2 Neuropsychological Testing in Sports Concussion Management: Test of an Evidence-Based Model When Baseline Is Unavailable

Peter Arnett, Jessica Meyer, Victoria Merritt, Erin Guty, Kaitlin Riegler, and Garrett Thomas

Introduction

Barth and colleagues' [1] seminal study using baseline neuropsychological testing as a model for sports concussion management set a standard that continues to be influential today. Many school-based sports medicine programs have adopted variations of their approach, and a range of recommendations have been made for the use of neuropsychological testing within that framework. Although the literature is variable regarding how best to use neuropsychological testing, most investigators recommend the use of pre-injury baseline neuropsychological testing as the best practice for sports concussion management [1–7]. Still, baseline data are not always available, and there is recognition that guidelines are needed for interpretation in such cases. In their "Consensus Statement on Concussion in Sport" article, McCrory and colleagues [6] suggested that an important area for future research was determining "best-practice" neuropsychological testing in cases where baseline data are not available. Also, in a position paper published under the aegis of the National Academy of Neuropsychology (NAN), Moser et al. [3] noted that neurocognitive tests can play a meaningful role in concussion management even in the absence of baseline testing. Nonetheless, neither article provides guidelines for how neuropsychological tests should be used when no baseline testing has been conducted. The most recent consensus statement published by McCrory and colleagues [8] indicated that baseline or preseason neuropsychological testing should not be mandatory, but also stating that it could be helpful in some situations. The consensus panel also indicated that conducting neuropsychological testing post concussion would be optimal.

The central goal of this chapter is to provide a test of the evidence-based model for using neuropsychological testing in the management of sports-related concussion when no baseline is available that we laid out in our chapter in the first edition

P. Arnett (✉) · J. Meyer · V. Merritt · E. Guty · K. Riegler · G. Thomas
Department of Psychology, Penn State University, University Park, PA, USA
e-mail: paa6@psu.edu

S. M. Slobounov, W. J. Sebastianelli (eds.), *Concussions in Athletics*,
https://doi.org/10.1007/978-3-030-75564-5_2

of this book. We first summarize and evaluate existing approaches, focusing on the merits and limitations of baseline testing, the timing of testing post concussion, and the additional value of neuropsychological tests in a sports concussion context. We then lay out the framework of our model and provide a test of it using five cognitive outcome variables not included in the algorithm itself by comparing "Recovered" and "Not Recovered" groups based upon the algorithm. It is not our intent to suggest that the model presented in this chapter should replace the baseline model. Furthermore, a discussion of the case for or against the use of neurocognitive testing in a sports concussion framework goes well beyond the scope of this chapter and has been discussed at length by other investigators [9, 10]. However, we do touch upon the merits and limitations of such tests, as well as the pros and cons of conducting baseline testing.

Summary of Literature Recommendations for the Use of Neuropsychological Testing in Sports Concussion

Use of Baseline Testing

Although the literature is variable regarding how best to use neuropsychological testing, many investigators recommend the use of pre-injury baseline neuropsychological testing as best practice for sports concussion management [1–7]. As Guskiewicz et al. [2] and Echemendia and colleagues [11] have articulated, the use of baseline testing for comparison with post-injury scores helps to control for idiosyncratic interindividual differences at baseline (e.g., ADHD, possible cumulative cognitive impact of prior concussions, cultural/linguistic differences, learning disorders, age, education, and proneness to psychiatric issues). Controlling for such extraneous factors by using baseline testing should make neuropsychological tests more sensitive to the impact of concussions on specific individuals.

Still, the baseline paradigm for sports concussion is not without limitations. It has been criticized because there is no empirical evidence that the use of baseline testing improves diagnostic accuracy [9, 12], reduces risk of further injury [10], or predicts decline better than would be expected by chance alone [11].

Another significant limitation of the baseline model is that, for the types of intervals often used in sports concussion testing, the test-retest reliability is not known for many typically used neuropsychological tests [10, 13–15]. Furthermore, the time between baseline and post-injury intervals can be years apart, whereas test-retest reliabilities are typically assessed over about 4–8 week intervals. Finally, commonly used neurocognitive tests in sports concussion often have less than optimal test-retest reliabilities for clinical decision-making [13, 15].

Consideration of test-retest reliability coefficients is critical because they are central to calculating the reliable change indices (RCIs) that are typically used to determine clinically significant change. If these reliability coefficients are low, then confidence intervals will be large and greater declines will be required post concussion for change to be detected. Tests with low test-retest reliability

coefficients, then, will be less sensitive to changes post concussion than those with higher values.

In practical terms, baseline testing is logistically complex and expensive. Also, practice effects are commonly seen with neuropsychological tests, something that can reduce sensitivity post concussion [16]. Overall, despite its utility in controlling for interindividual differences, the baseline model does have limitations. Given these considerations, using neuropsychological tests in the sports concussion framework when no baseline has been conducted should be considered.

Timing of Post-concussion Testing

There is no clear consensus on the timing of post-concussion neurocognitive testing. In Guskiewicz et al.'s [2] National Athletic Trainers' Association (NATA) Position Statement, the authors suggest that neurocognitive testing should ideally be conducted in the acute injury period to help determine the severity of the concussion, and then again when the athlete is symptom-free to help with return-to-play decisions. However, they do not provide any clear indication of when during the acute injury period that testing might ideally occur.

In the ImPACT Technical Manual [17] on the "Best Practices" page from the ImPACT website, the authors recommend post-concussion ImPACT testing 24–72 hours post concussion to assess whether declines have occurred from baseline and to help with concussion management in general. They also recommend testing after this acute period once the athlete is symptom free both at rest and with cognitive exertion.

