Journal of the International Neuropsychological Society (2016), **22**, 1038–1049.

doi:10.1017/S1355617716000916

Applying an Evidence-Based Assessment Model to Identify Students at Risk for Perceived Academic Problems following Concussion

Danielle M. Ransom,[1] Alison R. Burns,[2,3] Eric A. Youngstrom,[4] Christopher G. Vaughan,[2,3] Maegan D. Sady,[2,3] and Gerard A. Gioia[2,3]

[1]University of Miami Miller School of Medicine, Miami, Florida
[2]Children's National Health System, Washington, DC
[3]George Washington University School of Medicine, Washington, DC
[4]University of North Carolina, Chapel Hill, North Carolina

(Received March 14, 2016; Final Revision October 10, 2016; Accepted October 10, 2016)

Abstract

Objectives: The aim of this study was to demonstrate the utility of an evidence-based assessment (EBA) model to establish a multimodal set of tools for identifying students at risk for perceived post-injury academic problems. **Methods:** Participants included 142 students diagnosed with concussion (age: $M = 14.95$; $SD = 1.80$; 59% male), evaluated within 4 weeks of injury (median = 16 days). Demographics, pre-injury history, self- and parent-report measures assessing symptom severity and executive functions, and cognitive test performance were examined as predictors of self-reported post-injury academic problems. **Results:** Latent class analysis categorized participants into "high" (44%) and "low" (56%) levels of self-reported academic problems. Receiver operating characteristic analyses revealed significant discriminative validity for self- and parent-reported symptom severity and executive dysfunction and self-reported exertional response for identifying students reporting low *versus* high academic problems. Parent-reported symptom ratings [area under the receiver operating characteristic curve (AUC) = .79] and executive dysfunction (AUC = .74), and self-reported ratings of executive dysfunction (AUC = .84), symptoms (AUC = .80), and exertional response (AUC = .70) each classified students significantly better than chance ($ps < .001$). Hierarchical logistic regression indicated that, of the above, self-reported symptoms and executive dysfunction accounted for the most variance in the prediction of self-reported academic problems. **Conclusions:** Post-concussion symptom severity and executive dysfunction significantly predict perceived post-injury academic problems. EBA modeling identified the strongest set of predictors of academic challenges, offering an important perspective in the management of concussion by applying traditional strengths of neuropsychological assessment to clinical decision making. (*JINS*, 2016, *22*, 1038–1049)

Keywords: Brain injury, Post-concussion symptoms, Pediatrics, Receiver operating characteristic, Academics, Evidence based practice

INTRODUCTION

Children recovering from concussion face a range of problems that can negatively impact school learning and performance during recovery (Halstead et al., 2013; Ransom et al., 2015). Concussion symptoms, such as headaches, fatigue, impaired concentration, and slowed processing speed, may temporarily impair the student's ability to maintain pace in their academic curriculum (Sady, Vaughan, & Gioia, 2011). In turn, efforts to engage in cognitively challenging tasks while still symptomatic can exacerbate concussion symptoms and impede academic activities, including problems taking notes, difficulty understanding new material, and requiring more time for assignments (Sady et al., 2011). A multimodal assessment battery, including post-concussion symptoms and cognitive performance, has been recommended as the most effective means to evaluate concussion (Echemendia et al., 2013; McCrory et al., 2013). It is a logical extension that this approach may have value in identifying factors related to academic impairment for the student. Empirical evidence currently does not exist, however, to guide clinicians as to which tools may best

Correspondence and reprint requests to: Gerard Gioia, Chief, Division of Neuropsychology, Children's National Medical Center, 15425 Shady Grove Road, Suite 350, Rockville, MD 20850. E-mail: ggioia@childrensnational.org

detect and predict who is at greatest risk for adverse academic outcomes during recovery.

An evidence-based assessment (EBA) approach can address this question by examining the various sources of information gathered in the clinical setting (e.g., personal risk factors and validated assessment measures) to derive the best set of predictors of academic outcomes (Youngstrom, 2013). The evidence-based practice movement has been a cornerstone of modern healthcare since it was introduced in 1992 (Evidence-Based Medicine Working Group, 1992), although its application to clinical neuropsychology is still in the early stages (Chelune, 2010). By integrating empirical research with clinical expertise, evidence-based standards aim to maximize prediction of outcomes and improve quality of life (Chelune, 2010; Straus, Glasziou, Richardson, & Haynes, 2011). EBA methodology is advantageous in clinical settings as it can streamline assessments to answer specific clinical questions applied to a given patient, while reducing cost and burden of extensive evaluations. Examining a set of clinical factors that are predictive of current status or current treatment needs allows for a rapid, evidence-based, decision to occur at the time of the appointment. This has a methodological advantage over prediction of a delayed outcome, in that outcome status can change over time or become influenced by other factors not yet measured.

In the evaluation of a child with a concussion, specific tools can be administered to answer the question, "Is this student at a high risk for problems with academics?" The EBA can inform clinicians' school-related treatment planning during concussion recovery by refining the predictive capability of known risk factors, clinical measures, or functional impairment to the likelihood of post-injury school problems. Relying on predefined probabilities to determine relative risk for specific outcomes, high *versus* low probability of academic problems, allows the clinician to target individualized treatments, fostering a personalized approach to intervention (Straus et al., 2011; Youngstrom, 2014). Children with concussions who are determined to be at a higher risk of academic problems could receive additional guidance and support to reduce the chance of more persisting adverse outcomes (Gioia, 2014).

In defining the post-concussion risk for perceived academic problems, we applied the following six EBA procedures, outlined by Youngstrom (2014), Chelune (2010), Straus et al. (2011), and Hunsley and Mash (2007) *with responses relevant to the current study in italics*:

1. Evaluate presenting problems and complaints and convert these into answerable questions by defining a measurable outcome variable. *In this study, we answer the question as to which measures are most useful for determining who is at risk for perceived post-injury academic problems. We first define our outcome variable (high versus low self-reported academic problems), which in the absence of a "gold standard" can be done using latent class analysis modeling. These procedures have been described elsewhere to establish group membership when a true "gold standard" is not defined for the variable of interest (Pepe, 2004; Zhou, Obuchowski, & McClish, 2002).*
2. Obtain benchmarks from a sample that is representative of the clinical patients/setting for whom this information is intended to estimate the base rate of the clinical outcome in that particular setting. *In this study, we estimate the base rate of high levels of perceived academic problems within a specialty concussion clinic.*
3. Choose relevant test measures (or "index tests") that are posited to generate meaningful data to answer the specific question. *We examine common domains of functional disruption following concussion—symptom burden, executive dysfunction, cognitive exertion, and neuropsychological test performance—as predictors of perceived academic problems following a concussion.*
4. Determine the utility of each index test by examining receiver operating characteristic (ROC) analyses and diagnostic efficiency statistics (i.e., the trade-off between sensitivity and specificity). *We examine the predictive validity of measures of symptom severity, executive dysfunction, cognitive performance, and exertional response for the dichotomous outcome of high versus low self-reported academic problems.*
5. Compare index tests to one another to identify the most powerful index test(s) for answering the question in a given setting by examining the relative predictive values of each measure. *We use the Venkatraman method for comparing ROC curves (Venkatraman & Begg, 1996) to directly compare the predictive value for measures of symptom severity, executive dysfunction, cognitive performance, and exertional responses.*
6. Examine the unique contribution of key predictors above and beyond background variables (e.g., demographic variables) in predicting the outcome of interest. *In this study, we explore sex, school grade, and pre-injury patient history separately and in combination with index tests to further refine our multimodal battery for predicting perceived academic problems.*

This study aimed to demonstrate the utility of an EBA approach to establish a set of predictive measures to identify students at risk for perceived academic problems during concussion recovery. We hypothesized that each component of the battery: (1) would discriminate students reporting high and low levels of academic problems significantly better than chance (procedures 1 through 5), and (2) would provide significant incremental validity unique to each measure (procedure 6).

