pre and post injury. Such inaccurate findings can result in an athlete being returned to play before cognitive functioning has returned to baseline levels, which will place the athlete at increased risk of sustaining another concussion and/or experiencing harmful effects of multiple head injuries (Giza & Hovda, 2001, 2014; McKee et al., 2016).

In addition to motivation, neuropsychological test data are susceptible to the influence of other factors, including practice effects, and inattention (Beglinger et al., 2005). To minimize practice effects, alternative testing forms with different stimulus material are used during retesting, especially when

testing verbal memory. An approach utilizing a reliable change index for performance between baseline and post trauma and regression-based norms is then used to identify reasonable and reliable changes within an individual's performance over time (Iverson, Lowell, & Collins, 2003; Schatz & Robertshaw, 2014). Further, valid neuropsychological test administration generally includes multiple tests and can take up to several hours. For the purpose of sports related concussion testing, current approaches typically utilize brief test batteries that are more practical when testing large groups of athletes (Barr, 2001). Therefore, it is reasonable to expect that mental fatigue from testing would not affect results, especially during baseline as opposed to post-injury testing. However, mental or physical fatigue as a result of lack of sleep or jetlag, cannot be ruled out without self-report from an athlete, so such questions are included in questionnaires to be completed by athletes. Therefore, it is important to consider these factors when evaluating an athlete's performance on baseline measures, so that an accurate baseline is used for comparison in the event that a concussion is suspected.

Validity in Concussion Testing

Due to the multiple factors that can affect an individual's neuropsychological test performance, and thus yield inaccurate or skewed results, professionals use a test battery consisting of multiple tests to yield more reliable cognitive testing results for individuals. Furthermore, embedded

measures of validity are used to identify when an individual has performed unusually poor on a measure, indicating suboptimal effort. For example, the Rey Word Recognition Test and Dot Counting Tests have been used to identify instances in which an individual's performance may be invalid. These tests also comprise the performance validity tests used in the present study. Since effort can fluctuate throughout the course of a test battery administration, effort should be tested multiple times using multiple measures throughout testing (Proto et al., 2014).

In performance validity tests, cut off scores are identified, which present the minimum level of errors that have differentiated between normal and intentionally poor performance (Larrabee, 2014). Such measures have also been used to identify instances in which an individual is malingering neurocognitive dysfunction (Larrabee, Greiffenstein, Greven, and Bianchini, 2007). Larrabee and colleagues (2007) asserted that failure to meet or exceed cut off scores on two or more performance validity tests is sufficient enough to identify an individual who is malingering or exerting suboptimal effort during testing. An athlete may be motivated to perform poorly, or "fake bad" on a baseline assessment of neurocognitive function for concussion diagnostic purposes. If one performs poorly, or sub-optimally, on a measure during a baseline evaluation, he/she may believe they are less likely to be diagnosed with a concussion in the event of an injury

SCAT-5

The SCAT-5 consists of a symptom questionnaire assessing the severity of symptoms that may be associated with sports related concussions. The Standardized Assessment of Concussion (SAC) portion is a brief cognitive screening measure that assesses orientation, immediate and delayed memory, and concentration. Immediate memory is tested using a ten-item word list which is read aloud to the examinee and recall is recorded over three consecutive trials. The highest possible score in immediate memory is 30, with higher scores reflecting higher performance. The delayed memory trial occurs at the end of the SCAT-5, after a delay of at least five minutes. The concentration section of the SCAT-5 consists of two parts: digits backward and months backward. In digits backward, examinees are asked to repeat a sequence of numbers in the reverse order. Four items, ranging from a span of 3 digits to 6 digits, are included.

If the examinee responds incorrectly on the first trial of each item, a second trial is then administered. After two consecutive trial errors on a single item, the test is discontinued. The highest possible score on digits backward is 4. For months backward, the examinee is asked to recite the months of the year in reverse order, starting with December. Any uncorrected error in the form of an omission or sequencing error results in an assigned score of 0. The highest possible total concentration score is 5. The BESS portion assesses for postural

stability or balancing errors during a series of 3 positions held for a duration of 20 seconds each.

ImPACT

On the ImPACT, a widely used tool for the assessment and management of neurocognitive effects of sports concussion, algorithms are used to identify an invalid baseline test when an athlete exhibits a high amount of errors, beyond what would reasonably be expected on a particular subtest or module. The ImPACT will generate an invalid designation for a test taker if any of the following cut off scores are reached: a score of 30 or more on Xs and Os total incorrect, a score of 30 or more on the impulse control composite, less than 69% correct on word memory learning, less than 50% correct on design memory learning, and less than 8 total letters correct on Three Letters (ImPACT manual). In the event that any of these cut off scores are met, the administrator is to discuss the performance with the athlete, and identify possible causes of questionable validity, which can include difficulty understanding the instructions of one or more of the modules, inattention, or suboptimal effort. The athlete should then retake the entire test to obtain a valid score.

However, even when athletes are provided with instruction on how to "sandbag," or perform sub optimally on the ImPACT, it is difficult to accomplish successfully without invalidating the test (Erdal, 2012), indicating

the ImPACT is successful in preventing this phenomenon. Schatz and Glatts (2013) further demonstrated that intentionally performing poorly on the ImPACT at baseline without invalidating the test is difficult. Often, the ImPACT can be invalidated due to confusion of instructions on the Xs and Os module (Schatz, Moser, Solomon, Ott, & Karpf, 2012). For this module, a distracter task intended to interfere with memory rehearsal is used in which the examinee is asked to perform a specific action depending on whether a blue square or red circle is presented. Then, an assortment of Xs and Os are presented on screen, with three letters illuminated each trial and the examinee is instructed to remember the location of the illuminated letters. After each trial, the distracter task is presented, and the series of Xs and Os are again presented. Schatz and colleagues (2012) revealed no significant effects of sex or attention deficit disorder on valid ImPACT performance in a sample of high school and college students. Further, researchers revealed a rate of 4% to 11% of invalid baseline ImPACT tests among the sample. Compared to the performance of clinical populations on general neurocognitive testing, these rates are relatively low. Therefore, it appears that ImPACT serves as a reliable and valid tool for concussion diagnostic purposes, if those with invalid baseline scores are correctly identified and properly handled.

SCAT-5: Subtest Designs

On the SCAT-5, a digits-backward task is used to measure working

memory. For this task, subjects are read aloud a string of numbers by the administrator. Then, the subjects are asked to repeat the string of numbers in reverse order of how they are read. This requires the athlete to retain the string of numbers in their working memory while actively manipulating the information to repeat it backward. In addition, another SCAT-5 task (months backward), requires the athlete to recite the months of the year backward, and also requires an individual to simultaneously mentally "hold onto" the material while actively manipulating the information. The SCAT-5 subtest for measuring verbal immediate memory consists of a list of ten words read aloud to the athlete over three separate trials, for a total possible immediate memory score of 30. At the end of the SCAT-5 administration after a delay of approximately five minutes, the athlete is asked to recall as many words from the word list as possible, and the delayed memory score is obtained.

