


Influence of Special Education, ADHD, Autism, and Learning Disorders on ImPACT Validity Scores in High School Athletes

Julia E. Maietta¹ , Kimberly A. Barchard¹, Hana C. Kuwabara¹, Bradley D. Donohue¹, Staci R. Ross², Thomas F. Kinsora² and Daniel N. Allen^{1,*}

¹Department of Psychology, University of Nevada, Las Vegas, NV, USA

²Center for Applied Neuroscience, Las Vegas, NV, USA

(RECEIVED May 28, 2020; FINAL REVISION September 2, 2020; ACCEPTED September 17, 2020; FIRST PUBLISHED ONLINE December 9, 2020)

Abstract

Objective: The Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) is commonly used to assist with post-concussion return-to-play decisions for athletes. Additional investigation is needed to determine whether embedded indicators used to determine the validity of scores are influenced by the presence of neurodevelopmental disorders (NDs). **Method:** This study examined standard and novel ImPACT validity indicators in a large sample of high school athletes ($n = 33,772$) with or without self-reported ND. **Results:** Overall, 7.1% of athletes' baselines were judged invalid based on standard ImPACT validity criteria. When analyzed by group (healthy, ND), there were significantly more invalid ImPACT baselines for athletes with an ND diagnosis or special education history (between 9.7% and 54.3% for standard and novel embedded validity criteria) when compared to athletes without NDs. ND history was a significant predictor of invalid baseline performance above and beyond other demographic characteristics (i.e., age, sex, and sport), although it accounted for only a small percentage of variance. Multivariate base rates are presented stratified for age, sex, and ND. **Conclusions:** These data provide evidence of higher than normal rates of invalid baselines in athletes who report ND (based on both the standard and novel embedded validity indicators). Although ND accounted for a small percentage of variance in the prediction of invalid performance, negative consequences (e.g., extended time out of sports) of incorrect decision-making should be considered for those with neurodevelopmental conditions. Also, reasons for the overall increase noted here, such as decreased motivation, "sandbagging", or disability-related cognitive deficit, require additional investigation.

Keywords: Sports medicine, Neurodevelopmental disorders, Return to sport, Neuropsychological tests, Athletic injuries, Youth sports

INTRODUCTION

Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT; Lovell, 2018) is the most commonly utilized computerized battery for the assessment of cognitive abilities pre- and post-concussion (Kerr et al., 2015). ImPACT is used to track the severity of sports concussion and monitor recovery to make return-to-play decisions following concussive injury. ImPACT includes the assessment of symptoms associated with concussion using self-report, as well as performance-based neurocognitive abilities. These neurocognitive abilities are evaluated using six different

cognitive subtests that contribute to five composite scores (i.e., Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, and Impulse Control). Clinical use of ImPACT has become increasingly common in all levels of sport participation, including professional (e.g., NFL, NHL), collegiate (e.g., NCAA), and high school sports, which are the focus of this study.

Baseline (pre-concussion) and post-concussion assessment are used for within-athlete comparisons in the event of a concussion. Because baseline testing establishes a data point for interpreting the magnitude of decline in cognitive functioning following a concussion, it is important that baseline scores are valid. There is much evidence that these scores are susceptible to response sets, with the (albeit relatively uncommon) dissimulation of cognitive deficit or *sandbagging* being of particular concern (Erdal, 2012; Higgins et al., 2017). For this reason, ImPACT includes a number of embedded

*Correspondence and reprint requests to: Daniel N. Allen, Ph.D., Department of Psychology, University of Nevada Las Vegas, Box 455030, 4505 Maryland Parkway, Las Vegas, NV 89154-5030, USA. Tel.: +1 702 895 0121; Fax: +1 702 895 0195. Email: daniel.allen@unlv.edu

validity indicators that are used to determine whether athletes exerted sufficient effort to produce valid baseline testing scores. In addition to sandbagging, contributors to variable effort identified in the literature include sleep disturbance the night before testing, general lack of appreciation of the importance of baseline testing, and unintentional fluctuations of effort due to other factors (Erdal, 2012; Higgins et al., 2017; McClure et al., 2013; Rabinowitz et al., 2015; Schatz et al., 2017; Walton et al., 2018).