METHODS

Participants

The sample consisted of 142 children and adolescents (Age: M = 14.95 years; SD = 1.80; range, 11–18 years; male = 59%) evaluated nonconsecutively in an outpatient concussion clinic within 4 weeks of injury (median = 16 days;

range, 7–28 days). Participants completed a clinical assessment battery to measure perceived academic problems, symptoms, executive functioning, exertional symptom response, and cognitive performance. Criteria for study inclusion were as follows: clinician confirmed diagnosis of concussion within the past 28 days using the Center for Disease Control's published criteria (Center for Disease Control and Prevention & National Center for Injury Prevention and Control, 2015), 10 years of age or older to allow appropriate completion of self-report questionnaires due to reading/ language limitations, and the child had to have returned to school by the time of the evaluation.

A process of data screening was conducted within the clinical database to exclude ineligible cases ($n = 22$). Specifically, patients who presented to the clinic but who did not meet criteria for a diagnosis of concussion were excluded. In addition, individuals with highly atypical symptom presentations that were *not* consistent with standard developmental (e.g., attention deficit/hyperactivity disorder ADHD; learning disability, LD) or other neurological or psychiatric diagnoses (e.g., migraine, depression, etc.) were not eligible for the study. These include children with clinical presentations complicated by a high degree of psychosocial stressors (e.g., parental divorce, litigation, or school avoidance). Finally, cases with intellectual disability, diagnosis of severe anxiety/depression, and autism, were excluded due to possible limitations in the ability to accurately respond to self-report measures.

To maximize the generalizability of our sample to other patients seen in specialty clinics, participants were intentionally included with symptoms ranging from none/mild symptoms (i.e., recovered or nearly so) to those reporting high symptom levels (presumably still actively concussed). To further increase clinical representativeness and generalizability, children with premorbid LDs, ADHD, or behavior disorders were also included in the sample. The presence of these premorbid difficulties was coded in the database using information from the clinicians' interview as well as a parent-report history form that asked parents to indicate the presence of several pre-existing conditions. The majority of participants attended public school (65%), and parents reported average to above average grades before injury (86% reported their child primarily earned As/Bs).

Measures

Outcome measure

The Concussion Learning Assessment and School Survey (CLASS) is a questionnaire completed by children and adolescents (ages, 10 to 1 years) to assess the perception of academic functioning following concussion (Ransom et al., 2015). The CLASS has been found to help identify students who have not yet recovered from concussion by showing higher levels of concern, more self-reported post-injury academic problems, and greater difficulty in classes than their recovered peers (Ransom et al., 2015). Furthermore, the total number of self-reported academic problems is positively correlated with ratings of symptom severity across self- and parent-report (Ransom et al., 2015).

The outcome measure of interest in this study was a specific question on the CLASS, in which participants were asked to indicate which problems were present that were either new or had worsened post-injury ("check all that apply"). This question assesses seven possible problems (listed in parentheses) that were grouped in two broad categories: (1) problems associated with interfering symptoms (i.e., headaches, fatigue, problems concentrating), and (2) problems associated with diminished academic skills (i.e., difficulty taking notes, difficulty understanding class material, spending more time on homework, problems studying).

Index tests

The Post-Concussion Symptom Inventory (PCSI; Sady, Vaughan, & Gioia, 2014) was used as a measure of self and parent-reported symptom severity. The 17-item child version of the PCSI (3 response choices) was administered to 18 students under the age of 12 and the 21-item adolescent version was administered to 124 students aged 13–18 years (7 response choices). Adjusted post-injury score totals were computed by subtracting retrospective baseline (RBL) ratings of pre-injury symptom report from post-injury report. Mean score ratings were calculated between the child and adolescent versions.

To calculate a comparable metric across all versions of the PCSI given the different number of items and response categories, the child version mean scores were multiplied by two. We chose this multiplier because it resulted in a distribution of scores that was very similar (no significant differences in mean, standard deviation, or range) in those who completed the children's version of the PCSI compared to those who were over 13 years of age and completed the adolescent version. In addition, we ran all of our ROC and logistic regression analyses without these 18 participants, and there were no substantive differences in the results, other than a loss of power that made one of the interactions non-significant.

A modified version of the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Kenworthy, & Guy, 2000) was used as a parent- and self-report measure of executive functions. The modified BRIEF includes 34 and 35 items on the parent- and self-report versions, respectively, using a 5-point response scale. Three symptom validity (low endorsement items) are embedded in the measure to detect possible over-reporting and endorsement of unrealistic symptoms (i.e., forgets where bedroom is located, cannot remember friends' names, has difficulty chewing food). This modified version was found to produce factors along three dimensions; Emotional (i.e., Emotional Control), Behavioral Regulation (i.e., Inhibition), and Cognitive Regulation (i.e., Working Memory, Planning/Organization, and Task Completion; McGill, Gerst, Isquith, & Gioia, 2011). These factors have been shown to reliably measure change in behavioral manifestations of executive functioning following concussion (McGill et al., 2011). As with the PCSI, children and their parents provided RBL ratings of pre-injury functioning on each item together with ratings of current post-injury functioning. To calculate the post-injury change in functioning, raw score totals were adjusted by subtracting RBL ratings from post-injury ratings.

The Children's Exertional Effects Rating Scale (ChEERS) is a rating scale designed to capture dynamic change in symptom severity following cognitive activity. The ChEERS is completed by the patient before and after a session of neuropsychological testing (Sady, McGill, Gerst, & Gioia, 2013). Four symptoms were measured, including headache, fatigue, concentration problems, and irritability. Each symptom was presented with a scale from 0 (not experiencing symptom) to 10 (severe). Children ages 12 and younger were provided with cartoon faces to provide additional context for response options. The Exertional Effects Index (EEI) was calculated by subtracting the pre-test ratings from the post-testing ratings, and negative EEI scores (higher pre-test than post-test) were set to zero for analysis.