The most recent consensus conference [8] recommended that neurocognitive testing be conducted when clinically indicated and when athletes are symptom free by their own self-report; however, these authors provided the caveat that some cases (especially children and adolescents) may warrant neurocognitive testing prior to symptom resolution. They reasoned that such testing could help with school and home management. A position statement published by the American Medical Society for Sports Medicine [18] was agnostic on this issue, asserting that the evidence was unclear regarding the optimal timing of post-concussion neuropsychological testing. In sum, the available literature indicates that there is no clear consensus on the timing of neuropsychological testing post concussion.

The "Value-Added" of Neuropsychological Tests in a Sports Concussion Framework

Some investigators have argued that there is no "value added" to neuropsychological testing in the management of sports concussion, and that return-to-play decisions should strictly be based upon athletes' self-reported symptoms [9, 10]. However, research on this topic has revealed two important findings that counter such a recommendation: (1) A significant percentage of concussed athletes who

report full symptom resolution still show objective neurocognitive deficits—either declines from baseline [19] or when no baseline is available, worse neurocognitive performance than control participants [20]; and (2) neurocognitive tests can identify concussed athletes in the acute post-concussion period (within 2 days post concussion) who deny any symptoms but show objective declines from baseline [7].

Although the additional value of neurocognitive tests to the concussion management process is controversial, beyond such considerations there are problems with relying exclusively on self-report of cognitive functioning in guiding return-to-play decisions. First, athletes have a high motivation to minimize symptoms following concussion because of their desire to return to play (RTP), a process articulated in Echemendia and Cantu's [21] "Dynamic Model for Return-to-Play Decision Making." Second, there is extensive literature demonstrating that self-reports of cognitive functioning are only weakly correlated with actual performance on objective cognitive tests, even in individuals who are motivated and who have not experienced any injury to the brain [22].

Harmon and colleagues [18] argue that there are at least three circumstances where post-concussion neurocognitive testing may be warranted: (1) In situations where athletes are presumed to be at high-risk because of prior concussion; (2) with athletes who are likely to minimize or deny symptoms so that they can RTP; and (3) to identify athletes with persistent deficits. Thus, these authors appear to recommend post-concussion neurocognitive testing under limited circumstances. One problem with only administering neurocognitive tests to athletes who are likely to minimize or deny symptoms is that such individuals can only be definitively identified if neurocognitive testing is conducted. Otherwise, how does one know? A limitation of only administering tests to identify athletes with persistent deficits is that, again, how does one know if athletes have "persistent deficits" if they are not actually tested? As indicated above, self-report of symptoms is suspect for a variety of well-established reasons, so relying on an athlete's self-report of symptoms is not going to be useful in identifying persistent deficits.

A Proposed Evidence-Based Model for Neurocognitive Concussion Management When No Baseline Is Available

Following Ellemberg and colleagues' [14] observation that the absence of scientifically validated algorithms for neuropsychological test interpretation has resulted in clinicians and researchers using idiosyncratic decision rules, as well as McCrory et al.'s [8] recommendation for "Best-Practice" guidelines, we articulate a model for the use of neuropsychological tests in a sports concussion framework when no baseline is available. We then provide an empirical test of this model.

Figure 2.1 illustrates our algorithm. Before going into the details of this, we outline the tests in the battery on which the algorithm is based, which includes both computerized and paper-and-pencil tests. We then describe the evidence basis for each step of the algorithm. Note that there are separate decision rules for males and females. This is due to findings of sex differences

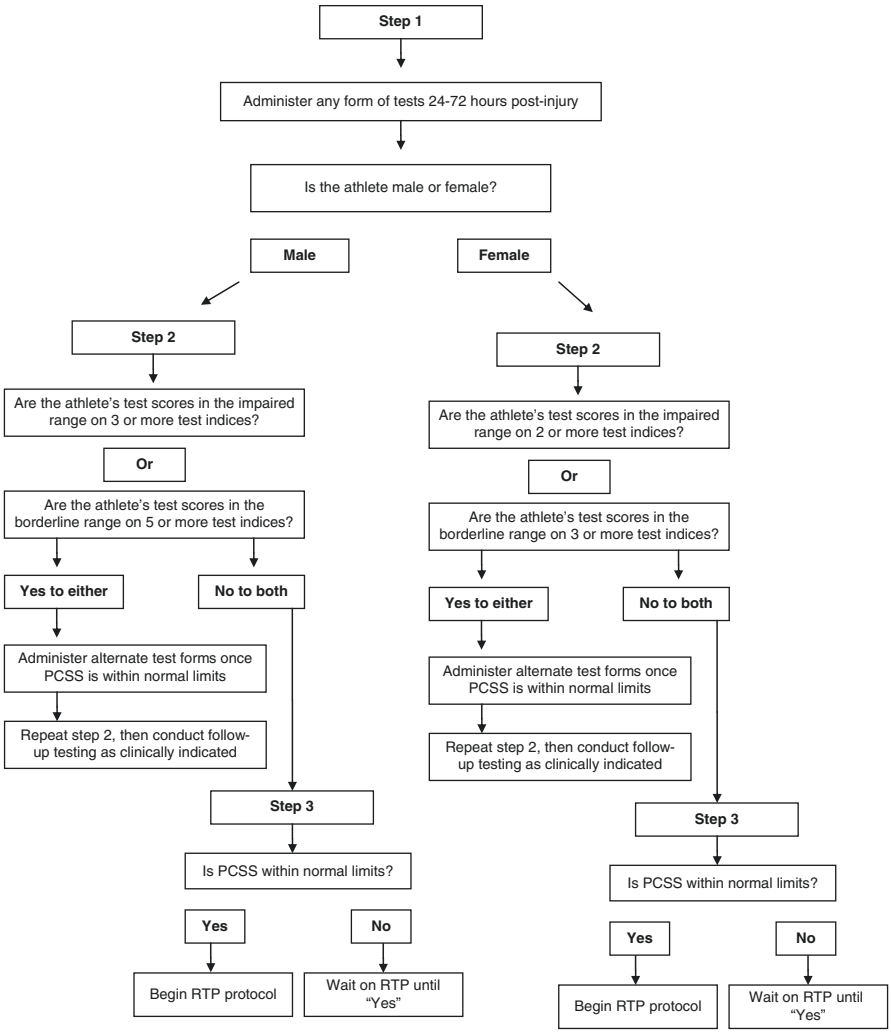

Fig. 2.1 Post-concussion neuropsychological testing algorithm when no baseline is available