ImPACT: Subtest Designs

The first module on the ImPACT, Word Memory, evaluates attentional processes and verbal recognition memory using a word discrimination task. A list of 12 target words are shown on screen. Then, 24 test words including the original 12 target words, as well as 12 distractor words with the same semantic category as the target words are presented individually and the examinee responds "yes" or "no" to whether or not this word was one of the target words. Following a delay of approximately twenty minutes, the examinee is

again tested on their recall of the same list. Five different versions of the word list are used to minimize the possibility of a practice effect between administrations.

Working memory and visual motor response speed are measured in the sixth module on ImPACT, titled Three Letters. Subjects are instructed to click as quickly as possible on numbers 1 through 25 presented on screen, in backward order. The test taker is then presented with three consonants on screen. The number grid reappears, and the subject repeats the first task. After 18 seconds, the grid disappears, and the subject is asked to recall the three consonants. A memory score (total number of correctly identified letters) and a score for the average number of correctly clicked numbers per trial are provided. Five trials of this task are presented.

Reliability

The reliability of an instrument refers to its ability to reflect a score that is minimally affected by error (Lovell et al. 2006). Test reliability is a continuous variable, with a high degree of reliability indicating an instrument's ability to reflect an accurate individual score in a specific population (Franzen, 2000). Therefore, reliability for different factors within the same test can vary from one another, depending on what is being measured and the population in which it is being measured. Internal consistency plays an important role in test

reliability and can be measured by comparing within individual performance on the same test at two separate times.

The test-retest reliability of ImPACT was studied by Schatz (2010) in a sample of 95 collegiate athletes who completed the test approximately two years apart. None of the athletes had sustained a concussion between testing. Intraclass correlation coefficient estimates for visual memory (0.65), processing speed (0.74), and reaction time (0.68) composite scores were found, suggesting stability in performance over time. The greatest variability was found in verbal memory (0.46) and the symptom scale scores (0.43). Further, reliable change indices and regression-based analyses revealed a rather small proportion of participants' scores showed reliable or significant change on the composite scores (0%-6%) between assessments. Such findings suggest that non-concussed collegiate athletes' performance on ImPACT at baseline remains considerably stable over a two-year period.

A separate study was conducted to evaluate the reliability of the ImPACT test over a shorter time period of one month between assessments (Schatz & Ferris, 2013). The following Pearson's correlation coefficients *(r)* and intra-class correlation coefficients (ICCs) were as follows: verbal memory: $r = 0.66$, ICC =.79; visual memory: $r = .43$, ICC =.60; visual motor speed r =.78, ICC = .88; reaction time: $r = .63$, ICC = .77. Dependent sample *t*-tests were conducted, revealing significant changes in performance on only visual

motor speed composite scores, with performance improving over time. This may have occurred due to a practice effect, and familiarity with the test, with not enough time in between testing to eliminate the effect. Reliable Change Indices revealed a significant number of participants fell outside 80% and 95% confidence intervals using regression-based measures.

Therefore, it appears that issues with test-retest reliability on ImPACT over shorter time periods is isolated to the visual motor speed composite. Meanwhile, repeated exposure to ImPACT test across one month do not lead to practice effects in memory performance or reaction time. Additional test-retest reliability studies have concluded that ImPACT is a reliable neurocognitive test battery at 45 and 50 days after the initial baseline assessment (Nakayama et al. 2014). Other reliability studies have reported acceptable intraclass correlation coefficients spanning testing intervals of 30 days to one year (ImPACT Administration and Interpretation Manual).

Chin and colleagues (2016) investigated test-retest reliability of the SCAT-3 in a sample of 2018 high school and collegiate athletes (Chin, Nelson, Barr, McCrory, & McCrea, 2016). Over a 7-day interval between testing, the Pearson correlation coefficient was 0.63 for symptoms, 0.49 for SAC, and 0.57-0.66 (male-female) for BESS. A longer test-retest interval of 196 days was tested and yielded less reliable Pearson coefficients: 0.45 for symptoms, 0.41 for SAC, and 0.53 for BESS. The differences in SCAT-3 and SCAT-5 lie

in the number of words in the immediate memory word list. The SCAT-5 immediate memory word list consists of 10 words, as opposed to 5 words in the SCAT-3, which resulted in a more normally distributed curve with respect to performance. Norheim, Kissinger-Knox, Cheatham, & Webbe (2018) demonstrated that the 10-item word list successfully eliminated the ceiling effects for immediate and delayed memory associated with the 5-item list. Studies examining the construct validity for SCAT-5 were not found in the literature, although Chin, Nelson, Barr, McCrory, and McCrea (2016) reported adequate validity for the SCAT-3.

However, the SCAT-5 components measuring immediate and delayed memory are similar to many well developed and commonly used verbal memory tests consisting of a word list being read aloud to the examinee across multiple trials and a delayed recall trial shortly after, including the California Verbal Learning Test. In addition, the digits backward component of the SCAT-5 is a shorter variation of the Digit Span subtest of the Wechsler Adult Intelligence Scale- 4th edition, which is a commonly and well supported cognitive measure of working memory (Ruchinskas, 2019). While the research on the reliability and validity of the SCAT-5 is limited and the existing findings are not overwhelmingly supportive, few tools of its kind exist. Due to its similarity to well supported measures of verbal and working memory, it is expected to measure these constructs.

Validity

Construct validity for the ImPACT test was examined in a sample of 72 athletes by Iverson, Lovell, and Collins (2005) using the Symbol Digits Modalities Test (SDMT), a commonly used tool in sports concussion research. The SDMT measures attention and processing speed, specifically scanning and tracking aspects (Strauss, Sherman, & Spreen, 2006). Athlete's performance on SDMT correlated strongly with ImPACT processing speed ($r = 0.70$, $p < .01$), and a moderate correlation was found with ImPACT reaction time ($r = .60$, $p < .01$) (Iverson, Lovell and Collins, 2005). As hypothesized, performance on the SDMT was less correlated with verbal memory ($r = .46$, $p < .01$), and an even weaker correlation was found with visual memory ($r = .37$, $p < .01$). In a smaller sample of 30 college student volunteers, Schatz and Putz (2006) utilized the Trail Making Test A & B (TMT) and the Wechsler Ault intelligence Scale-Revised (WAIS-R) Digit Symbol subtest to measure construct validity for the ImPACT. Findings supported construct validity for reaction time as it was significantly correlated with TMT A ($r = .64$, $p < .05$) and TMT B ($r = .44$, $p = < .01$), and significantly correlated with WAIS-R Digit Symbol ($r = .46$, $p < .01$). In addition, WAIS-R Digit symbol also correlated significantly with ImPACT processing Speed index (0.51). Maerlender (2010) studied construct validity of ImPACT by examining correlations with a comprehensive cognitive test battery (N=54) consisting of the California

Verbal Learning Test (CVLT) of verbal memory, the Brief Visuospatial Test-Revised (BVMT-R) of visual memory, Trail Making Tests A and B, Verbal Fluency, Delis Kaplan Executive Function System (DKEFS) Color Word memory for cognitive speed, Conners' Continuous Performance Test (CPT) for reaction time, and the Paced Auditory Serial Addition Test (PASAT) for working memory. Such tests are sensitive to cognitive functions associated with mTBI (Maerlender, 2010).