Participants underwent a battery of computerized cognitive tests. Fifteen children ages 11 to 12 completed the Multimodal Assessment of Cognition & Symptoms for Children (MACS; Newman, Reesman, Vaughan, & Gioia, 2013) and 127 participants ages 12 years, 6 months to 18 completed the Immediate Post Concussion Assessment and Testing (ImPACT; Lovell, Collins, Podell, Powell, & Maroon, 2000). The MACS generates two composite standard scores, Memory and Speed. On ImPACT, two similar composite scores were calculated from the four ImPACT subcomposite scores by transforming the scores from percentiles to standard scores and then averaging the two memory and two speed measures (Schatz & Maerlender, 2013).

Procedure

Parents provided demographic information, including the child's age, level of schooling, gender, race, ethnicity, and the presence of specific pre-injury mental health and medical diagnoses (e.g., anxiety, depression, attention deficit/hyperactivity disorder, LD, and headache/migraines; see Table 1). Children and their parents completed rating scales, including the outcome data from the CLASS, and performed cognitive testing described above at the time of their clinical evaluation. Since information was gathered in the clinic setting, the order of administration (e.g., questionnaires or testing first) varied. Exertion ratings were standardly collected at the start and end of the combined testing/questionnaire session. The contemporaneous administration of the predictor and criterion tests is generally considered desirable, in that having a short lag reduces the amount that the target condition is likely to change (in this case, number of self-reported academic problems) relative to current symptom or cognitive status (Bossuyt et al., 2003).

While it would be valuable to know how a current assessment may predict future outcomes such as recovery, identifying current status and to predict current treatment needs is similarly useful (Findling et al., 2013; Kraemer, 1992). Data were collected with Institutional Review Board approval (Children's National Health System), with informed consent obtained from parents and assent from children and adolescents. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Data Analysis

Data analyses were performed with SPSS statistical software version 22 (IBM SPSS Statistics, IBM Corporation), Latent Gold 5.0 (Statistical Innovations, Inc.; Vermunt & Magidson, 2005), and R 3.2 (R Core Team, 2013). We first engaged in detailed data examination and cleaning. We examined distributions on all continuous measures (e.g., PCSI, modified BRIEF monitor) and eliminated individuals whose scores exceeded three standard deviations from the mean, to avoid undue influence of statistical outliers on the results.

We also used several methods to screen for unusual or excessive reporting on rating scales. On the PCSI, we examined previously collected samples, both uninjured and injured (see Sady et al., 2014), for symptoms that were rarely endorsed at a high level. In these samples, less than 5% of both groups endorsed more than five rare symptoms in the extreme (a rating of 2 on the 8- to 12-year-old PCSI, or a rating of 5–6 on the 13- to 18-year-old PCSI). In our sample, only one individual endorsed five rare symptoms at a high level. Since this individual did not endorse an overall extreme level of symptoms and they were not identified as a multivariate outlier in subsequent analyses, they were retained.

On the modified BRIEF-Monitor, we examined the three symptom validity items that were included to screen for possible over-reporting and endorsement of unrealistic symptoms. Two children and two parents endorsed two of three items, and one child and one parent endorsed three of three; none of these comprised a parent-child pair. Examination of the outcome variables for the elevated cases found that the number of reported school problems was within 1 *SD* of the mean, so these individuals were also not excluded.

Finally, we examined the influence of time since injury on our variables, given that the range was from 2 to 28 days. The two outcome groups differed significantly, with the high perceived problems group being on average 2.5 days longer post-injury (see Table 1). However, when added to the models, days since injury did not predict group membership after accounting for self-report questionnaires, so we did not include it in the final set of analyses. We also evaluated the predictors with respect to time since injury, as we wanted to determine whether there was any evidence of systematic recall bias on preinjury ratings (which were used to compute adjusted ratings). There were no significant correlations between number of days since injury and mean retrospective ratings on any of the self-report or parent report measures ($r < |.05|$; $p > .56$ for all four).

Following data screening, we applied the following set of statistical procedures to achieve our aims, corresponding to the six EBA procedures identified above. First, to establish a distinct grouping outcome variable for self-reported academic problems on the CLASS, Latent Class Analysis (LCA) identified a two-class solution based on the distribution of the responses in the study sample, following standard procedures for a "missing gold standard" scenario (Hagenaars & McCutcheon, 2002; Zhou et al., 2002). Individuals within a given latent class category are, therefore, similar to one another with respect to their responses on the CLASS (i.e., high *vs.* low self-reported academic problems). The LCA is a latent variable model, essentially a factor analysis with a categorical factor, and class membership is assigned probabilistically based on the constellation of observed responses. Independent sample *t* tests quantified criterion validity, or the association of each predictor to the LCA category. Because LCA is a latent variable approach, there were not set classification thresholds on each observed indicator. For example, moderate scores on several items could result in classification in the "high problems" group, but so could severe scores on a smaller set of items.

For the second and third steps, we examined the frequency of high and low self-reported academic problems to establish the base rate within our specialty concussion clinic, and we evaluated the relationship between each index test and the outcome to assess relevance.

Fourth, ROC analyses were conducted to evaluate the trade-off between diagnostic specificity and sensitivity for each of the index tests in distinguishing students reporting high and low academic problems. The area under the curve (AUC) quantified the accuracy of the measure in predicting group membership. Chance levels of diagnostic performance correspond to an AUC of .50, while an AUC of 1.0 reflects perfect performance. Benchmarks provided by Rice and Harris (2005) indicate that AUCs in the mid .50s are small and not clinically useful; AUCs in the mid .60s are medium in size and may provide some incremental clinical value, but are not sufficient in isolation; AUCs in the low to mid .70s reflect large, clinically informative values; AUCs in the low .80s are excellent; AUCs in the high .80s are exceptional under clinically rigorous designs and are often highly informative; while AUCs in the .90s are extraordinary. In practice, Youngstrom (2014) highlighted that rating scales producing AUCs greater than .90 are more likely to indicate design flaws rather than a high level of discriminative validity (e.g., Youngstrom, Meyers, Youngstrom, Calabrese, & Findling, 2006).

Fifth, significance tests (Venkatraman & Begg, 1996) were conducted to directly compare index tests in their ability to discriminate individuals reporting high and low academic problems.

Sixth, a hierarchical logistic regression was used to test the incremental validity of the index tests at discriminating outcome groups, after accounting for the effects of demographic and pre-injury history. Additionally, logistic regression models explored whether the utility of particular tests depends upon other factors (i.e., potential demographic/background moderators). Demographic and pre-injury history variables, as well as self- and parent-report of symptoms and executive dysfunction (PCSI and the modified BRIEF), cognitive performance measures, and ratings of exertional response (EEI), and the interactions between demographic and pre-injury history and all other predictors were entered hierarchically into a logistic regression model. Continuous variables were mean-centered to reduce multicollinearity and improve interpretability of interaction terms. The order of entry for variables in the regression model was chosen to examine the incremental predictive validity of assessment measures with increasing examiner and patient burden (i.e., demographics, questionnaires, cognitive testing/EEI).

By using this order of entry, a more burdensome (time- and cost-consuming) measure would only be recommended for clinical use in identifying risk of self-reported academic problems if it showed incremental predictive power above less burdensome measures. Pairwise interactions between demographics and questionnaires, cognitive testing, and the EEI (21 total) were included to assess whether the discriminative power is significantly better amongst a particular subsample (e.g., whether the PCSI was equally beneficial for both genders). All nonsignificant interactions were iteratively eliminated from the final regression model to conserve power and improve precision of estimates. Two significant interactions were retained within the final model.