in base rates of impairment using this same battery in Division I collegiate athletes [23].

Measures

The battery we use as the basis for our model includes both computerized and paper-and-pencil measures. Although the use of paper-and-pencil measures can be logistically more complex and expensive than using computerized tests alone because they require face-to-face administration, including such tests is likely to

increase the sensitivity of the battery. Also, if neuropsychological tests are only used post concussion, then the cost of administration is considerably lower.

Computerized tests Computerized tests include the ImPACT [24] and the Vigil Continuous Performance Test (CPT) [25]. The following summary indices from the ImPACT are included: Verbal Memory Composite, Visual Memory Composite, Visuomotor Speed Composite, and Reaction Time Composite. Although more recent versions of the ImPACT are available, we based our algorithm on the 2.0 version because of the availability of data for our evidence-based model. This version appears to be highly correlated with more recent (including online) versions of the ImPACT. Average Delay (a reaction time index) is used for the Vigil.

Paper-and-pencil tests These measures include: the Hopkins Verbal Learning Test-Revised (HVLT-R) [26] (total correct immediate and delayed recall), the Brief Visuospatial Memory Test-Revised (BVMT-R) [27] (total correct immediate and delayed recall), the Symbol-Digit Modalities Test (SDMT) [28] (total correct within 90 seconds), a modified Digit Span Test [29] (total correct forward and backward sequences), the PSU Cancellation Task [30] (total correct within 90 seconds), Comprehensive Trail Making Test Trails 2 and 4 or 3 and 5 (CTMT) [31] (completion times for both parts), and the Stroop Color-Word Test (SCWT) [32] (time to completion for both Color-Naming and Color-Word conditions). Thus, across computerized and paper-and-pencil measures, there are 17 test indices.

When multiple assessments are completed, we suggest that alternate forms be used at each repeated administration of the test battery. The ImPACT has such alternate forms built into the program; alternate forms are available for all of the above paper-and-pencil tests with the exception of the modified Digit Span Test and Stroop Color-Word Test.

Self-report To measure post-concussion symptoms, we use the Post-Concussion Symptom Scale (PCSS). This measure includes a list of 22 common post-concussion symptoms. Examinees rate the extent to which they are currently experiencing each symptom on a scale from 0 to 6, with 0 indicating the absence of the symptom, and 6 indicating severe symptoms.

Algorithm of Decision Rules

As Fig. 2.1 shows, each step of the algorithm after the initial neuropsychological testing involves a question, and then an action depending on the answer to the question.

Step 1 The action at Step 1 is to administer the test battery at 24–72 hours post injury. The evidence basis for this stems from animal models showing that many elements of the neurochemical cascade in the brain following concussion peak at about 48 hours post injury, and the decrease in glucose metabolism that occurs at

about 48 hours post injury, are correlated with cognitive dysfunction in adult rats [33–35]. Also, neurocognitive research in humans has shown that the greatest cognitive impact post concussion typically occurs within 24–72 hours post injury [1, 16, 36, 37], though there is considerable individual variability [37]. As such, testing athletes during this time interval should provide a likely estimate of the full impact of the concussion on the brain as manifested by neurocognitive test results. Also, if the athlete is free of neurocognitive impairment at this early stage (relative to base rates), then no further neurocognitive testing would necessarily need to be conducted post concussion, and the return-to-play (RTP) decision could be made based on other factors (e.g., self-reported symptoms, vestibular signs, etc.). If the athlete does show signs of neurocognitive impairment at this point, then the objective neurocognitive data could be used to assist in getting temporary academic accommodations while symptomatic (e.g., deferral of exams and other assignments, testing in a room free from distraction, extra time on exams, etc.). A more detailed rationale for testing at this early time point post concussion, and possibly before self-reported symptom resolution, is provided below in the section entitled, "Why Recommend Testing During the Acute Concussion Phase?"

Step 2 The algorithm has different Step 2s for males and females because the study on which these specific decision rules are based revealed slightly different base rates for males and females. In this study, we examined baseline performance in 495 collegiate athletes on the same test battery outlined in this article [23], and impairment on a test was defined as performing 2 SDs or more below the mean of other athletes; borderline impairment was defined as 1.5 SDs or more below the mean. These criteria were used since currently there is no agreed upon definition of abnormally poor test performance on neuropsychological tests following concussion, and also to allow for some flexibility in decision-making.