Analyses revealed significant correlations between the aforementioned traditional pencil and paper testing domains and appropriate ImPACT composite scores, with the exception of the Impulse Control factor. Maerlnder (2010) concluded that the cognitive domains measured by ImPACT have good construct validity with standard neuropsychological tests. However, they noted the utility of ImPACT as a screening tool that should be administered with additional tests of neuropsychological functioning to provide a valid assessment. Overall, findings of multiple studies support the reliability and validity of ImPACT. Results using the ImPACT are largely stable over time, and accurately measure the constructs of memory and reaction time.

Convergent validity, also known as concurrent validity, is the relationships between two scores from different tests that aim to measure the same concept. In this study, the convergent validity of the following domains will be examined between SCAT-5 and ImPACT:

Objectives

The objectives of this study are the following: examine whether immediate and delayed memory performance is consistent across SCAT-5 and ImPACT tests by comparing SCAT-5 immediate memory to ImPACT word memory, and SCAT-5 delayed memory to ImPACT delayed memory. In addition, concentration and attention performance will be compared across SCAT-5 total concentration score and ImPACT Three Letters scores. Finally, group performance in these domains (verbal memory and concentration) will be compared based on order of administration in baseline testing. See Table 1 for the subtest comparisons for each domain in this study.

Table 1

Cognitive Domains and Subtest Comparisons

Domain	Scat Subtest	ImPACT Subtest
Verbal Memory	Immediate memory Delayed Memory	Word Memory Delayed Memory
Concentration and Attention	Concentration	Three Letters

Hypotheses

Since the two tests aim to measure the same constructs within the same athlete, performance within measures of verbal memory and concentration should be similar within individual participants. Within verbal memory, both immediate and delayed performance measures will be compared. Scores on

ImPACT total percent correct were chosen for comparison to SCAT-5 immediate memory due to its measure of participant accuracy in recalling and identifying target words. Additionally, attention and concentration performance should also be similar within individual athlete performance. For attention and concentration, it is predicted that individual athlete performance will be similar across SCAT-5 concentration score and ImPACT three letters average counted correctly. The SCAT-5 concentration score was chosen due to its inclusion of subtasks involving retaining and sequentially manipulating information. This SCAT-5 concentration score is comprised of performance on a digit's backward subtest, in which participants are required to retain a string of numbers in memory and repeat them in backwards order. The other component of the SCAT-5 concentration score is derived from the months backward subtest that similarly requires the participant to utilize working memory to repeat the months in a backward sequence. Lastly, the ImPACT three letters subtest average counted correctly score was chosen for comparison with the SCAT-5 concentration score, as this subtest employs a similar cognitive task in which participants must click on numbers 1 through 25 in a backwards sequence.

The different modalities of presentation of testing (computerized versus paper and pencil test) should not affect performance in athletes. This issue of presentation order will be examined by comparing group performance on the

SCAT-5 between groups who completed the SCAT-5 first, and those who completed the SCAT-5 second. Relatedly, group performance on ImPACT will be compared between those who completed ImPACT first and those who completed ImPACT second within the testing battery. Increased performance on the test administered second is not to be expected, as the SCAT-5 and ImPACT are reasonably disparate in terms of tasks and stimuli overall, and therefore a practice effect is not predicted. Possible explanations for potential findings of increased performance on one type of modality can include the environment in which the test is administered, as the SCAT-5 is administered individually, and the ImPACT is generally administered in a group setting. As the same protocol is being used across multiple sessions and multiple teams, an effect of team membership or time of test administration is not expected to emerge.

Method

Participants

Participants included in the final data analysis for the study included 747 participants, 484 men and 263 women, comprised of collegiate student-athletes administered the assessment battery by experienced examiners in the 2016-2017, 2017-2018, and 2018-2019 athletic seasons. All athletes attended Florida Institute of Technology and were required to complete baseline testing in order to participate in their sport. Age of participants ranged from 18 to 25.

The collegiate athletic teams are comprised of 16 different sports teams (specific sports included in this study are outlined below). All participants provided their signature verifying their consent to participate.

Instrumentation

The Sport Concussion Assessment Tool–5th Edition (SCAT-5) is a standardized instrument for evaluating athletes for sports related concussion. The Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) is a computerized battery of tests designed to detect changes in neurocognitive functioning following sports-related concussion. Both tests are utilized during baseline evaluations for concussion at Florida Institute of Technology.

Procedure

Athletes were administered both the SCAT-5 and ImPACT. Individual teams were scheduled together. In addition, participants completed the Patient Health Questionnaire as a brief screener for psychological distress. During each testing session, athletes were divided randomly into groups and administered either the computerized testing first, followed by the remaining paper format tests, or vice versa. The ImPACT test was administered in a computer lab with no more than 22 students per group, with one or two examiners present. An invalid score on ImPACT, determined using ImPACT's built in validity indicator, resulted in a second administration of the baseline ImPACT in order to obtain a valid score. SCAT-5 and the remaining paper

tests were administered individually in classrooms or smaller research rooms. Total testing time was approximately two hours per group. The duration of the delay between each athlete's completion of the SCAT-5 and ImPACT ranged from one minute to one hour. This is due to the nature of the baseline testing administration currently utilized in concussion management at Florida Institute of Technology, in which large groups of athletes are to be tested across multiple sessions. This delay for each participant was not documented.

Statistical analyses

Data were analyzed using IBM Statistical Package for the Social Sciences version 25 (SPSS-25). For the first objective, within-person comparisons were made for SCAT immediate memory and ImPACT immediate memory percent correct and delayed memory percent correct. For the second objective, within-person correlations were also calculated for digits backward and three letters for the attention and concentration domain. For the third research objective, average scores were calculated for two groups based on whether ImPACT was administered first or second, and the score differences between groups were examined using a *t*-test.

Results

The sample of 747 participants consisted of 484 men and 263 women. The age of participants ranged from 18 to 25 years ($M = 19.5$, $SD = 1.54$). The frequencies for each age group was 18 (36%), 19 (19%), 20 (20%), 21 (14%),

22 (8%), 23 (1%), 24 (2%) 25 (<1%). Age was non-normally distributed, with a moderate skewness of .725 (*SE* = .09) and kurtosis of .102 (*SE* = 0.18). The largest percent of the sample by sport was represented by football (23.6%), followed by lacrosse (11.5%), soccer (11.1%), rowing (9.1%), swimming (8.6%). Other sports included baseball, basketball, bowling, cheerleading, dance, golf, softball, tennis, track and field, volleyball, and cross country.

Verbal Memory

It was hypothesized that SCAT-5 total immediate memory scores will have a positive correlation with ImPACT word memory total percent correct. Results of the Pearson correlation indicated that there was a significant positive association between SCAT-5 total immediate memory correct and ImPACT word memory total percent correct, $r(747) = .29$, $p < .001$), reflecting a small effect. See Table 2 for all verbal memory score correlations.