Recent evidence also suggests that high school and collegiate athletes with a history of attention-deficit/hyperactivity disorder (ADHD) and/or academic difficulties (e.g., special education, learning disorder; LD) perform more poorly on ImPACT neurocognitive composite scores than their non-neurodevelopmental diagnosis peers. For example, Elbin et al. (2013) reported that a large sample ($n = 938$) of high school and collegiate athletes with ADHD, LD, or both (ADHD/LD) demonstrated lower performance on all ImPACT composite scores compared to athletes without NDs. Kaye et al. (2019) also found that Division I and club sport athletes with a history of ADHD performed significantly worse on ImPACT Verbal Memory and Visual Memory Composites compared to athletes without neurodevelopmental diagnoses. Gardner et al. (2017) also reported that athletes (aged 10–21 years old) with ADHD performed significantly worse than matched controls on ImPACT Verbal Memory, Visual Memory, Visual Motor Speed, and Reaction Time composite scores at baseline and on post-concussion testing. Others also report similar findings for athletes with ADHD and LD (Kaye et al., 2019; Manderino & Gunstad, 2018a; Peltonen et al., 2019; Poysophon & Rao, 2018; Salinas et al., 2016). Some suggest that athletes with neurodevelopmental disorders (NDs) demonstrate lower performance possibly due to the underlying attention and reading requirements of computerized neurocognitive testing (Lovell, 2018; Schatz et al., 2012). This is important because current embedded validity indicators use cut scores that identify lower than expected levels of performance based on healthy, non-ADHD/LD samples. Therefore, use of these cut scores may result in increased estimates of invalid baseline performance when applied to those with neurodevelopmental conditions. Little research is available that examines this matter, but the studies that are available support this concern (Manderino & Gundstad, 2018a; Manderino et al., 2019).

Manderino and Gunstad (2018a) examined ImPACT standard embedded validity indicators in NCAA collegiate athletes with ADHD ($n = 65$), academic difficulties ($n = 53$), or ADHD and academic difficulty ($n = 19$). When compared to a control sample with no history of ADHD or academic difficulty, the ADHD groups were more likely to have invalid baseline scores. In a follow-up study examining NCAA athletes, Manderino et al. (2019) found that lowering the criterion scores for novel embedded validity indicators (Schatz & Glatts, 2013) resulted in lower rates of invalid baseline scores for those with ADHD, academic difficulty, and ADHD with academic difficulty. They provide adjusted cut

scores for these athletes that might be used in clinical practice to help decrease the identification of valid profiles as invalid. In addition to Schatz and Glatts (2013), other authors have proposed novel validity indicators (e.g., Higgins et al., 2017) all of which are designed to improve the detection of invalid scores. A recent systematic review by Gaudet and Weyandt (2017) indicates that these and other novel embedded indicators do increase the sensitivity of ImPACT to underperformance in analogue studies. The research suggesting increased invalid baseline performance associated with neurodevelopmental diagnosis history (ND) is concerning for obvious reasons, possibly most importantly, because it limits the accuracy of return-to-play decisions by increasing the likelihood of a valid baseline being incorrectly interpreted as unusable (invalid). In these instances, valid information which could be used is not considered in making return-to-play decisions, so an athlete may be held out of play for longer than necessary or returned to play too soon, which could result in a number of negative consequences such as another injury (i.e., second impact) or more severe post-concussive symptoms (Castile et al., 2012; Chrisman et al., 2013; Covassin et al., 2008).

Based on these findings, the current study aims to investigate the following research questions: (1) Do the standard and novel embedded validity indicators identify higher rates of invalid baseline performance in high school athletes with ND compared to those with no ND history? (2) Based on previous literature suggesting age (Abeare et al., 2018), gender (Schatz et al., 2012), and sport-based differences in invalid baseline performance, does the addition of ND predict invalid performance above and beyond these other demographic variables? In addition to these questions, multivariate base rates of criterion failure on validity indicators will be presented for athletes with neurodevelopmental history based on the current ImPACT embedded indicators. Currently, there is no research reporting multivariate base rates for ImPACT validity indicators in athletes with a history of ND. Multivariate base rates allow for consideration of the clinical significance of low scores across multiple tests. Multivariate, rather than a single test score, analysis aids in clinical interpretation of neuropsychological tests results by helping to protect against overinterpretation of one or two low scores within a battery (e.g., Brooks et al., 2013; Cook et al., 2019; Holdnack, 2019), since normal non-impaired populations often obtain one score in the impaired range (Binder et al., 2009). We examine multivariate base rates in this study, because they may be particularly useful in protecting against overinterpretation of ImPACT validity indicators in athletes with a history of ND. Table 1 provides an overview of the indicators examined in this study.