In this study, less than 10% of males had five or more borderline scores, and less than 10% of females had three or more borderline scores. Additionally, less than 10% of males had three or more impaired scores, and less than 10% of females had two or more impaired scores. We used these base rates as a foundation for the decision rules in our model. In light of such data, male athletes who are tested post concussion who show impairment on 3 or more tests and female athletes who show impairment on 2 or more tests evidence highly unusual performance that is likely to reflect the impact of their concussion (see Fig. 2.1). Similarly, male athletes who are tested post concussion who show borderline scores on 5 or more tests and female athletes who show borderline scores on 3 or more tests display highly unusual performance that is likely to reflect the impact of their concussion. The application of these data in decision rules is shown at Step 2 in Fig. 2.1. Ideally, concussion programs adopting this algorithm would be advised to use base rate of impairment data collected from athletes participating in their specific programs. In this way, the data used are likely to be most valid for that group of athletes for a particular neurocognitive test battery. If such base rates differ from what we report, relevant values could simply replace what we report from our athletes in the algorithm. If base rates of

impairment are not available, it should be noted that other studies using test batteries of comparable length have reported similar base rates of impairment using a similar number of test indices in healthy older adults [38, 39], as well as children and adolescents [40].

If male or female athletes receive a "yes" response at Step 2, for either the impaired or borderline criterion, then the action is to "Administer Alternate Test Forms Once PCSS is Within Normal Limits." The evidence basis for this stems from findings showing that even when athletes report that they are symptom free, many still show evidence for objective cognitive impairment [19]. Additionally, relying on self-report of cognitive functioning when determining when athletes can RTP may be inaccurate given the consistently replicated low correlation found between objective neurocognitive test performance and self-reported neurocognitive functioning [22]. Thus, any athlete should have to perform within normal limits neurocognitively prior to returning to play, and such decisions should not be based on self-reported cognitive functioning alone. Following this recommendation after a "yes" response, the algorithm indicates, "Repeat Step 2, Then Conduct Follow-Up Testing as Clinically Indicated."

Step 3 If either male or female athletes have a "no" response at Step 2, then the algorithm moves to Step 3 to consider the following question: "Is PCSS Within Normal Limits?" The determination of "within normal limits" is made using normative data from our sample of collegiate athletes at baseline on the PCSS. Scores falling within the broad average range (i.e., standard score of 80 or above) are considered "within normal limits." If the answer to this question is "yes," then the recommendation is to begin the Return-to-play (RTP) protocol. If the answer is "no," then the recommendation is to wait on starting the RTP until the PCSS is within normal limits.

One complicating issue involves cases where athletes have a "yes" response at Step 2 (meeting the below base rate impaired or borderline criterion), yet report being within normal limits in terms of their symptom report. Given that the recommendation following such an outcome is to "Administer Alternate Test Forms Once PCSS Is Within Normal Limits," how does one proceed? There are no clear evidence-based guidelines for how to proceed here in terms of the precise timing of the next post-concussion testing point. A broad guideline would be to recommend testing the athlete again between 5 and 10 days post concussion, given that many studies show that most collegiate athletes show full cognitive recovery by that point [1, 16, 30, 36, 41–43]. With that said, other research shows that some collegiate athletes do not recover within that window and take longer than 2 weeks for their neurocognitive functioning to normalize [43, 44]. Thus, more research will clearly be needed to refine this broad guideline. Studies that examine the duration for normalization of brain functioning in athletes who report being normal in terms of symptom report but show impairments neurocognitively would be ideal. Given the current state of the literature, the most prudent approach would be to rely more on individualistic clinical concussion management strategies employed by skilled

clinicians to determine temporal sequencing of testing in these cases [45]. Factors such as the urgency with which a return-to-play decision needs to be made (e.g., if a crucial game is imminent vs. the athlete's sport not being "in season"), as well as other individualistic factors (e.g., prior concussion history, the presence of clinically significant depression) would need to be considered. Thus, the model allows for considerable flexibility at this stage due, in part, to the absence of clear research evidence to guide decision-making, but also due to idiosyncratic factors that are nearly always going to be at play in the clinical management of concussion.

Why Recommend Testing During the Acute Concussion Phase?

One potentially controversial recommendation in our algorithm is to routinely test athletes in the acute stage more systematically post concussion. Many athletes are likely to still be experiencing some symptoms at the 24–72 hour post-concussion point, and some investigators and clinicians have asserted that such testing should be avoided on a number of grounds. First, given that athletes are still symptomatic, some posit that such testing cannot contribute anything to the RTP decision, because clinicians are typically not going to put athletes back to play who are still experiencing self-reported symptoms. Second, it has been suggested that such testing could exacerbate the athlete's symptoms, and there is some evidence to support this concern [46]. These are reasonable concerns; however, we assert that the value of such acute testing outweighs the potential minor risk of a temporary increase in symptoms. The caveat to this, of course, involves cases where symptoms are so severe that testing could be harmful in exacerbating already severe symptoms, or where the nature of such symptoms would likely substantially interfere with test performance (e.g., severe dizziness, nausea, or headache, among others). This is where individualistic concussion management again becomes important [45].

One benefit of such testing in the early acute phase is to help document the severity of the concussion. Athletes who show more impairments at this acute stage could be managed more conservatively once RTP procedures have begun than those who were back to their likely premorbid cognitive level, or nearly back to such a level. Another benefit, as noted earlier, is that early objective documentation of deficits could result in athletes quickly being able to secure needed academic accommodations during their recovery period. A third benefit of acute testing is that it may show that the athlete is in fact back to baseline neurocognitively, even at this early stage. If this is the case, then more rapid RTP could potentially occur. Although an athlete's medical well-being must always be the most important consideration of sports medicine professionals, athletes performing at a high level of sport (e.g., Division I college, the basis of our algorithm) could suffer significant harm in terms of their status on the team and ability to compete in important games and maintain their scholarships, as well as their mental health, if they are held out of play for an unnecessarily long period of time.

A final benefit of conducting systematic testing during this acute period post concussion and at other systematic time points is that the neurocognitive results

following any future concussion could be compared with the results from the previous concussion to assess whether the range and severity of cognitive impairments increases. If an athlete is tested at different points following different concussions, then such systematic comparisons would not be possible. Athletes who suffer multiple concussions and show an increased range and severity of cognitive impairments with each successive concussion can then be treated more conservatively.