RESULTS

The following results were obtained within the aforementioned six procedures.

1. *Evaluate presenting problems and complaints and convert these into answerable questions by defining a measureable outcome.* Latent class analysis was conducted on the patient reported number of academic problem on the CLASS to establish two criterion groups, high and low perceived academic problems. Table 1 presents the study demographics, pre-injury characteristics, and clinical measures, comparing latent class groups with *t* tests. Significant differences were found between groups on self/parent-reported pre-injury history of ADHD/LD, anxiety, depression, and/or headache/migraine, in addition to self- and parent-report of symptoms and executive dysfunction (PCSI and the modified BRIEF) and ratings of exertional response (EEI). No differences in demographics, aspects of pre-injury history, or cognitive performance measures were found between the two latent class criterion groups.
2. *Obtain benchmarks from a sample that is representative of the clinical patients/setting for whom this information is intended, in order to estimate the base rate of the clinical outcome in that particular setting.* In our sample of adolescents who presented to a specialty concussion clinic, 79 (56%) were categorized as reporting a "low" number of academic problems (M = 1.61; SD = 1.10) and 63 (44%) were categorized as reporting a "high" number of perceived academic problems (M = 5.29; SD = 1.25) based on patient report on the CLASS.
3. *Choose relevant test measures (or "index tests") that are posited to generate meaningful data to answer the specific questions.* To examine the pairwise relation

Table 1. Descriptive statistics for clinical and demographic variables and bivariate tests of association with perceived school problems status at visit 1 (N = 142)

Variable	School problems: high (n = 63)	School problems: low (n = 79)	Test statistic	p-Value	Effect size
Age in years					
Mean	15.14	14.79	t (140 df) = 1.14	.26	d = .19
SD	1.68	1.89			
Female	n = 28 (44%)	n = 30 (38%)	χ^2 (1 df) = 0.61	.44	phi = .07
Race (Caucasian %)	n = 50 (83%)	n = 58 (77%)	χ^2 (1 df) = 0.75	.39	phi = .07
Level of schooling (high school)	n = 41 (65%)	n = 49 (62%)	χ^2 (1 df) = 0.14	.71	phi = .03
Pre-injury history					
ADHD/LD	n = 17 (27%)	n = 15 (19%)	χ^2 (1 df) = 1.28	.26	phi = .09
Anxiety/Mood disorder	n = 16 (25%)	n = 16 (20%)	χ^2 (1 df) = 0.53	.47	phi = .06
Headaches/migraines	n = 31 (49%)	n = 28 (35%)	χ^2 (1 df) = 2.73	.10	phi = .14
At least one of the above	n = 46 (73%)	n = 42 (53%)	χ^2 (1 df) = 5.86	.02	phi = .20
At least two of the above	n = 16 (25%)	n = 14 (18%)	χ^2 (1 df) = 1.24	.27	phi = .09
All three of the above	n = 2 (3%)	n = 2 (3%)	χ^2 (1 df) = 0.05	.82	phi = .02
Injury: sport-related concussion	n = 54 (87%)	n = 61 (77%)	χ^2 (1 df) = 2.26	.13	phi = .13
Injury characteristics					
Loss of consciousness	n = 9 (15%)	n = 7 (9%)	χ^2 (1 df) = 1.12	.29	phi = .09
No recall of impact	n = 26 (41%)	n = 22 (28%)	χ^2 (1 df) = 2.82	.09	phi = .14
Retrograde amnesia	n = 12 (19%)	n = 11 (14%)	χ^2 (1 df) = 0.69	.41	phi = .07
Anterograde amnesia	n = 18 (27%)	n = 14 (18%)	χ^2 (1 df) = 2.12	.15	phi = .12
Seizures	n = 1 (2%)	n = 1 (1%)	χ^2 (1 df) = 0.02	.89	phi = .01
Days since injury *M (SD)*	18.29 (6.04)	15.81 (6.04)	t (140 df) = 2.43	.02	d = 0.41
PCSI Self-Report mean score					
M	1.26	.43	t (140 df) = 6.78	<.001	d = 1.18
(SD)	(.83)	(.55)			
PCSI Parent-Report mean score					
M	1.51	.58	t (140 df) = 5.81	<.001	d = .98
(SD)	(1.03)	(.87)			
Modified BRIEF Self-Report raw total					
M	22.48	7.41	t (140 df) = 6.50	<.001	d = 1.12
(SD)	(15.70)	(10.79)			
Modified BRIEF Parent-Report raw total					
M	14.37	6.03	t (140 df) = 3.77	<.001	d = .63
(SD)	(14.64)	(11.70)			
Exertional Effects Index					
M	3.75	2.10	t (140 df) = 3.41	<.001	d = .48
(SD)	(3.13)	(3.75)			
Cognitive measures (ImPACT/MACS):					
Processing Speed SS					
M	90.87	93.52	t (140 df) = .92	.36	d = .16
(SD)	(15.79)	(17.99)			
Memory SS					
M	91.75	94.96	t (140 df) = 1.18	.24	d = .19
(SD)	(16.50)	(15.73)			

between the criterion variable and each of the clinical predictors, we examined chi square analyses, analyses of variance, point biserial, and Pearson correlations (Tables 1 and 2). Report of high academic problems was associated with pre-injury history (i.e., history of ADHD/LD, depression, anxiety, or headache/migraine) in addition to greater overall post-concussion symptom severity, executive problems, and exertional response, but was not associated with cognitive performance measures. We also ran a supplemental confirmatory factor analysis to check that the measures covered distinct constructs, rather than a single method variance factor. A single factor solution, consistent with the shared method hypothesis, fit badly: chi-squared (2 df) = 43.68; $p < .000005$; comparative fit index (CFI) = .905 Tucker Lewis index (TLI) = .716, whereas a two-factor model for the distinct constructs provides more acceptable fit, chi-squared (1 df) = 16.28; $p < .00005$; CFI = .965; TLI = .792. Overall, the findings provided preliminary evidence that the majority of

Table 2. Correlations among variables

Variable	Female	High school	Pre-injury history[a]	PCSI-Self	PCSI-Parent	Modified BRIEF-Self	Modified BRIEF-Parent	EEI	Processing Speed SS[c]	Memory SS[c]
School problems: high[b]	−.07[d]	.03[d]	.20**[d]	.51***[e]	.44***[e]	.50***[e]	.30***[e]	.28***[e]	−.08[e]	−.10[e]
Gender: female		.13[d]	.14[d]	.25**[e]	.18*[e]	.21**[e]	.21**[e]	.09[e]	.03[e]	.06[e]
School level: high school			−.01[d]	−.09[e]	.07[e]	.11[e]	.13[e]	.09[e]	.17*[e]	.07[e]
Pre-injury history[a]				.04[e]	.17*[e]	−.04[e]	−.01[e]	−.06[e]	.15[e]	.04[e]
PCSI Self					.58***	.68***	.42***	.39***	−.28***	−.25**
PCSI Parent						.52***	.65***	.30***	−.14	−.15
Modified BRIEF Self							.55***	.26**	−.23**	−.19*
Modified BRIEF Parent								.22**	−.14	−.15
EEI									−.16	.11
Processing Speed SS[c]										.52***

[a]Pre-injury history includes diagnoses of ADHD, learning disability, anxiety, depression, or personal history of headaches/migraines.
[b] Coded such that low academic problems = 0 and high academic problems = 1.
[c]Standard Score; all others are raw scores, adjusted for retrospective ratings of pre-injury functioning.
[d]Phi Coefficient.
[e]Point-biserial correlation; all others are Pearson *r* correlations.
**p* < .05, ***p* < .01, ****p* < .001, two-tailed.

selected index tests are relevant to the prediction of self-reported academic problems.