Secondly, it was hypothesized that SCAT-5 delayed memory score will have a positive correlation with ImPACT word memory hits delayed. Results of the Pearson correlation indicated that there was a significant positive association between SCAT-5 delayed memory score and ImPACT word memory hits delayed, $r(746) = .23$, $p < .001$). The largest correlation found was between SCAT-5 delayed memory score and ImPACT word memory total percent correct $r(746) = .33$, $p < .001$), reflecting a medium effect. See Table 3 for correlation values between SCAT-5 and ImPACT verbal memory subtests.

Table 2

Verbal Memory Descriptive Statistics

Variable	*M*	*SD*
1. SCAT-5 Immediate Memory	20.52	3.25
2. ImPACT Word Memory Percent Correct (%)	93.75	6.36
3. SCAT-5 Delayed Memory	6.59	1.83
4. ImPACT Word Memory Total Correct	11.59	.76
5. ImPACT Delayed Memory	10.69	1.43

Table 3

Verbal Memory Correlations

Variable	SCAT-5 Immediate Memory	ImPACT Word Memory Percent Correct (%)	SCAT-5 Delayed Memory	ImPACT Word Memory Total Correct
1. SCAT-5 Immediate Memory				
2. ImPACT Word Memory Percent Correct (%)	.29**			
3. SCAT-5 Delayed Memory	.64**	.33**		
4. ImPACT Word Memory Total Correct	.17**	.62**	.18**	
5. ImPACT Delayed Memory	.20**	.75**	.23**	.47**

**. Correlation is significant at the 0.01 level (2-tailed).

Concentration and Attention

It was hypothesized that SCAT-5 concentration scores will have a positive correlation with ImPACT Three Letters average counted correctly. Results of the Pearson correlation indicated there was a significant positive association between SCAT-5 Concentration score with ImPACT Three Letters average counted correctly $r(747) = .34$, $p < .001$), reflecting a medium effect. Table 5 shows Pearson correlation results for several attention and concentration subtests. SCAT-5 subtests include digits backward, months backward, and concentration total. ImPACT scores included in the Pearson correlation are shown for Three Letters total letters correct, Three Letters percentage letters correct (%), Three Letters average counted correctly, and Three Letters total sequence correct.

Table 4

Attention and Concentration Descriptive Data

Variable	*M*	*SD*
1. Digits Backward	2.89	.95
2. Months	.85	.36
3. Concentration Total	3.74	1.10
4. Three Letters Total Letters Correct	13.96	1.54
5. Three Letters Percentage Letters Correct (%)	93	10
6. Three Letters Average Counted Correctly	18.44	4.05
7. Three Letters Total Sequence Correct	4.44	.77

Table 5

Concentration and Attention Correlations

Variable	Digits Backward	Months	Concentration	Total Letters Correct	% Letters Correct	Average Counted Correctly
1. Digits Backward						
2. Months	.22**					
3. Concentration Total	.94**	.53**				
4. Three Letters Total Letters Correct	.12**	.04	.12**			
5. Three Letters Percentage Letters Correct (%)	.12**	.04	.12**	1**		
6. Three Letters Average Counted Correctly	.31**	.21**	.34**	.17**	.17**	
7. Three Letters Total Sequence Correct	.11*	.04	.11**	.93**	.93**	.18**

**Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Order of Administration

It was hypothesized that athletes performing SCAT-5 first would perform similarly to athletes completing SCAT-5 second across measures of verbal memory, attention, and concentration. An independent-samples t-test was conducted to compare both SCAT-5 immediate and delayed memory scores for SCAT-5 first and SCAT-5 second groups. An independent samples *t*-test revealed no significant group differences in SCAT-5 immediate memory scores between those who completed SCAT-5 first and second. Further, no significant group differences in SCAT-5 delayed memory between those who completed SCAT-5 first and second. See Table 6 for results of the independent samples-test comparing SCAT-5 verbal memory scores between groups who completed ImPACT first versus ImPACT second.

Table 6

Differences Between Groups on SCAT-5 Verbal Memory

Subtest	Group	*M*	*SD*	*t*	*p*	*df*
Immediate Memory	ImPACT first	20.6	3.23	.30	.77	745
	ImPACT second	20.5	3.26			
Delayed Memory	ImPACT first	6.51	1.84	1.25	.21	744
	ImPACT second	6.67	1.81			

Regarding order of administration and word memory on ImPACT, an independent-samples t-test was conducted to compare memory scores for ImPACT first and ImPACT second groups. Results revealed no significant differences in scores between the two groups for ImPACT word memory hits, verbal memory composite scores, and word memory hits delayed. See Table 7 for values.

Table 7

Differences Between Groups on ImPACT Verbal Memory

Subtest	Group	*M*	*SD*	*t*	*p*	*df*
Word Memory Hits	ImPACT first	11.64	.76	1.72	.09	745
	ImPACT second	11.54	.76			
Verbal Memory Composite	ImPACT first	88.66	9.79	-.07	.95	745
	ImPACT second	88.71	9.68			
Word Memory Hits Delayed	ImPACT first	10.75	1.44	1.14	.26	745
	ImPACT second	10.63	1.42			

Regarding concentration and order of administration, an independent-samples t-test was conducted to compare SCAT-5 digits backward, months backward, and total concentration scores for ImPACT first and ImPACT second groups. Results revealed no significant differences between scores on SCAT-5 digits backward, months backwards, and total concentration scores in those who completed ImPACT first and those who completed ImPACT second. See Table 8 for results of the independent samples t-test.

Table 8

Comparison of Group Performance on SCAT-5 Concentration Subtests

Subtest	Group	*M*	*SD*	*t*	*p*	*df*
Digits Backward	ImPACT first	2.92	.92	.75	.45	745
	ImPACT second	2.97	.98			
Months Backward	ImPACT first	.85	.36	.38	.70	703
	ImPACT second	.84	.37			
Total Concentration	ImPACT first	3.76	1.06	.65	.52	745
	ImPACT second	3.71	1.13			

For the ImPACT measures Three Letters total correct, Three Letters total sequence correct, and Three Letters total correct on ImPACT, Levene's test was not significant, suggesting that variances were equivalent. However, the assumption of normality of distribution was violated according to the Shapiro-Wilk test. Therefore, the nonparametric Mann-Whitney *U*-test was conducted to compare group performance based on order of administration**.** The Mann-Whitney test found no difference between the group of athletes who were administered ImPACT first versus second for any of these three variables.

Discussion

Verbal Memory

Analyses of this study aimed to examine whether immediate and delayed memory performance was consistent across SCAT-5 and ImPACT tests. Results suggest that while performance across SCAT-5 and ImPACT measures of immediate and delayed verbal memory are positively correlated, the small to medium effect size was less than expected. The strongest correlation found was between SCAT-5 delayed memory score and ImPACT word memory total percent correct, as the ImPACT word memory total percent correct score reflects performance on both immediate and delayed trials.