Limitations

Our algorithm represents an initial attempt to develop systematic guidelines for decision-making post concussion in cases where baseline data are not available. Although we provide systematic decision rules, there is much room for individualistic concussion management, and we spell out a number of examples where such factors come in to play. The neuropsychological test battery we recommend is relatively lengthy and logistically complex; however, applying it in cases where baseline testing has not been conducted significantly reduces such complexity. Also, the algorithm can be adapted to different (possibly shorter) test batteries and different athlete groups when base rates of impairment data can be derived from such groups.

Testing the Model

In this section, we provide a preliminary test of our model in a group of concussed collegiate athletes. We aimed to test the validity of the model by comparing algorithm-determined "Recovered/Not Recovered" groups of concussed collegiate athletes along a number of dimensions. Our primary hypothesis concerned examining post-injury cognitive tests between these groups that we conceptualized as other indicators of recovery beyond the number of cognitive impairments used for the algorithm.

Primary Hypothesis Compared with the Not Recovered group, the Recovered group will show significantly better performance on indices from five cognitive test indices not included in the algorithm itself when tested post concussion: Immediate and Delayed Recall on the Story subtest from the Rivermead Behavioral Memory Test (RBMT); Immediate and Delayed Recall from the Affective Word List (AWL); and the Controlled Oral Word Association Test (COWAT).

Method

Participants

The sample was derived from athletes who were referred to us for evaluation between 2002 and 2019 at our Division I university. All participants were referred by either an athletic trainer or team physician for concussion testing after sustaining a sports-related concussion (SRC).

Recovered Versus Not Recovered Participants Using Base Rate of Impairment at Baseline Algorithm

In terms of dichotomizing groups using this algorithm, participants were included in the "Not Recovered" group if they met criteria for either the "borderline" or "impaired" algorithm (see Fig. 2.1). Participants were included in the "Recovered" group if they fell below the threshold for both the "borderline" and "impaired" algorithm. As shown in Table 2.1, this resulted in a total of 96 (80%) "Recovered" athletes and 24 (20%) "Not Recovered" athletes.

Measures

Post-Concussion Cognitive Outcome Variables Not Included in the Algorithm

The RBMT—Story Memory involves examinees having to recall two different stories read to them. They are asked to recall these immediately and then after about a 30-minute delay. The Affective Word List (AWL) [47] has a similar format as a traditional list-learning task. The examinee is read three trials of a list of 16 words, 8 positively valenced and 8 negatively valenced, and is then asked to recall as many of the words as possible. After a 20- to 25-min delay, examinees are again asked to recall as many of the words as they can remember. The AWL was developed to measure affective bias, specifically to assess the proportion of positive versus negative words recalled as a way of detecting depressive tendencies through a performance-based task. However, we have also found that it can be used as a traditional list-learning task, and it has been shown to have good convergent and discriminant validity as a memory test [48]. The COWAT is a speeded measure of verbal fluency that requires examinees to generate as many words as they can that begin with a particular letter of the alphabet. Total words generated across three letter trials was used as the dependent variable for the COWAT.

Results

Preliminary Analyses: Recovered Versus Not Recovered Participants Using Combined Algorithm

As shown in Table 2.1, the groups did not differ significantly in age or days since concussion. However, the groups did differ in terms of sex distribution. Specifically, a much higher proportion of females comprised the Not Recovered group (45.8%) compared with the Recovered (20.3%) group, X^2 (N = 120) = 11.37, p = 0.001. Given these differences, we conducted some post-hoc analyses (see below) to try and understand these differences further.

The groups were also compared on the number of borderline and impaired scores. Not surprisingly, given the selection criteria for the groups, the Not Recovered group had significantly more borderline scores (mean = 6.58 (3.41),

Table 2.1 Participant demographic and impairment variables for recovered versus not recovered groups using combined algorithm criteria

	Recovered (n = 96)		Not recovered (n = 24)			
Continuous variables	*M*	*SD*	*M*	*SD*	*t-value*	*p*
Age (years)	18.96	1.27	18.83	1.27	−0.43	0.67
Days since concussion	20.56	51.38	26.79	45.22	0.54	0.59
Impaired scores[1]	0.44	0.65	4.58	2.98	6.78[3]	<0.001
Borderline scores[2]	0.88	1.00	6.58	3.41	8.11[3]	<0.001
Categorical variables	*N*	%	*N*	%	X^2	*p*
Sex					11.37	=0.001
Male	82	85.4	13	54.2		
Female	14	14.6	11	45.8		
Ethnicity						
Caucasian	109	77.9	17	53.1		
African American	22	15.7	11	34.4		
Asian American	4	2.9	0	0.0		
Biracial/multiracial	3	2.1	4	12.5		
Other	2	1.4	0	0.0		
Sport						
Baseball	3	3.1	0	0.0		
Cheerleading	1	1.0	0	0.0		
Football	25	26.0	9	37.5		
Golf	2	2.1	0	0.0		
Men's basketball	17	17.7	0	0.0		
Men's ice hockey	6	6.3	1	4.2		
Men's lacrosse	16	16.7	1	4.2		
Men's soccer	2	2.1	1	4.2		
Rugby	12	12.5	0	0.0		
Softball	1	1.0	1	4.2		
Swimming/diving	1	1.0	0	0.0		
Women's basketball	1	1.0	4	16.7		
Women's ice hockey	0	0.0	2	8.3		
Women's lacrosse	2	2.1	2	8.3		
Women's soccer	1	1.0	2	8.3		
Wrestling	6	6.3	0	0.0		
Other	0	0.0	1	4.2		