4. *Determine the utility of each index test by examining diagnostic efficiency statistics (i.e., the trade-off between sensitivity and specificity) via ROC analyses:* Table 3 and Figure 1 present the ROC curves for the eight clinical measures. According to Rice and Harris' criteria (2005), self-reported ratings on both the PCSI and the modified BRIEF demonstrated excellent accuracy in distinguishing participants reporting high academic problems from those reporting low academic problems, earning AUCs of 0.80 ($p < .001$) and 0.84 ($p < .001$), respectively. These findings indicate that a randomly selected student reporting high academic problems scores higher than a student reporting low academic problems 80% of the time on the PCSI and 84% of the time on the modified BRIEF. Parent report on the PCSI and the modified BRIEF demonstrated large, clinically informative accuracy, with AUCs of .79 ($p < .001$) and .74 ($p < .001$), respectively. In contrast, performance on cognitive measures did not significantly distinguish students reporting high and low academic problems (Processing Speed AUC = .56; Memory AUC = .57, *ns*), although the EEI measuring the exertional response resulting from this cognitive effort demonstrated a large, clinically informative AUC of .70 ($p < .001$).
5. *Compare index tests to one another to identify the most powerful index test(s) for answering the question in a given setting by examining the relative predictive values of each measure.* Venkatraman tests (Venkatraman & Begg, 1996) were used to directly compare the index tests with clinically informative AUCs to one other to examine significant differences in their ability to discriminate among students reporting high and low academic problems. Of all possible comparisons (eight total), the only significant difference was observed between ratings on the modified BRIEF, with self-report identifying students reporting high academic problems significantly better than parent-report ($p < .03$). However, this finding was not significant after correcting for multiple comparisons.
6. *Examine the unique contribution of the key predictors as well as background variables (e.g., demographic variables) in predicting perceived academic problems.* First, Pearson and point biserial correlations quantified the relation among predictors to check the degree of multicollinearity (Table 2). To examine the second study hypothesis, hierarchical logistic regression assessed the incremental utility of index tests to predict group membership of low *versus* high levels of self-reported academic problems. Nonsignificant interactions were trimmed from the final model.

Results indicated that demographics and pre-injury history did not provide significant prediction when entered in the first block of the model ($p = .07$; see Table 4). Comparison of log-likelihood ratios for models with and without self-and parent-report measures showed statistically significant improvement with addition of self-report measures in the second block ($\chi^2 = 54.43$; $p < .001$), but not after addition of parent-report measures in the third block ($\chi^2 = 2.22$; $p = .33$).

Table 3. AUC from ROC analyses identifying students reporting school problems at visit 1 with index tests and moderators

				95% Confidence interval	
Index test	Area under curve	Standard error	*p*-Value	Lower	Upper
Modified BRIEF Self-Report	.84	.03	<.001	.78	.91
PCSI Self-Report	.80	.04	<.001	.73	.87
PCSI Parent Report	.79	.04	<.001	.72	.87
Modified BRIEF Parent-Report	.74	.04	<.001	.66	.83
Exertional Effects Index	.70	.04	<.001	.61	.78
Cognitive performance:					
Processing Speed	.56	.05	.22	.47	.66
Memory	.57	.05	.17	.47	.66

Note. Benchmarks provided by Rice and Harris (2005) indicate that AUCs in the mid .50s are small and not clinically useful; AUCs in the mid .60s are medium in size and may provide some incremental clinical value, but are not sufficient in isolation; AUCs in the low to mid .70s reflect large, clinically informative values; AUCs in the low .80s are excellent; AUCs in the high .80s are exceptional under clinically rigorous designs and are often highly informative; while AUCs in the .90s are extraordinary.

In the fourth block, there was a significant interaction between self-reported PCSI and level of schooling on the prediction of self-reported academic problems, and between self-reported executive dysfunction and pre-injury history on the prediction of self-reported academic problems (χ^2 = 11.742; p = .003). Specifically, the discriminative power of self-reported PCSI scores is stronger for high school students compared with elementary/ middle school students (B = 1.87; p = .03).

However, ROC analysis of elementary/middle school students alone showed that self-reported PCSI significantly predicted perceived academic problems, albeit not as strongly as high school students (AUC = .67; p = .04 and AUC = .88; $p < .001$, respectively). Furthermore, greater self-reported adjusted scores on the modified BRIEF significantly increased the likelihood of high self-reported academic problems, particularly in individuals with a pre-injury history of ADHD/ LD, depression, anxiety, or headache/migraine (B = 0.13; p = .02). However, self-reported BRIEF scores demonstrated equally significant discriminative power in individuals with and without pre-injury risk factors (AUC = .85; $p < .001$ for both groups). In blocks 5 and 6, addition of cognitive performance and exertional response did not explain a unique portion of the variance above and beyond demographics and self- and parent-reported measures (χ^2 = 1.13; p = .77).

In sum, self-reported ratings of symptoms predicted perceived academic problems in all students, but did so more strongly for high school students, and self-reported executive dysfunction predicted high perceived academic problems for all individuals. Parent report of symptoms and executive dysfunction, performance on cognitive measures, and exertional response did not account for additional significant variance

DISCUSSION

Children and adolescents who sustain a concussion are at risk for perceived academic problems during and immediately following recovery, including such difficulties as understanding new concepts, headaches or fatigue interfering with productivity, and difficulty completing work on time (Ransom et al., 2015). Defining the most useful predictors of who is most likely to perceive academic problems would better guide the clinician in using an efficient and targeted evaluation to improve clinical decision-making and management. This study applied an EBA approach, using a set of six statistical procedures, to examine a multimodal set of assessment measures—including patient demographics, pre-injury history, parent- and self-report of symptoms and executive functions, cognitive performance measures, and exertional response—to identify those measures that best differentiate children at risk for perceived post-injury academic problems.