Factors which may have influenced scores and thus leading to the finding of small effect sizes are related to design differences of each test. First,

the number of words used as stimuli varies between SCAT-5 and ImPACT, as there are 10 words in SCAT-5 repeated across three trials for a total of 30 possible points, compared to 12 words in ImPACT. As the number of items to be recalled by a participant grows, more variability in performance is allowed as ceiling effects are eliminated (Norheim, Kissinger-Knox, Cheatham, & Webbe (2018). Considering the SCAT immediate memory word list was increased from 5 words in the SCAT-3 to 10 words in the later SCAT-5 edition, it may be concluded that a larger number of words in the list is preferred. As seen in table 2, the mean SCAT-5 average score was well below the highest possible score of 10, indicating that the 10-item word list is adequate in eliminating the ceiling effect in this sample. In contrast, a ceiling effect was demonstrated for ImPACT memory scores, as the mean score was close to the maximum possible score. This discrepancy may have contributed to the finding of a weak correlation between SCAT-5 and ImPACT memory performance in this study.

Additionally, the type of verbal memory tasks differs between the two tests. More specifically, the word memory subtest on ImPACT evaluates attentional processes and verbal recognition memory through a word discrimination task while the SCAT-5 uses a spontaneous immediate recall task. Research by Sternberg and Tulving (as cited in Blumenfeld & Ranganath, 2007) found that recall performance is strongly influenced by organizational

strategies which can be measured by analyzing characteristics of items that are recalled and the order in which they are recalled. For example, a phenomenon known as semantic clustering is found when participants tend to recall semantically similar items together from a list of categorically related words, which occurs even if the items were distributed throughout the worst list. Another similar phenomenon is subjective organization, which occurs when the same list is used repeatedly and tested, such as in the SCAT-5. Here, participants tend to recall sets of items in the same order across different recall trials. Researchers explain these phenomena by theorizing that associations are formed between items during encoding, which are subsequently used to guide retrieval (Sternberg & Tulving as cited in Blumenfeld & Ranganath, 2007). Therefore, the cognitive processes used in word discrimination in ImPACT differs from the recall task used in SCAT-5. While an individual must be attending during both tasks in order to be able to recall or discriminate later in both tests, the ImPACT may provide an advantage. For example, participants are shown 24 test words, which include 12 target words and 12 distractor words during the ImPACT verbal memory discrimination task. They are then asked to choose whether each of the 24 test words was in fact a target word they were originally shown through a forced choice task. Mathematically, they have a 50% chance of choosing a correct response for each trial, even if the test word was not actually recalled and the participant guesses during their

response. In SCAT-5, the task is comparatively simpler in nature, but it is more difficult for a participant to employ guessing to increase scores, as no distractor words are presented in the simple immediate recall task on SCAT-5. However, the same list of words is presented three times, and performance should thereby increase with each trial. The ImPACT Word Memory Total Percent score takes overall performance into account, that is whether a participant was able to correctly identify a target word, and also whether they were able to respond "no" to correctly identify a distractor word.

Additionally, another design factor which comes into play is the modality in which verbal stimuli is presented in each test. SCAT-5 verbal memory words are read aloud by the examiner, while ImPACT verbal memory is presented to the participant visually on screen. Research on the neural processing of verbal working memory has demonstrated that similar brain regions, specifically prefrontal and parietal regions, are involved in both auditory and visual verbal working memory; however, there are important modality differences in the way neural signals are generated, processed, and routed during verbal working memory tasks (Crottaz-Herbette, Anagnoson, & Menon 2004). A study conducted by Deboth and Dominowski (1978) investigated the possible interaction of individual differences in learning with mode of presentation. A sample of 160 college students learned four lists of 20 words each, two lists were presented auditorily and two lists were presented to

participants visuals. Results of their study demonstrated reliable individual differences in learning, however, researchers were unable to reliably classify participants in terms of auditory or visual preference (Deboth & Dominowski, 1978). Therefore, it is unlikely that differences in presentation modality of verbal stimuli contributed to differences in individual scores in verbal memory on ImPACT and SCAT-5.

Regarding delayed memory, the amount of time between the first presentation of words in each test varies, thereby possibly affecting performance. As time increases, the ability to recall or discriminate between the original words decreases. In the SCAT-5, the delayed recall task is administered approximately 5 minutes following the immediate recall task, whereas the delay in ImPACT is 20 minutes, which is considerably longer. Results comparing SCAT-5 delayed memory score and ImPACT word memory hits delayed revealed that individual delayed memory performance is significantly positively correlated across tests; however, the effect is small. These differences in the length of the delay between immediate and delayed memory between the two tests may have contributed to variability in individual delayed memory performance.

The SCAT-5 employs a repeated learning trial method in which the same list of words is read aloud to the examinee over three trials. The Rey Auditory Verbal Learning Test (RAVLT; Frederick, 2003) is a relatively more

complex, commonly used neuropsychological measure of verbal learning and memory which also includes repeated trials of a 15-item word list. An analysis of 58 groups of nonclinical adults and children on the free recall trials of the RAVLT has demonstrated a learning curve, in which verbal recall generally increases across repeated trials (Poreh, 2005). Therefore, it is reasonable to assume that the repeated trial method employed in SCAT-5 may contribute to individual differences in performance across SCAT-5 and ImPACT used in this study, as the ImPACT only uses one trial in administration.

Further analysis in this study revealed a larger effect size for the correlation between SCAT-5 delayed memory score and ImPACT word memory total percent correct. Word memory total percent correct on ImPACT considers the total correct responses for both target and distractor words, thereby tapping into discriminative abilities, which again is a more complex task than the simple recall task employed in the SCAT-5. This suggests that the word memory total percent score derived from ImPACT yields more similar performance in participants to the SCAT-5 total delayed recall score. Further, the inclusion of accuracy in identifying distractors as well as target words in the ImPACT word memory total percent correct score allows for more variability in performance, thus yielding more specific data for test interpretation. For example, if an examinee responded “yes” to all possible words, a score measuring only true positive responses would falsely be

interpreted as high. By identifying false positive responses, a more accurate interpretation of performance is reflected. Further, this characteristic of the ImPACT word memory total percent correct score therefore has more variability.

As mentioned earlier, research by Schatz (2010) on the test-retest reliability of ImPACT composite scores in collegiate athletes who completed the test approximately two years apart suggests overall stability in performance over time. However, the greatest variability was found in verbal memory and the symptom scale scores. While the present study compared within individual verbal memory performance across SCAT-5 and ImPACT, variability in verbal performance appears to be supported by other findings (Schatz, 2010).