Note: *M* = Mean, *SD* = Standard deviation. [1] Number of impaired scores reflects the number of scores out of the 17 indices from the algorithm that are 2 SD or more below the mean of 100 (i.e., at or below a standard score of 70). [2] Number of borderline impaired scores reflects the number of scores out of the 17 indices from the algorithm that are 1.5 SD or more below the mean of 100 (i.e., at or below a standard score of 78). [3] Degrees of freedom for this t-score = 1, 24, because of significant Levene's Test for Equality of Variances for this variable

median = 6.50) compared with the Recovered group (mean = 0.88 (1.00), median = 1.00), t (1, 24)[1] = 8.11, $p < 001$. The Not Recovered group also had significantly more impaired scores (mean = 4.58 (2.98), median = 4.00) compared with the Recovered group (mean = 0.44 (0.65), median = 0.00), t (1, 24)[1] = 6.78, $p < 001$.

[1] The dfs are lower here because of an adjustment in the degrees of freedom due to the fact that Levene's test for Equality of Variances between the groups was significant.

Thus, using the algorithm to define these groups led to a clear demarcation in terms of the relative cognitive difficulties of the groups post concussion.

Hypothesis Testing Analyses on Recovered Versus Not Recovered Participants Using Combined Algorithm

Primary Hypothesis: A Multivariate ANOVA (MANOVA) was conducted comparing the Recovered and Not Recovered groups on all five cognitive indices. The Multivariate F was significant, indicating that the groups were different on the five factors overall, $F(5, 114) = 4.69$, $p = 0.001$, $\eta p^2 = 0.17$. Univariate tests showed that the groups were significantly different on the RBMT—Immediate Recall ($F(1, 118) = 4.89$, $p < 0.05$, $\eta_p^2 = 0.04$), RBMT—Delayed Recall ($F(1, 118) = 10.75$, $p = 0.001$, $\eta_p^2 = 0.08$), and AWL—Immediate Recall ($F(1, 118) = 12.98$, $p < 0.001$, $\eta_p^2 = 0.10$). There was a statistical trend for AWL—Delayed Recall ($F(1, 118) = 3.49$, $p < 0.10$, $\eta_p^2 = 0.03$), and the effect for the COWAT was not significant ($F(1, 118) < 1.0$, ns). Consistent with predictions, all effects were in the direction of the Not Recovered group having lower scores than the Recovered group (See Table 2.2).

Post Hoc Analyses

Unexpectedly, females comprised more than twice the percentage of the Not Recovered compared with the Recovered group. Using our base rate algorithm, fewer tests need to fall in the impaired or borderline range in females compared with males for selection in the Not Recovered group. Given this, we wanted to explore whether the males in our Not Recovered group had more impaired tests than the females. Although the males had more impairments (5.5 vs. 3.5) and more borderline impairments (7.6 vs. 5.4), the differences between groups were not significant, $t(1, 22) = 1.62$, ns, and $t [1, 20] = 1.67$, ns, respectively. Still, these mean differences were relatively large.

Table 2.2 Mean performance on key outcome variables for recovered versus not recovered groups using combined borderline and impaired criteria

Measure	Recovered		Not recovered				
	M	*SD*	*M*	*SD*	F-test	*p*	η_p^2
Cognitive Indices[1]							
Multivariate F (df = 5, 114)					4.69	<0.001	0.17
Univariate F (1, 170)							
RBMT—Immediate Recall	101.62	14.39	94.11	16.75	4.89	<0.05	0.04
RBMT—Delayed Recall	100.93	14.61	89.93	15.05	10.75	=0.001	0.08
AWL—Immediate Recall	26.20	5.62	21.54	5.83	12.98	<0.001	0.10
AWL—Delayed Recall	7.14	3.79	5.58	2.95	3.49	=0.06	0.03
COWAT	16.34	7.36	15.00	3.22	0.76	ns	0.01

Note: *M* = Mean, *SD* = Standard deviation. Cognitive indices include the RBMT—Story subtest Immediate and Delayed Recall, AWL—Immediate and Delayed Recall, and the COWAT. *RBMT* Rivermead Behavioral Memory Test, *AWL* Affective Word List, *COWAT* Controlled Oral Word Association Test. RBMT indices are in standard scores; AWL and COWAT indices are raw scores. η_p^2 effect size magnitudes: small = 0.01 to 0.05; medium = 0.06 to 0.13; large = 0.14 and above

Next, we explored whether females in the Not Recovered group had worse scores on the key cognitive outcome indices compared with males, such that the disproportionate number of females in this group might be driving the group difference with the Recovered group. A MANOVA was conducted with sex as the grouping variable and the five cognitive outcome indices as the dependent variables. The multivariate F was significant, F [5, 16] = 2.95, $p < 0.05$, $\eta_p^2 = 0.45$, with all of the univariate F values being significant (p at least <0.05) for all of the individual tests. Remarkably, all of the effects were in the direction of females performing better than males in the Not Recovered group.

We also examined whether the female athletes were different from males in terms of the number of days post concussion when they were tested, and the groups did not significantly differ, t [1, 20] = −0.28, ns. Finally, we examined whether there were motivational/effort differences between males and females in the Not Recovered group. Using the ImPACT Impulse Control Composite (ICC), as well as an Examiner Rating (scale of 1–7, with higher scores reflecting higher motivation), the groups were again not different, t [1, 20] = −0.53, ns, and t [1, 20] = −0.28, ns, respectively.