We posited that each component of the battery (1) would discriminate students reporting high and low levels of academic problems significantly better than chance, and

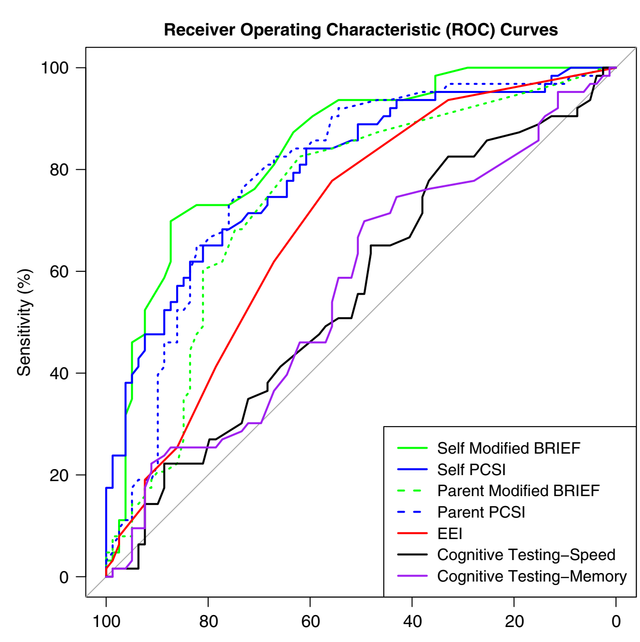

Fig. 1. ROC curves for the eight clinical measures.

Table 4. Logistic regression model identifying students reporting high levels of post-injury school problems

	B	Standard error	p-Value	Odds ratio
Block 0:				
Intercept	–0.99	0.54	0.06	0.37
Block 1: $\chi 2$ = 7.00 , p = .07				
Gender (female)	0.36	0.50	0.47	1.44
School level (high school)	–0.01	0.50	0.98	0.99
Pre-injury characteristics[a]	1.26	0.53	0.02	3.54
Block 2: $\chi 2$ = 54.43, p < .001				
Modified BRIEF Self-Report	0.03	0.03	0.27	1.03
PCSI Self-Report	0.07	0.55	0.90	1.08
Block 3: $\chi 2$ = 2.22 , p = .33				
PCSI Parent-Report	0.47	0.31	0.13	1.59
Modified BRIEF Parent-Report	–0.01	0.02	0.66	0.99
Block 4: $\chi 2$ = 11.74 , p = .003				
High school x PCSI Self-Report	1.87	0.84	0.03	6.47
Pre-injury history x Modified BRIEF Self-Report	0.13	0.06	0.02	1.14
Block 5: $\chi 2$ = 1.13 , p = .77				
EEI	0.08	0.08	0.36	1.08
Cognitive measures[b]:				
Processing Speed	–0.009	0.02	0.57	0.99
Memory	0.002	0.02	0.92	1.00

[a]Presence of pre-injury history of ADHD, LD, anxiety, depression, or history of headache/migraine.
[b]ImPACT and MACS Processing Speed and Memory composite standard scores. High scores indicate greater difficulty.

(2) would provide significant incremental validity unique to each measure. We first defined the clinical outcome using Latent Class Analysis to categorize individuals into high and low self-reported academic problems and examined base rates within our specialty concussion clinic sample. Next, we used ROC analyses to examine the discriminative validity of the assessment battery's components in identifying students reporting high and low levels of academic problems. We directly compared the utility of each measure to one another through a series of Venkatraman tests. Finally, we conducted a hierarchical logistic regression to explore the incremental validity of measures above and beyond other measures.

Results demonstrated partial support for the first study hypothesis, as we found significant discriminative validity of self- and parent-reported measures of symptom severity and executive functions, as well as the student's exertional response during testing. In contrast, cognitive performance measures did not significantly differentiate students reporting high and low levels of academic problems. The second study hypothesis was partially supported as well, indicating that self-reported symptoms and executive dysfunction were the best predictors of perceived academic problems, although the strength of prediction for symptom status was dependent on level of schooling. Parent-report of symptoms or executive dysfunction and exertional response, although significant by themselves, did not provide significant incremental validity above and beyond self-report measures in predicting perceived academic outcome. Cognitive performance did not predict perceived academic problems above and beyond other measures.

Several implications can be drawn from this study. First, self-reported measures of symptom severity and executive dysfunction are the strongest predictors of perceived academic problems during concussion recovery. Although parent-reported measures and exertional response are valid indicators for identifying at-risk students, clinicians evaluating children with concussions might consider prioritizing self-report of post-concussion symptoms and executive dysfunction to reduce the cost and burden of an extensive evaluation. Second, this study highlights the methodological advantages of an EBA approach in identifying clinically relevant tools in answering important questions related to school outcomes.

EBA can provide a unique perspective for the field of neuropsychology. Refining research questions in the context of EBA (procedure 1) can yield specific, measurable outcomes that are well suited to clinical assessment. While this study explored relative risk for perceived post-concussion academic problems, such a model could easily be applied to a variety of other clinical questions relevant to children with concussions (e.g., risk for prolonged recovery) or other aspects of neuropsychology (e.g., risk for post-treatment effects in pediatric cancer survivors). To evaluate clinical outcomes without a clearly defined "gold standard" for identification, we demonstrate the use of latent class analysis to provide a meaningful outcome criterion (McCutcheon, 1987; Zhou et al., 2002).

Evaluating the utility of neuropsychological assessment tools with an EBA approach (procedures 3 and 4) can define their applicability to important diagnostic and functional outcomes. ROC analysis can provide a clinically relevant measure of the strength of the relationship between a clinical predictor and the outcome of interest, perhaps more so than traditional effect sizes such as a bivariate correlation or Cohen's d (Kraemer et al., 1999), which are helpful for understanding group-level differences but do not address classification accuracy for individual cases.

The AUC statistic of the ROC analysis reflects the predictive validity of each assessment measure and makes a compelling case for their clinical utility in selecting a test battery. Examining clinical factors that predict a patient's current status allows for an empirically based decision to occur at the time of the appointment which can be more advantageous than predicting a delayed outcome, as the outcome status can change over time or become influenced by other factors. In the evaluation of pediatric concussion, self- and parent-reported measures of symptoms and executive functions demonstrated significant predictive utility in identifying students reporting high academic problems, above and beyond performance on cognitive measures and would, therefore, be considered as primary indicators of this important outcome for a school-aged child.

Examining the incremental validity of the measures within the neuropsychological assessment (procedure 6) defines the "added value" of each component. By exploring the relative amount of variance explained by measures in a hierarchical model, a parsimonious set of clinical measures can be defined with empirical evidence to most efficiently answer the question at hand. This study highlighted the unique contribution of self-reported symptom measures (PCSI and modified BRIEF) to identify students at high risk for perceived adverse academic effects relative to the more limited information added by parent-reported measures.

Furthermore, more fine-grained examination of the interaction effects between level of schooling and self-report measures also showed the strength of symptom report in high school students. An important consideration regarding the EBA approach is that although some measures in the battery may not add a unique contribution to the prediction of an outcome, corroborating data may, at times, provide clinically meaningful information. For example, athletes sometimes underreport post-injury symptoms in an effort to prematurely return to competition; parent report may better elucidate the level of post-injury symptoms and/or executive dysfunction in the event of questionable self-report validity. Although the parent reports in this study did not provide significant incremental validity beyond the self-report, they nevertheless did demonstrate a strong relationship with the outcome variable, and could substitute as a surrogate predictive indicator if needed.