Attention and Concentration

The second objective of this study aimed to compare concentration and attention performance across the two tests. It was hypothesized that SCAT-5 concentration score would have a positive correlation with ImPACT Three Letters average counted correctly. As can be seen in table 3, results of the analyses supported this hypothesis. Similar to the verbal memory findings, results revealed a significant positive association between the two scores, with a medium effect size. While significant, the amount of variability in scores on one baseline measure accounted for by individual performance on the other comparably similar cognitive domain measure is small. The effect size for the

correlation of performance on the compared tasks involving concentration and attention were not as large as expected. One possible factor accounting for this result is the that the design of the specific subtests used by each concussion baseline measure were not entirely the same. For example, the three letters average counted correctly subtest on ImPACT requires individuals to respond as quickly as possible, adding another processing speed component to the task and thereby increasing the cognitive load placed on the participant. This differs from the SCAT-5 concentration subtests which are untimed. The added level of a timed task could account for individual differences in scores on the compared tasks. Further, the SCAT-5 concentration subtest digits backwards, is also presented auditorily, whereas ImPACT Three Letters is a visual-motor task.

The ImPACT Three Letters component is primarily a measure of memory, with distractor tasks which tap into attention and concentration constructs. As such, it is not less surprising none of the Three Letter scores were very strongly correlated with SCAT-5 concentration subtests.

While studies examining the construct validity for SCAT-5 are scant in the literature, results of this study suggest that the SCAT-5 concentration measures are adequate. This is consistent with findings from Chin, Nelson, Barr, McCrory, and McCrea (2016), which reported adequate validity for the SCAT-3.

Order of Administration

Further analyses conducted aimed to compare group performance based on order of administration in baseline testing. As predicted, results suggest that no significant group differences were found in participants' scores among those who completed ImPACT first or second. The results suggesting the absence of significant differences in group performance based on order of presentation, thereby providing support against the concern that completing one test may diminish or enhance athlete performance on another measure due to factors such as fatigue or practice effect. To that end, this finding supports the current approach of utilizing brief test batteries in testing large groups of athletes for the purpose of sports-related concussion testing (Barr, 2001). Further, the tests used in this study include alternate forms in order to decrease nonclinical improvements in performance over consecutive test-retest trials often employed in sports concussion related testing (Beglinger et al., 2005). Additionally, the only ImPACT score that has been demonstrated to be significantly vulnerable to practice effects within a short period of time, 2 months, is the visual motor speed composite (Schatz & Ferris, 2013), which was not a factor in this study. Therefore, it is unlikely that practice effects contributed to the rather small correlation found between the measures in the study.

Limitations

As stated earlier, the duration of the delay between each athlete's completion of the SCAT-5 and ImPACT ranged from one minute to one hour. It is to be noted that the delay from completion of ImPACT or SCAT-5 to the initiation of the following test was not measured. As discussed earlier, potential effects of the variable delays between tests include the potential for interference from the word lists.

Additionally, the tests included in the concussion battery used in this study are limited. While there are relatively equivalent memory tests on the SCAT-5 and ImPACT, choosing similar attention and concentration tasks from each test was more difficult. ImPACT generates composite scores for reaction time and visual-motor speed through time or speed-based tasks, which are not found on the SCAT-5. Nonetheless, it is argued that these tasks fit within the constructs of attention and concentration. Therefore, correlational analyses in this study were limited by the different types of subtests included in each of the instruments to measure certain cognitive constructs.

The sample utilized in this study represent a narrow portion of the general population due to demographics including age and level of education. Therefore, results may not be generalizable to individuals from other demographics, such as older adults. Future baseline concussion management

studies may consider studying the effect of concussion history and attentional disorders on within-individual performance across separate measures.

Implications

The impact of these findings is meaningful as the current protocol used for concussion management affects concussion diagnosis, and consequently the health of collegiate athletes and other populations affected by mTBI. The results of this study support that widely utilized measures of cognition in concussion management have convergent validity within the domains of verbal memory, attention and concentration; as stability was found within individual athlete performance across the related measures of this study, findings confirm that these tests are adequate in measuring the related constructs.

Notably, the small to medium effect sizes found for the significant positive correlations between comparable SCAT-5 and ImPACT subtest scores support that components of baseline concussion tests measure similar cognitive domains while maintaining their own unique utility. For example, the SCAT-5 includes a balance component, orientation questions, and a somewhat more in-depth report of athlete symptomatology chronicity, which is not found on ImPACT, and thus provides valuable information in the detection of concussion in its own right. However, other cognitive domain performance relating to concussion, such as processing speed and visual motor speed, is not measured by the SCAT-5. Meanwhile, the ImPACT subtests include several

subtests which measure these tasks, thereby providing additional fundamental information regarding the athlete's cognitive status. The high degree of variability in the presentation of concussive sequalae, including cognitive functioning, underscores the importance of using multiple measures that are slightly different from each other during diagnostic testing. Similarly, broader neuropsychological research beyond the scope of concussion diagnosis recommends the use of more than on instrument to measure a similar cognitive domain (Barr, 2001). As such, findings of the study provide further support of the use of multiple measures in order to accurately detect and diagnose concussion. Overall, this study supports the clinical utility in administering both SCAT-5 and ImPACT baseline concussion tests in a complimentary fashion.

Lastly, the nonsignificant findings on order of administration in this study reject the notion that this factor may extraneously affect athlete performance. Thus, the currently used concussion protocol by concussion management at this university is supported. Findings on extraneous factors provide valuable information on the internal and external validity of these measures, which ultimately impact concussion diagnosis and recovery within student-athletes. Future research should aim to illuminate other factors influencing variability in cognitive performance as they relate to concussion testing.

References

Bailey, C.M, Echemendia, R.J, & Arnett, P.A. (2006). The impact of motivation on neuropsychological performance in sports-related mild traumatic brain injury. *Journal of the International Neuropsychological Society, 12,* 475-484.

Barr, W. B. (2001). Methodologic issues in neuropsychological testing. *Journal of Athletic Training, 36*(3), 297.

Barr, W. B., & McCrea, M. (2001). Sensitivity and specificity of standardized neurocognitive testing immediately following sports concussion. *Journal of the International Neuropsychological Society, 7*(6), 693-702.

Beglinger, L. J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D. A., Crawford, J., Fastenau, P.S., & Siemers, E. R. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology, 20*(4), 517-529.

Chin, E. Y., Nelson, L. D., Barr, W. B., McCrory, P., & McCrea, M. A. (2016). Reliability and validity of the Sport Concussion Assessment Tool–3 (SCAT3) in high school and collegiate athletes. *The American Journal of Sports Medicine, 44*(9), 2276-2285.

Crottaz-Herbette, S., Anagnoson, R. T., & Menon, V. (2004). Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. *Neuroimage, 21*(1), 340-351.

DeBoth, C. J., & Dominowski, R. L. (1978). Individual differences in learning: Visual versus auditory presentation. *Journal of Educational Psychology, 70*(4), 498-503. doi:http://dx.doi.org.portal.lib.fit.edu/10.1037/0022-0663.70.4.498

Erdal, K. (2012). Neuropsychological testing for sports-related concussion: how athletes can sandbag their baseline testing without detection. *Archives of Clinical Neuropsychology*, *27*(5), 473-479.

Franzen, M. D. (2000). Validity as applied to neuropsychological assessment. *Reliability and Validity in Neuropsychological Assessment* (pp. 33-54). Springer, Boston, MA.

Frederick, R. I. (2003) A review of Rey's strategies for detecting malingered neuropsychological impairment. *Journal of Forensic Neuropsychology, 2*(3-4), 1-25.