Discussion of Model Test

Our test of the model showed partial support. We found that the "Not Recovered" group performed significantly worse than the "Recovered" group on three of the five outcome variables: the RBMT Immediate and Delayed Recall, and the AWL Immediate Recall. One interesting finding from the study was that there were a highly disproportionate number of females in the Not Recovered group, over double the percentage found in the Recovered group. With follow up analyses, we found that females performed significantly better than males on all of the cognitive outcome variables. As such, it appears that the worse performance of the Not Recovered compared with the Recovered group was largely driven by poorer performance of *males* in the Not Recovered group. These differences may be due to the fact that, using our base rate algorithm, fewer tests need to fall in the impaired or borderline range in females compared with males for selection in the Not Recovered group. Although there were no significant differences in number of impairments between males and females in the Not Recovered group, the mean differences were fairly large. It seems likely that, given that more cognitive impairments are necessary for males to reach criteria for being "Not Recovered," their overall greater cognitive difficulties were manifested in our cognitive outcome variables, even though the latter were not part of the algorithm.

We conducted other follow-up tests and found that males did not differ from females in the Not Recovered group on two motivational indices (ICC from the ImPACT, and an Examiner Rating of observed effort on testing), so such factors could not explain the sex differences observed on the cognitive outcome variables.

Future Directions

Future work should include additional studies to validate the algorithm in other samples independent of our lab group. Ideally to extend what we have presented as a preliminary test in this chapter, groups of collegiate athletes with and without concussions could be tested at the same time intervals as suggested by our model. Examining base rates of impairment on the test battery in individuals with ADHD and/or learning disorders would also be a valuable focus for future work.

Our recommendations are tentative, given the limited evidence available for some aspects of the proposed algorithm. For example, several factors still need to be empirically established, including the ideal timing of post-concussion testing during the acute injury period, and the ideal temporal sequence of testing once athletes are normative symptomologically, but still impaired neurocognitively. However, we hope that our algorithm provides a template for improving neurocognitive concussion management in collegiate athletes.

Acknowledgments There are no conflicts of interest involved with this manuscript, and no sources of financial support.

References

1. Barth JT, Alves WM, Ryan TV, Macciocchi SN, Rimel RW, Jane JA, et al. Mild head injury in sports: neuropsychological sequelae and recovery of function. In: Levin HS, Eisenberg HM, Benton AL, editors. Mild head injury. New York: Oxford University Press; 1989. p. 257–75.
2. Guskiewicz KM, Bruce SL, Cantu RC, Ferrara MS, Kelly JP, McCrea M, et al. Recommendations on management of sport-related concussion: summary of the National Athletic Trainers' association position statement. Neurosurgery. 2004;55:891–5.
3. Moser RS, Iverson GL, Echemendia RJ, Lovell MR, Schatz P, Webbe FM, et al. NAN position paper: neuropsychological evaluation in the diagnosis and management of sports-related concussion. Arch Clin Neuropsychol. 2007;22:909–16.
4. Aubry M, Cantu R, Dvorak J, Graf-Baumann T, Johnston K, Kelly J, et al. Summary and agreement statement of the first International Conference on Concussion in Sport, Vienna 2001. Br J Sports Med. 2002;36:3–7.
5. McCrory P, Johnston K, Meeuwisse W, Aubry M, Cantu R, Dvorak J, et al. Summary and agreement statement of the 2nd International Conference on Concussion in Sport, Prague 2004. Br J Sports Med. 2005;39:196–204.
6. McCrory P, Meeuwisse W, Johnston K, Dvorak J, Aubry M, Molloy M, et al. Consensus Statement on Concussion in Sport – the 3rd International Conference on Concussion in Sport held in Zurich, November 2008. Br J Sports Med. 2009;43(Suppl 1):i76–90.
7. Van Kampen DA, Lovell MR, Pardini JE, Collins MW, Fu FH. The "value added" of neurocognitive testing after sports-related concussion. Am J Sports Med. 2006;30:1630–5.
8. McCrory P, Meeuwisse W, Dvořák J, Aubry M, Bailes J, Broglio S, et al. Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016. Br J Sports Med. 2017;51:838–47.
9. Randolph C, McCrea M, Barr WB. Is neuropsychological testing useful in the management of sport-related concussion? J Athl Train. 2005;40:139–54.
10. Randolph C. Baseline neuropsychological testing in managing sport-related concussion: does it modify risk? Curr Sports Med Rep. 2011;10(1):21–6.