METHODS

Participants

Participants included 33,772 high school athletes aged 13–19 (mean age = 15.0, $SD = 1.2$; 43.7% female; mean education = 9.0, $SD = 1.4$) from a larger longitudinal, statewide

Table 1. ImPACT validity criteria used for comparisons

Source	Variable	Subtest score description	Cut score
ImPACT standard	X's and O's Total Incorrect	Number incorrect on a spatial memory task.	>30
	Impulse Control Composite	Composite includes number incorrect on a choice reaction time task (X's and O's distractor) and number of commissions on a Stroop task (Color Match).	>30
	Word Memory Learning % Correct	% of correct hits and distractors on the immediate recall portion of a word memory task.	<69%
	Design Memory Learning % Correct	% of correct hits and distractors on immediate recall portion of a design memory task.	<50%
	Three Letters Total Letters Correct	Number of letters correctly recalled (out of three) after 18-s delay across five trials.	<8
Higgins, Denney, and Maerlender (2017)	Word Memory Learning % Correct	% of correct hits and distractors on the immediate recall portion of a word memory task.	≤90%
	Word Memory Delayed Memory % Correct	% of correct hits and distractors on the 20-min delay portion of a word memory task.	≤81%
	Design Memory Total % Correct	% of correct hits and distractors on immediate recall portion of a design memory task.	≤64.75%
	X's and O's Total Correct (interference)	Number incorrect on a choice reaction time task (distractor for the X's and O's spatial memory task).	≤107.5
Manderino et al. (2019)	Word Memory Correct Distractors*	Number of correct distractors (out of 24) from both the immediate and delayed recall trials.	<18
	Design Memory Correct Distractors*	Number of correct distractors (out of 24) from both the immediate and delayed recall trials.	<10

Note. % = percentage.

*Immediate and delayed conditions are summed for this cut score.

sample of athletes of who completed ImPACT to establish a pre-concussion baseline prior to beginning their respective sport season between 2008 and 2016. Athletes from the larger longitudinal sample were not included in this study if they reported a history of epilepsy, brain surgery, meningitis, treatment for substance and/or alcohol use, and treatment for a psychiatric disorder (e.g., depression, anxiety). The athletes were then separated into the following distinct categories based on self-reported diagnostic history: ADHD only (3.8%), LD only (1.5%), comorbid ADHD/LD (0.7%), autism (inclusive of comorbid autism with ADHD/LD; 0.2%), special education history with no diagnosis reported (SpED; 1.3%), and healthy athletes with no diagnosis (92.5%). Overall, 7.1% of the baselines were judged to be invalid based on the standard embedded ImPACT indicators that flag an athlete's report as "Baseline++". Sport type was classified as collision, contact, limited contact, or no contact based on previous research denoting these categories for each sport reported (Brett & Solomon, 2017, Rice, 2008). Collision sports (e.g., football) are those with a purposeful collision with other players/objects and consisted of 33.3% of the sample. Contact sports (e.g., soccer) involve routine contact with other players/objects but no purposeful collision (31.5% of the sample). Limited contact sports (e.g., volleyball) have less frequent contact with other players/objects and (18.9%), and noncontact sports (e.g., golf) have rare or unpredicted contact with other players/objects (16.4%).

Overall, 8.7% of the sample had a self-reported concussion history. Analyses were performed with and without athletes who report a concussion history, with the same substantive conclusions found both ways. Given this, the results reported here are from the full sample (including those with concussion history).

Measure

ImPACT utilizes six subtests – Word Memory, Design Memory, X's and O's, Symbol Match, Color Match, and Three Letters – to assess various aspects of cognitive performance. Scores from these subtests are averaged to form the five composite scores – Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, and Impulse Control. In clinical settings, four of these composite scores (Verbal Memory, Visual Memory, Visual Motor Speed, and Reaction Time) are used to track the acute effects of the concussion as well as the speed of recovery following concussive injury. ImPACT has demonstrated adequate reliability with test-retest reliability with intraclass correlation coefficients between .62 and .85 (Elbin et al., 2011).