Limitations

Several limitations are evident in the present study. One limitation lies in the use of a self-reported criterion variable (e.g., self-reported post-injury school problems) without an independent, corroborating measure of school problems such as course grades or teacher report of performance. Although participants were asked to report new/worsened post-injury academic problems, recall bias or heightened anxiety/concern may have impacted perception of past or current difficulties. Recall bias is a systematic influence on ratings of behaviors or other phenomena that occurred in the past. In the field of concussion, a "good old days" bias has been reported in adults, such that ratings of preinjury functioning tend to be more positive than reality would suggest (Iverson, Lange, Brooks, & Rennison, 2010; Mittenberg, DiGiulio, Perrin, & Bass, 1992), but much of the research on this phenomenon relates to long post-injury intervals (many months to years).

Iverson et al. (2010) found that the bias was most apparent in individuals who also failed multiple validity tests (not simply those involved in litigation, as previous research had suggested). Brooks et al. (2014) found that parents' preinjury ratings declined over the first month post-injury, suggestive of the good old days bias in children with concussion; however, this longitudinal design introduces the possibility of the decrease being accounted for by the test–retest attenuation that is commonly seen with multiple administrations of rating scales (e.g., Gioia, 2015; Gioia et al., 2000). In fact, the screening analyses in the current cross-sectional design showed no association between time since injury and level of retrospective preinjury ratings. Similarly, days since injury was not a significant predictor of our outcome once more relevant variables (such as current symptom status) were entered into the models. Taken together, there was no evidence of systematic bias in recall over the first month post-injury in this study.

In addition, validity of self-report could not be confirmed in the absence of objective measures. There are multiple reasons participants could have reported a high level of academic problems, including individuals who were truly experiencing problems and those who rated a high number of problems due to anxiety, a feeling of helplessness or of being overwhelmed with missed work, or even exaggeration or feigning. While we screened our data to avoid including those with obvious symptom exaggeration, it is possible that some participants were not presenting their true status. It is possible that exertional effects may have generally influenced responses on child self-report measures, as highly exertional patients may not have responded as carefully to questions. We would not expect these effects to be specific to any one measure, however, as the order of test administration varied across examiners.

Method variance is a second limitation in that the strongest relationships were found among the self-report measures of symptoms, executive dysfunction, and academic problems. Supplemental analyses confirmed that a single, method variance factor fit these indicators poorly, allaying concern that results were driven by shared source variance and highlighting that results from this study are not driven by self-report measures predicting self-report measures. It also is reassuring that parent-reported measures also discriminated the classes, albeit not showing incremental validity after controlling for self-report, showing validity for a distinct source. Future research should examine academic outcomes across reporters, including self-report, parent-report, teacher report, and other external indicators of academic performance (e.g., test grades, report cards). Similarly, future studies can use this EBA model to examine the utility of a range of predictors, as results from this study are limited to the use of these specific measures (PCSI, modified BRIEF, ChEERS) and may not apply to other measures.

An inability to account for change in cognitive performance from pre-injury to post-injury may explain the limited utility of the cognitive performance measures. For other clinical predictors, a change score (i.e., the difference between reports of RBL and current symptoms/executive dysfunction) allowed for an examination of change in functioning compared to an individual's own pre-injury level of problems, which may be more sensitive to perceived concussion related academic problems than a single post-injury test score. Further investigation of the link between cognitive performance and problems might be bolstered by the addition of a "hold" test as a control for preinjury performance (i.e., a reading or vocabulary test).

An additional limitation is the sole inclusion of students and their parents who sought evaluation and treatment in a specialty concussion clinic, possibly reflecting a more affected sample with a more prominent self/parent-reported pre-injury history of ADHD/LD, anxiety/mood concerns, and/or headaches/migraine. Finally, our analyses were limited to a clinical sample and did not include an uninjured contrast group. This was by design, however, as these results were meant to generalize best to concussion clinic settings, where the key comparison is not between those with or without a concussion, but rather among those with a concussion to identify who is at risk for perceived academic problems.

CONCLUSION

This study demonstrates a systematic, multi-procedure method highlighting the utility of an EBA model in establishing a multimodal toolset for identifying students at risk for self-reported post-injury academic problems. Students recovering from concussion are at risk for perceived academic problems during recovery, as predicted by measures of symptom severity, executive dysfunction, and exertional response. Application of an EBA model offers an important perspective in the management of pediatric concussion by defining the key tools that best predict this adverse outcome for children and adolescents. These findings inform clinicians about the measures that best predict perceived academic problems following concussion, allowing them to make an informed decision on the selection of research-validated assessment measures.

The relationships between self-reported symptoms/executive functioning problems and perceived school problems elucidated in this study provide a springboard for further investigation of several questions. For example, it will be important to look at when these perceived problems resolve in relation to recovery, how well perceived problems map onto more objective measures of decreased school performance, and how perceived problems relate to other important indicators of injury status. Additional research should also investigate the types of accommodations and strategies that are effective in alleviating, or at least minimizing long-term impact of, these difficulties during recovery.

These findings are a first step in defining the relationship between predictors of school-related outcomes, and an important next step is to provide metrics for clinical use, including what specific scores on the identified measures result in a "high risk" designation for any given individual. Follow-up work by this author team will provide guidance on the clinical application of the EBA model, articulating the at-risk scores on the PCSI, modified BRIEF, and ChEERS that would alert the clinician to take action to meet concurrent treatment needs and to identify those at risk for possible post-injury school problems.

ACKNOWLEDGMENTS

This work was supported by the Eunice Kennedy Shriver Intellectual and Developmental Disabilities Research Center (IDDRC; P30HD040677) and the Clinical and Translational Science Award (UL1TR000075). *Conflict of interest.* None.

REFERENCES

Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L.M., … de Vet, H.C.W. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *British Medical Journal*, *326*, 41–44.

Brooks, B.L., Kadoura, B., Turley, B., Crawford, S., Mikrogianakis, A., & Barlow, K.M. (2014). Perception of recovery after pediatric mild traumatic brain injury is influenced by the "good old days" bias: Tangible implications for clinical practice and outcomes research. *Archives of Clinical Neuropsychology*, *29*(2), 186–193. http://doi.org/10.1093/arclin/act083

Center for Disease Control and Prevention, & National Center for Injury Prevention and Control. (2015, February 26). What are the signs and symptoms of concussion? Retrieved from http://www.cdc.gov/concussion/signs_symptoms.html

Chelune, G.J. (2010). Evidence-based research and practice in clinical neuropsychology. *The Clinical Neuropsychologist*, *24*(3), 454–467. http://doi.org/10.1080/13854040802360574

Echemendia, R.J., Iverson, G.L., McCrea, M., Macciocchi, S.N., Gioia, G.A., Putukian, M., & Comper, P. (2013). Advances in neuropsychological assessment of sport-related concussion. *British Journal of Sports Medicine*, *47*(5), 294–298. http://doi.org/10.1136/bjsports-2013-092186