11. Echemendia RJ, Bruce JM, Bailey CM, Sanders JF, Arnett PA, Vargas G. The utility of post-concussion neuropsychological data in identifying cognitive change following sports-related MTBI in the absence of baseline data. Clin Neuropsychol. 2012;26:1077–91.
12. Randolph C, Kirkwood MW. What are the real risks of sport-related concussion, and are they modifiable? J Int Neuropsychol Soc. 2009;15:1–9.
13. Broglio SP, Ferrara MS, Macciocchi SN, Baumgartner TA, Elliott R. Test–retest reliability of computerized concussion assessment programs. J Athl Train. 2007;42:509–14.
14. Ellemberg D, Henry LC, Macciocchi SN, Guskiewicz KM, Broglio SP. Advances in sport concussion assessment: from behavioral to brain imaging measures. J Neurotrauma. 2009;26:2365–82.
15. Mayers LB, Redick TS. Clinical utility of ImPACT assessment for postconcussion return-to-play counseling: psychometric issues. J Clin Exp Neuropsychol. 2012;34:235–42.
16. Rosenbaum AM, Arnett PA, Bailey CM, Echemendia RJ. Neuropsychological assessment of sports-related concussion: measuring clinically significant change. In: SSWS, editor. Foundations of sport-related brain injuries. Norwell, MA: Springer; 2006. p. 137–71.
17. Lovell M. ImPACT version 2.0 clinical user's manual. Pittsburgh, PA: ImPACT Applications Inc.; 2002.
18. Harmon KG, Drezner JA, Gammons M, Guskiewicz KM, Halstead M, Herring SA, et al. American Medical Society for Sports Medicine position statement: concussion in sport. Br J Sports Med. 2013;47:15–26.
19. Broglio SP, Macciocchi SN, Ferrara MS. Neurocognitive performance of concussed athletes when symptom free. J Athl Train. 2007;42:504–8.
20. Fazio VC, Lovell MR, Pardini JE, Collins MW. The relation between post concussion symptoms and neurocognitive performance in concussed athletes. NeuroRehabilitation. 2007;22:207–16.
21. Echemendia RJ, Cantu RC. Return to play following brain injury. In: Lovell MR, Echemendia RJ, Barth JT, Collins MW, editors. Traumatic brain injury in sports: an international neuropsychological perspective. Lisse, the Netherlands: Swets & Zeitlinger B.V.; 2004.
22. Lezak MD, Howieson DB, Loring DW. Neuropsychological assessment. 4th ed. New York: Oxford University Press; 2004.
23. Barwick FH, Rabinowitz AR, Arnett PA. Base rates of impaired neuropsychological test performance among healthy collegiate athletes. in revision.
24. Lovell M, Collins M, Podell K, Powell J, Maroon J. ImPACT: immediate post-concussion assessment and cognitive testing. Pittsburgh, PA: NeuroHealth Systems, LLC; 2000.
25. Cegalis JA, Cegalis S. The vigil/W continuous performance test (manual). New York: ForThought; 1994.
26. Benedict RHB, Schretlen D, Groninger L, Brandt J. Hopkins verbal learning TestñRevised: normative data and analysis of inter-form and test-retest reliability. Clin Neuropsychol. 1998;12(1):43–55.
27. Benedict RHB. Brief visuospatial memory test – revised: professional manual. Psychological Assessment Resources: Odessa, FL; 1997.
28. Smith A. Symbol digit modalities test (SDMT) manual (revised). Western Psychological Services: Los Angeles; 1982.
29. Wechsler D. Wechsler adult intelligence scale-III (WAIS-III). New York: Psychological Corporation; 1997.
30. Echemendia RJ, Putukian M, Mackin RS, Julian L, Shoss N. Neuropsychological test performance prior to and following sports-related mild traumatic brain injury. Clin J Sport Med. 2001;11:23–31.
31. Reynolds CR. Comprehensive trail making test (CTMT). Pro-Ed: Austin, TX; 2002.
32. Trenerry MR, Crosson B, DeBoe J, Leber WR. Stroop neuropsychological screening test. Psychological Assessment Resources: Odessa, FL; 1989.
33. Barkhoudarian G, Hovda DA, Giza CC. The molecular pathophysiology of concussive brain injury. Clin Sports Med. 2011;30:33–48.

34. Giza CC, DiFiori JP. Pathophysiology of sports-related concussion. Sports Health. 2011;3:46–51.
35. Giza CC, Hovda DA. The neurometabolic cascade of concussion. J Athl Train. 2001;36:228–35.
36. Belanger HG, Vanderploeg RD. The neuropsychological impact of sports-related concussion: a meta-analysis. J Int Neuropsychol Soc. 2005;11:345–57.
37. Wilde EA, McCauley SR, Barnes A, Wu TC, Chu Z, Hunter JV, et al. Serial measurement of memory and diffusion tensor imaging changes within the first week following uncomplicated mild traumatic brain injury. Brain Imaging Behav. 2012;6:319–28.
38. Brooks BL, Iverson GL, White T. Substantial risk of 'accidental MCI' in healthy older adults: base rates of low memory scores in neuropsychological assessment. J Int Neuropsychol Soc. 2007;13:490–500.
39. Palmer BW, Boone KB, Lesser IM, Wohl MA. Base rates of "impaired" neuropsychological test performance among healthy older adults. Arch Clin Neuropsychol. 1998;13:503–11.
40. Brooks BL, Sherman EMS, Iverson GL. Healthy children get low scores too: prevalence of low scores on the NEPSY-II in preschoolers, children, and adolescents. Arch Clin Neuropsychol. 2010;25:182–90.
41. Covassin T, Schatz P, Swanik CB. Sex differences in neuropsychological function and post-concussion symptoms of concussed collegiate athletes. Neurosurgery. 2007;61:345–51.
42. Covassin T, Elbin R, Harris W, Parker T, Kontos A. The role of age and sex in symptoms, neurocognitive performance, and postural stability in athletes after concussion. Am J Sports Med. 2012;40:1303–12.
43. Echemendia RJ, Iverson GL, McCrea M, Broshek DK, Gioia GA, Sautter SW, et al. Role of neuropsychologists in the evaluation and management of sport-related concussion: an inter-organization position statement. Arch Clin Neuropsychol. 2012;27:119–22.
44. McClincy MP, Lovell MR, Pardini J, Collins MW, Spore MK. Recovery from sports concussion in high school and collegiate athletes. Brain Inj. 2006;20:33–9.
45. Lovell M. The management of sports –related concussion: current status and future trends. Clin J Sport Med. 2009;28:95–111.
46. Meyer JE, Arnett PA. Changes in symptoms in concussed and non-concussed athletes following neuropsychological assessment. Dev Neuropsychol. 2015;40(1):24–8.
47. Ramanathan DN, Rabinowitz AR, Barwick FH, Arnett PA. Validity of affect measurements in evaluating symptom reporting in athletes. J Int Neuropsychol Soc. 2012;18:101–7.
48. Meyer J, Arnett PA. Validation of the affective word list as a measure of verbal learning and memory. J Clin Exper Neuropsychol. 2015;37:316–24.