Athlete performance is considered to be invalid if one of the athlete's scores falls above (or below depending on the cut score) any of the five standard ImPACT embedded invalidity indicators. See Table 1 for an overview of the indicators used in this study.

Procedure

ImPACT was administered by trained school personnel in group settings at the high schools where the athletes participated in their sports. School personnel was trained in how to properly administer ImPACT (e.g., keep a quiet testing environment, etc.) and standard instructions were utilized prior to the administration of the test (ImPACT includes instructions on the screen for the athlete). Baseline assessment was conducted in groups with spaces between each athlete completing testing. This study utilized de-identified archival data, which was deemed exempt by the local social/behavioral institutional review board for the protection of human subjects.

Statistical Analyses

To answer the first research question (Are baseline scores for athletes with a history of ND more likely to be flagged as invalid?), chi-square analysis was used to compare the rate of baseline performance that is deemed invalid between groups using standard ImPACT embedded indicators as well as the two additional novel indicators (Higgins et al., 2017; Manderino et al., 2019). *Post hoc* comparisons were made using *Z* tests of two proportions with Benjamini and Hochberg (1995) correction to reduce the likelihood of Type 1 error due to multiple comparisons (the false discovery rate was set at .05). Odds ratios were computed for each of the diagnoses that were significantly different ($p < .05$) on *post hoc* testing.

To answer the second research question (Do demographic characteristics such as age, sex, sport, and neurodevelopmental history predict performance that is flagged as invalid?), a binomial logistic regression was conducted to determine if demographic characteristics (age, sex, sport, and neurodevelopmental history) predicted performance deemed invalid based on the standard ImPACT indicators.

Additionally, multivariate base rates were computed for the standard ImPACT indicators and were stratified by age, sex, and neurodevelopmental history. Multivariate base rates were computed following the dichotomization of clinical diagnosis (i.e., those with or without a history of ADHD or non-ADHD neurodevelopmental condition).

RESULTS

Frequency of Invalid Neurodevelopmental Baselines

Table 2 presents the results of the chi-square tests. *Post hoc* analyses and the odds ratios for each significant comparison are presented in Table 3.

Standard ImPACT indicators

The rate of baseline performance that was flagged as invalid differed significantly based on diagnosis history ($\chi^2 = 128.95$,

Table 2. Chi-square results for invalid baselines

Validity criteria	χ^2 (df)	Diagnostic group	Invalid profiles (%)
ImPACT standard	128.95 (5)*	Healthy	6.7
		ADHD only	9.7
		LD only	15.1
		ADHD/LD	16.2
		Autism	12.9
Higgins et al. (2017)	206.34 (5)*	SpED	13.2
		Healthy	32.7
		ADHD only	41.5
		LD only	49.3
		ADHD/LD	54.3
Manderino et al. (2019)	156.22 (5)*	Autism	37.1
		SpED	50.8
		Healthy	13.1
		ADHD only	17.7
		LD only	24.0
		ADHD/LD	27.8
		Autism	21.0
		SpED	24.5

Abbreviations: ADHD = attention-deficit/hyperactivity disorder; LD = learning disorder; ADHD/LD = comorbid ADHD and LD; SpED = special education history with no other neurodevelopmental disorder reported.

* $p < .001$.

$p < .001$). *Post hoc* analyses revealed that when compared to athletes with no ND history, the baseline scores of athletes with a history of ADHD, LD, comorbid ADHD/LD, or SpED were significantly more likely to be flagged as invalid based on standard embedded ImPACT validity indicators (i.e., one or more cut score failure). As seen in Table 3, athletes with ADHD had more invalid baselines than did healthy athletes, but they also displayed significantly fewer invalid baselines than those with LD or ADHD/LD. All other groups did not differ significantly from each other.