Giza, C. C., & Hovda, D. A. (2001). The neurometabolic cascade of concussion. *Journal of athletic training, 36*(3), 228–235.

Giza, C.C., & Hovda, D. A. (2014). The new neurometabolic cascade of concussion. *Neurosurgery, 75*(4), S24-S33.

Hubertus, V., Marklund, N., & Vajkoczy, P. (2019). Management of concussion in soccer. *Acta Neurochirurgica*, 1-9.

ImPACT administration and interpretation manual, ImPACT Applications, Inc., Pittsburgh, PA, United States, 2016.

Iverson, G.L., Lovell, M.R., Collins, M.W. (2003). Interpreting change on ImPACT following sport concussion. *The Clinical Neuropsychologist, 17*(4). 460-467.

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*(6), 683-689.

Larrabee, G. L., Greiffenstein, M. F., Greve, K.W., & Bianchini, K. J. (2007). *Refining diagnostic criteria for malingering* (pp. 334-369). In G. L. Larrabee (Ed.), Assessment of malingered neuropsychological deficits. New York: Oxford University Press.

Lovell, M. R., Iverson, G. L., Collins, M. W., Podell, K., Johnston, K. M., Pardini, D., & Maroon, J. C. (2006). Measurement of symptoms following sports-related concussion: reliability and normative data for the post-concussion scale. *Applied Neuropsychology, 13*(3), 166-174.

Maerlender, A., Flashman, L., Kessler, A., Kumbhani, S., Greenwald, R., Tosteson, T., & McAllister, T. (2010). Examination of the construct validity of ImPACT computerized test, traditional, and experimental neuropsychological measures. *The Clinical Neuropsychologist, 24*(8), 1309 -1325.

McKee, A. C., Cairns, N. J., Dickson, D. W., Folkerth, R. D., Keene, C. D., Litvan, I., Perl, D.P, Stein, T.D., Vonsattel, J.P., Stewart, W., Tripodis, T., Crary, J. F., Bienek, K.F., Dams-O'Connor, K., Alvarez, V.E., Gordon, W.A., TBI/CTE group (2016). The first NINDS/NIBIB consensus meeting to define neuropathological criteria for the diagnosis of chronic traumatic encephalopathy. *Acta Neuropathologica, 131*(1), 75–86.

Mullally, W. J. (2017). Concussion. *The American Journal of Medicine, 130*(8), 885-892.

Nakayama, Y., Covassin, T., Schatz, P, Nogle, S., & Kovan, J. (2014). Examination of the test-retest reliability of a computerized neurocognitive test battery. *American Journal of Sports Medicine, 42*(8), 2000-2005.

Norheim N., Kissinger-Knox A., Cheatham M., Webbe, F. (2018). Performance of college athletes on the 10-item word list of SCAT5. *BMJ Open Sport & Exercise Medicine, 4*, e000412. doi:10.1136/bmjsem-2018-000412

Poreh, A. (2005). Analysis of mean learning of normal participants on the rey auditory-verbal learning test. *Psychological Assessment, 17*(2), 191-199. doi:http://dx.doi.org.portal.lib.fit.edu/10.1037/1040-3590.17.2.191

Proto, D. A., Pastorek, N. J., Miller, B. I., Romesser, J. M., Sim, A. H., & Linck, J. F. (2014). The dangers of failing one or more performance validity tests in individuals claiming mild traumatic brain injury-related postconcussive symptoms. *Archives of Clinical Neuropsychology, 29*(7), 614-624.

Ruchinska, R. (2019). Wechsler Adult Intelligence Scale-4th edition digit span performance in subjective cognitive complaints, amnestic mild cognitive impairment, and probable dementia of the Alzheimer type. *The Clinical Neuropsychologist, 33*(8), 1436-1444.

Schatz, P. (2010). Long-term test retest reliability of baseline cognitive assessments using ImPACT. *American Journal of Sports Medicine, 38*(1). 47-53. DOI: 10.1177/0363546509343805

Schatz, P., & Ferris, C.S. (2013). One-month test-retest reliability of the ImPACT test battery. *Archives of Clinical Neuropsychology, 28*(5), 499-504. doi: 10.1093/arclin/act034.

Schatz, P., & Glatts, C. (2013). "Sandbagging" baseline test performance on ImPACT, without detection, is more difficult than it appears. *Archives of Clinical Neuropsychology*, *28*(3), 236-244.

Schatz, P., & Putz, B.O. (2006). Cross-validation of measures used for computer-based assessment of concussion. *Applied Neuropsychology, 13*(3), 151-159. DOI: 10.1207/s15324826an1303_2

Schatz, P., & Robertshaw, S. (2014). Comparing post-concussive neurocognitive test data to normative data presents risks for under-classifying "above average" athletes. *Archives of Clinical Neuropsychology, 29*(7), 625-632.

Schatz, P., Moser, R. S., Solomon, G. S., Ott, S. D., & Karpf, R. (2012). Prevalence of invalid computerized baseline neurocognitive test results in high school and collegiate athletes. *Journal of Athletic Training*, *47*(3), 289-296.

Strauss, E., Sherman, E., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford University Press. 617-626.

Appendix

OFFICE OR OFF-FIELD ASSESSMENT

Please note that the neurocognitive assessment should be done in a distraction-free environment with the athlete in a resting state.

STEP 1: ATHLETE BACKGROUND

Sport / team / school: _____

~~Date / time of injury:~~ _____

Years of education completed: _____

Age: _____

Gender: M / F / Other

Dominant hand: left / neither / right

How many diagnosed concussions has the athlete had in the past?: _____

When was the most recent concussion?: _____

How long was the recovery (time to being cleared to play) from the most recent concussion?: _____ (days)

Has the athlete ever been:

Hospitalized for a head injury?	Yes	No
Diagnosed / treated for headache disorder or migraines?	Yes	No
Diagnosed with a learning disability / dyslexia?	Yes	No
Diagnosed with ADD / ADHD?	Yes	No
Diagnosed with depression, anxiety or other psychiatric disorder?	Yes	No

Current medications? If yes, please list:

_____

_____

_____

_____

_____

_____

Name: _____

DOB: _____

Examiner: _____

Date: _____

2

STEP 2: SYMPTOM EVALUATION

The athlete should be given the symptom form and asked to read this instruction paragraph out loud then complete the symptom scale. For the baseline assessment, the athlete should rate his/her symptoms based on how he/she typically feels and for the post injury assessment the athlete should rate their symptoms at this point in time.