Findling, R.L., Jo, B., Frazier, T.W., Youngstrom, E.A., Demeter, C.A., Fristad, M.A., … Horwitz, S.M. (2013). The 24-month course of manic symptoms in children. *Bipolar Disorders*, *15*(6), 669–679. http://doi.org/10.1111/bdi.12100

Gioia, G.A. (2014). Medical-school partnership in guiding return to school following mild traumatic brain injury in youth. *Journal of Child Neurology*, 883073814555604. http://doi.org/10.1177/0883073814555604

Gioia, G.A. (2015). Multimodal evaluation and management of children with concussion: Using our heads and available evidence. *Brain Injury*, *29*(2), 195–206. http://doi.org/10.3109/02699052.2014.965210

Gioia, G.A., Isquith, P.K., Guy, S.C., & Kenworthy, L. (2000). TEST REVIEW behavior rating inventory of executive function. *Child Neuropsychology*, *6*(3), 235–238. http://doi.org/10.1076/chin.6.3.235.3152

Evidence-Based Medicine Working Group. (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. *JAMA*, *268*(17), 2420–2425. http://doi.org/10.1001/jama.1992.03490170092032

Hagenaars, J.A., & McCutcheon, A.L. (2002). *Applied latent class analysis*. New York: Cambridge University Press.

Halstead, M.E., McAvoy, K., Devore, C.D., Carl, R., Lee, M., & Logan, K. (2013). Returning to learning following a concussion. *Pediatrics*, 948–957. http://doi.org/10.1542/peds.2013–2867

Hunsley, J., & Mash, E.J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology*, *3*(1), 29–51. http://doi.org/10.1146/annurev.clinpsy.3.022806.091419

Iverson, G.L., Lange, R.T., Brooks, B.L., & Rennison, V.L.A. (2010). "Good old days" bias following mild traumatic brain injury. *The Clinical Neuropsychologist*, *24*(1), 17–37. http://doi.org/10.1080/13854040903190797

Kraemer, H.C. (1992). *Evaluating medical tests: Objective and quantitative guidelines*. Newbury Park, CA: Sage.

Kraemer, H.C., Kazdin, A.E., Offord, D.R., Kessler, R.C., Jensen, P.S., & Kupfer, D.J. (1999). Measuring the potency of risk factors for clinical or policy significance. *Psychological Methods*, *4*(3), 257–271. http://doi.org/10.1037/1082-989X.4.3.257

Lovell, M.R., Collins, M., Podell, K., Powell, J., & Maroon, J. (2000). *ImPACT: Immediate post-concussion assessment and cognitive testing*. Pittsburgh, PA: Neurohealth Systems, LLC.

McCrory, P., Meeuwisse, W.H., Aubry, M., Cantu, B., Dvořák, J., Echemendia, R.J., … Turner, M. (2013). Consensus statement on concussion in sport: The 4th International Conference on Concussion in Sport held in Zurich, November 2012. *British Journal of Sports Medicine*, *47*(5), 250–258. http://doi.org/10.1136/bjsports-2013-092313

McCutcheon, A. (1987). *Latent class analysis*. Thousand Oaks, CA: Sage.

McGill, C., Gerst, E., Isquith, P., & Gioia, G. (2011). Evidence of validity for a monitoring version of the Behavior Rating Inventory of Executive Function (BRIEF). *Journal of the International Neuropsychological Society*, *17*(s1), 297.

Mittenberg, W., DiGiulio, D.V., Perrin, S., & Bass, A.E. (1992). Symptoms following mild head injury: Expectation as aetiology. *Journal of Neurology, Neurosurgery, and Psychiatry*, *55*(3), 200–204.

Newman, J.B., Reesman, J.H., Vaughan, C.G., & Gioia, G.A. (2013). Assessment of processing speed in children with mild TBI: A "first look" at the validity of pediatric ImPACT. *The Clinical Neuropsychologist*, *27*(5), 779–793. http://doi.org/10.1080/13854046.2013.789552

Pepe, M.S. (2004). *The statistical evaluation of medical tests for classification and prediction*. New York, NY: Oxford University Press.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from, http://www.R-project.org/

Ransom, D.M., Vaughan, C.G., Pratson, L., Sady, M.D., McGill, C.A., & Gioia, G.A. (2015). Academic effects of concussion in children and adolescents. *Pediatrics*, *135*, 1043–1050. http://doi.org/10.1542/peds.2014-3434

Rice, M.E., & Harris, G.T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*(5), 615–620. http://doi.org/10.1007/s10979-005-6832-7

Sady, M.D., McGill, C.A., Gerst, E.H., & Gioia, G.A. (2013). Standardized assessment of cognitive exertion in mTBI and non-injured children. *Journal of the International Neuropsychological Society*, *19*(S1), 194. http://doi.org/10.1017/S1355617713000362

Sady, M.D., Vaughan, C.G., & Gioia, G.A. (2011). School and the concussed youth: Recommendations for concussion education and management. *Physical Medicine and Rehabilitation Clinics of North America*, *22*(4), 701–719, ix. http://doi.org/10.1016/j.pmr.2011.08.008

Sady, M.D., Vaughan, C.G., & Gioia, G.A. (2014). Psychometric characteristics of the postconcussion symptom inventory in children and adolescents. *Archives of Clinical Neuropsychology*, *29*(4), 348–363. http://doi.org/10.1093/arclin/acu014

Schatz, P., & Maerlender, A. (2013). A two-factor theory for concussion assessment using ImPACT: Memory and speed. *Archives of Clinical Neuropsychology*, *28*(8), 791–797. http://doi.org/10.1093/arclin/act077

Straus, S.E., Glasziou, P., Richardson, W.S., & Haynes, R.B. (2011). *Evidence-based medicine: How to practice and teach it* (4th ed.), Edinburgh: Churchill Livingstone.

Venkatraman, E.S., & Begg, C.B. (1996). A distribution-free procedure for comparing receiver operating characteristic curves from a paired experiment. *Biometrika*, *83*(4), 835–848. http://doi.org/10.1093/biomet/83.4.835

Vermunt, J., & Magidson, J. (2005). *Latent GOLD Choice 4.0 User's Guide*. Belmont, MA: Statistical Innovations Inc.

Youngstrom, E.A. (2013). Future directions in psychological assessment: Combining evidence-based medicine innovations with psychology's historical strengths to enhance utility. *Journal of Clinical Child and Adolescent Psychology*, *42*(1), 139–159. http://doi.org/10.1080/15374416.2012.736358

Youngstrom, E.A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology*, *39*, 204–221. http://doi.org/10.1093/jpepsy/jst062

Youngstrom, E.A., Meyers, O., Youngstrom, J.K., Calabrese, J.R., & Findling, R.L. (2006). Diagnostic and measurement issues in the assessment of pediatric bipolar disorder: implications for understanding mood disorder across the life cycle. *Developmental Psychopathology*, *18*(4), 989–1021.

Zhou, X.-H., Obuchowski, N., & McClish, D. (2002). *Statistical methods in diagnostic medicine*. New York: Wiley. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/1541-0420.00266/abstract