Higgins et al. (2017) indicators

The rate of baseline performance that was flagged as invalid differed significantly based on diagnosis history ($\chi^2 = 206.34$, $p < .001$). *Post hoc* analyses revealed that, when compared to athletes without ND history, athletes with ADHD, LD, comorbid ADHD/LD, or SpED were significantly more likely to be flagged as invalid based on the failure of one or more of the Higgins et al. (2017) cut score indicators. With the exception of the autism group, all of the ND groups had significantly higher rates of invalid baseline performance compared to controls, with the ADHD group having significantly more invalid baselines scores than healthy athletes, but significantly fewer invalid baselines than the LD, ADHD/LD, or SpED groups (see Table 3). The autism group had significantly fewer invalid baselines than the ADHD/LD group, but did not differ

Table 3. *Post hoc* comparisons by neurodevelopmental diagnosis

Validity criteria	Comparison	Odds ratio	Comparison	Odds ratio	Comparison	Odds ratio
ImPACT standard	Healthy to ADHD	1.5***	ADHD to LD	1.7	LD to ADHD/LD	1.1
	Healthy to LD	2.5***	ADHD to ADHD/LD	1.8**	LD to autism	0.8
	Healthy to ADHD/LD	2.7***	ADHD to autism	1.4	LD to SpED	0.9
	Healthy to Autism	2.1	ADHD to SpED	1.4	Autism to ADHD/LD	1.3
	Healthy to SpED	2.1***	SpED to ADHD/LD	1.3	Autism to SpED	1.0
Higgins et al. (2017)	Healthy to ADHD	1.5***	ADHD to LD	1.4**	LD to ADHD/LD	1.2
	Healthy to LD	2.0***	ADHD to ADHD/LD	1.7	LD to Autism	0.6
	Healthy to ADHD/LD	2.4***	ADHD to autism	0.8	LD to SpED	1.1
	Healthy to Autism	1.2	ADHD to SpED	1.5	Autism to ADHD/LD	2.0*
	Healthy to SpED	2.1***	SpED to ADHD/LD	1.1	Autism to SpED	1.8
Manderino et al. (2019)	Healthy to ADHD	1.4***	ADHD to LD	1.5**	LD to ADHD/LD	1.2
	Healthy to LD	2.1***	ADHD to ADHD/LD	1.8***	LD to autism	0.8
	Healthy to ADHD/LD	2.5***	ADHD to autism	1.2	LD to SpED	1.0
	Healthy to Autism	1.8	ADHD to SpED	1.5**	Autism to ADHD/LD	1.4
	Healthy to SpED	2.1***	SpED to ADHD/LD	1.2	Autism to SpED	1.2

Note. Odds ratios are only presented for those *post hoc* comparisons that were significant. Significance levels were adjusted based on the Benjamini–Hochberg correction for multiple comparisons.

Abbreviations: ADHD = attention-deficit/hyperactivity disorder; LD = learning disorder; ADHD/LD = comorbid ADHD and LD; SpED = special education history with no other neurodevelopmental disorder reported.

p* < .05. *p* < .01. ****p* < .001.

Table 4. Binomial logistic regression results

Variable	<i>B</i>	SE	Wald	<i>df</i>	<i>p</i>	Odds ratio	95% CI for odds ratio
Age	-.10	.02	29.65	1	.000	0.91	0.87–0.94
Gender (male)	.13	.05	6.58	1	.010	1.14	1.03–1.26
Sport category			17.50	3	.001		
Collision	.23	.07	10.10	1	.001	1.25	1.09–1.45
Contact	.28	.07	16.21	1	.000	1.33	1.16–1.52
Limited contact	.26	.08	11.51	1	.001	1.30	1.12–1.51
ND history			114.02	5	.000		
ADHD	-.39	.10	15.72	1	.000	0.67	0.56–0.82
LD	-.89	.13	48.79	1	.000	0.41	0.32–0.53
Autism	-.83	.38	4.64	1	.031	0.44	0.21–0.93
SpED	-.76	.14	28.05	1	.000	0.47	0.35–0.62
ADHD/LD	-.95	.18	27.64	1	.000	0.39	0.27–0.55

Note. Gender is for males compared to females. For sport category, noncontact sports are the reference group. ND history = neurodevelopmental diagnosis. For ND history, healthy controls are the reference group.

Abbreviations: ADHD = attention-deficit/hyperactivity disorder; LD = learning disorder; ADHD/LD = comorbid ADHD and LD; SpED = special education history with no other neurodevelopmental disorder reported.

significantly from any other groups. All other groups did not differ significantly from each other.

Manderino et al. (2019) indicators

The rate of baseline performance that was flagged as invalid differed significantly based on diagnosis history ($\chi^2 = 128.95$, $p < .001$). *Post hoc* analyses revealed that when compared to non-ND athletes, those with ADHD, LD, comorbid ADHD/LD, or SpED were significantly more likely to be flagged as invalid based on failure of one or more of the Manderino et al. (2019) cut score indicators. The athletes with

ADHD had significantly more invalid baselines than did healthy athletes, but they also displayed significantly fewer invalid baselines than the LD, ADHD/LD, or SpED groups (see Table 3). All other groups did not differ significantly from each other.