Please Check: ☐ **Baseline** ☐ **Post-Injury**

Please hand the form to the athlete

	none	mild		moderate		severe	
Headache	0	1	2	3	4	5	6
"Pressure in head"	0	1	2	3	4	5	6
Neck Pain	0	1	2	3	4	5	6
Nausea or vomiting	0	1	2	3	4	5	6
Dizziness	0	1	2	3	4	5	6
Blurred vision	0	1	2	3	4	5	6
Balance problems	0	1	2	3	4	5	6
Sensitivity to light	0	1	2	3	4	5	6
Sensitivity to noise	0	1	2	3	4	5	6
Feeling slowed down	0	1	2	3	4	5	6
Feeling like "in a fog"	0	1	2	3	4	5	6
"Don't feel right"	0	1	2	3	4	5	6
Difficulty concentrating	0	1	2	3	4	5	6
Difficulty remembering	0	1	2	3	4	5	6
Fatigue or low energy	0	1	2	3	4	5	6
Confusion	0	1	2	3	4	5	6
Drowsiness	0	1	2	3	4	5	6
More emotional	0	1	2	3	4	5	6
Irritability	0	1	2	3	4	5	6
Sadness	0	1	2	3	4	5	6
Nervous or Anxious	0	1	2	3	4	5	6
Trouble falling asleep (if applicable)	0	1	2	3	4	5	6
Total number of symptoms:							of 22
Symptom severity score:							of 132
Do your symptoms get worse with physical activity?						Y	N
Do your symptoms get worse with mental activity?						Y	N
If 100% is feeling perfectly normal, what percent of normal do you feel?							

If not 100%, why?

_____

_____

_____

Please hand form back to examiner

3

STEP 3: COGNITIVE SCREENING

Standardised Assessment of Concussion (SAC)*

Name: _____

DOB: _____

Examiner: _____

Date: _____

ORIENTATION

What month is it?	0	1
What is the date today?	0	1
What is the day of the week?	0	1
What year is it?	0	1
What time is it right now? (within 1 hour)	0	1
Orientation score		of 5

IMMEDIATE MEMORY

The Immediate Memory component can be completed using the traditional 5-word per trial list or optionally using 10-words per trial to minimise any ceiling effect. All 3 trials must be administered irrespective of the number correct on the first trial. Administer at the rate of one word per second.

I am going to test your memory. I will read you a list of words and when I am done, repeat back as many words as you can remember, in any order. For Trials 2 & 3: I am going to repeat the same list again. Repeat back as many words as you can remember in any order, even if you said the word before.

List G	Trial 1	Trial 2	Trial 3	Alternative List (H)	Alternative List (I)
Finger	0 1	0 1	0 1	Baby	Jacket
Penny	0 1	0 1	0 1	Monkey	Arrow
Blanket	0 1	0 1	0 1	Perfume	Pepper
Lemon	0 1	0 1	0 1	Sunset	Cotton
Insect	0 1	0 1	0 1	Iron	Movie
Candle	0 1	0 1	0 1	Elbow	Dollar
Paper	0 1	0 1	0 1	Apple	Honey
Sugar	0 1	0 1	0 1	Carpet	Mirror
Sandwich	0 1	0 1	0 1	Saddle	Saddle
Wagon	0 1	0 1	0 1	Bubble	Anchor
Trial Total					
				Immediate memory score total:	/30

CONCENTRATION

DIGITS BACKWARDS

Please circle the Digit list chosen (A, B, C, D, E, F). Administer at the rate of one digit per second reading DOWN the selected column.

I am going to read a string of numbers and when I am done, you repeat them back to me in reverse order of how I read them to you. For example, if I say 7-1-9, you would say 9-1-7.

Item	List A	Trial 1	Administer Trial 2 only if zero on Trial 1	Trial 2
1	4-9-3	0 1	6-2-9	0 1
2	3-8-1-4	0 1	3-2-7-9	0 1
3	6-2-9-7-1	0 1	1-5-2-8-6	0 1
4	7-1-8-4-6-2	0 1	5-3-9-1-4-8	0 1
	Trial Totals			

Item	List B	Trial 1	Administer Trial 2 only if zero on Trial 1	Trial 2
1	5-2-6	0 1	4-1-5	0 1
2	1-7-9-5	0 1	4-9-6-8	0 1
3	4-8-5-2-7	0 1	6-1-8-4-3	0 1
4	8-3-1-9-6-4	0 1	7-2-4-8-5-6	0 1
	Trial Totals			

Item	List C	Trial 1	Administer Trial 2 only if zero on Trial 1	Trial 2
1	1-4-2	0 1	6-5-8	0 1
2	6-8-3-1	0 1	3-4-8-1	0 1
3	4-9-1-5-3	0 1	6-8-2-5-1	0 1
4	3-7-6-5-1-9	0 1	9-2-6-5-1-4	0 1
	Trial Totals			

Digit Score Total:	/4

MONTHS IN REVERSE ORDER

Now tell me the months of the year in reverse order. Start with the last month and go backward. So you'll say December, November. Go ahead.

Dec - Nov - Oct - Sept - Aug - Jul - Jun - May - Apr - Mar - Feb - Jan	0 1
Months Score	of 1
Concentration Total Score (Digits + Months)	of 5

Davis GA, *et al. Br J Sports Med* 2017;**0**:1–8. doi:10.1136/bjsports-2017-097506SCAT5

4

STEP 4: NEUROLOGICAL SCREEN

See the instruction sheet (page 7) for details of test administration and scoring of the tests.

Can the patient read aloud (e.g. symptom check-list) and follow instructions without difficulty?	Y	N
Without moving their head or neck, can the patient look side-to-side and up-and-down without double vision?	Y	N
Can the patient perform the finger nose coordination test normally?	Y	N
Can the patient perform tandem gait normally?	Y	N

BALANCE EXAMINATION

Modified Balance Error Scoring System (mBESS) testing[6]

Which foot was tested (i.e. which is the non-dominant foot) ☐ Left ☐ Right

Testing surface (hard floor, field, etc.) _____

Footwear (shoes, barefoot, braces, tape, etc.) _____

Condition	Errors
Double leg stance	of 10
Single leg stance (non-dominant foot)	of 10
Tandem stance (non-dominant foot at the back)	of 10
Total Errors	of 30

Name: _____

DOB: _____

Examiner: _____

Date: _____

5

STEP 5: DELAYED RECALL:

The delayed recall should be performed after 5 minutes have elapsed since the end of the Immediate Recall section. Score 1 pt. for each correct response.

Do you remember that list of words I read a few times earlier? Tell me as many words from the list as you can remember in any order.

Time Started _____

Please record each word correctly recalled. Total score equals number of words recalled.

_____

_____

_____

Total number of words recalled accurately: _____ of 10

6

STEP 6: DECISION

	Date & time of assessment:		
Domain			
Symptom number (of 22)			
Symptom severity score (of 132)			
Orientation (of 5)			
Immediate memory	of 15 of 30	of 15 of 30	of 15 of 30
Concentration (of 5)			
Neuro exam	Normal Abnormal	Normal Abnormal	Normal Abnormal
Balance errors (of 30)			
Delayed Recall	of 5 of 10	of 5 of 10	of 5 of 10

SCORING ON THE SCAT5 SHOULD NOT BE USED AS A STAND-ALONE METHOD TO DIAGNOSE CONCUSSION, MEASURE RECOVERY OR MAKE DECISIONS ABOUT AN ATHLETE'S READINESS TO RETURN TO COMPETITION AFTER CONCUSSION.