Prediction of Invalid Baselines Using Demographic Variables

To assess the effect of demographic factors (i.e., age, sex, sport, and having a neurodevelopmental diagnosis) on the prediction of performance that is deemed invalid (using

standard ImPACT indicators), a binomial logistic regression was conducted. The full model (using age, sex, sport, and neurodevelopmental history to predict invalid performance) was statistically significant ($\chi^2(10) = 155.75, p < .001$), but explained only 1.2% of the variance (Nagelkerke R^2). All four predictor variables were significant (see Table 4). Compared to females, males had 1.14 times higher odds to have performance flagged invalid. Decreasing age was associated with increased odds of invalid baseline. Collision sports had 1.25 times higher odds of invalid performance relative to non-contact sports. Logistic regression indicated neurodevelopmental diagnoses were associated with decreased rates of invalid performance; however, this result is based on the number of invalid tests rather than the proportion of invalid tests within each diagnostic group (which a chi-square test would assess as we have above). Specifically, each neurodevelopmental group's sample size gets progressively smaller (as does the number of invalid cases), but the relative proportion of invalid to valid cases increases as described above in the chi-square results. Irrespective of this specific interpretation of the Wald Test and the odds ratio of receiving invalid performance, neurodevelopmental diagnosis, overall, was still predictive in the binary logistic regression model ($p < .001$).

Multivariate Base Rates

Multivariate base rates are presented for the standard embedded ImPACT validity indicators. Currently, ImPACT automatically flags athletes who fall below or above the cut-off point for one or more indicators (Table 1 specifies these cut scores). In order to maximize sample size to increase the stability of multivariate base rates for the athletes with ND, we combined athletes across non-ADHD ND diagnoses and stratified rates for non-ADHD ND and healthy, non-ND athletes (see Table 5). Due to differences between ADHD and other ND groups, multivariate base rates for ADHD athletes are presented separately in Table 6. To determine stratification of age for multivariate base rates, chi-square tests (with *post hoc* analyses using the Benjamini-Hochberg correction) were conducted to assess significant differences between age groups (e.g., the proportion of failed cut scores in 13- vs. 14-year old, 14- vs. 15-year old, etc.). Results revealed that the rate of cut score failure differed between age groups ($\chi^2 = 213.08, p < .001$). *Post hoc* analyses indicated that rates of cut score failure did not differ for the following age groups – 13- and 14-year-old athletes ($p > .05$), 15- and 16-year-old athletes ($p > .05$), and 17- and 18-year-old athletes ($p > .05$). The following age groups did differ significantly in terms of frequency of cut score failure – 14- and 15-year-old athletes ($p < .001$) and 16- and 17-year-old athletes ($p < .05$). A chi-square test (with Bonferroni correction for Type 1 error) also indicated that stratification for multivariate base rates was needed for male *versus* female athletes ($\chi^2 = 39.30, p < .001$). Due to a low number of athletes ($n = 34$), the 19-year-old age group was excluded from these analyses. The 17- to 18-year-old females with non-ADHD

Table 5. Multivariate base rates of invalid cut score failure for LD, ADHD/LD, and special education history neurodevelopmental groups

NVIF	Aged 13–14				Aged 15–16				Aged 17–18			
	Male		Female		Male		Female		Male		Female	
	Healthy (<i>n</i> = 7831)	ND (<i>n</i> = 329)	Healthy (<i>n</i> = 6689)	ND (<i>n</i> = 183)	Healthy (<i>n</i> = 6990)	ND (<i>n</i> = 326)	Healthy (<i>n</i> = 5740)	ND (<i>n</i> = 200)	Healthy (<i>n</i> = 2375)	ND (<i>n</i> = 129)	Healthy (<i>n</i> = 1589)	ND (<i>n</i> = 54)
1	5.0%	10.9%	4.8%	10.4%	5.8%	10.4%	5.6%	10.5%	6.0%	12.4%	5.1%	5.6%
2	0.8%	1.2%	0.7%	2.2%	1.1%	4.0%	1.0%	3.0%	1.5%	2.3%	1.0%	3.8% ^c
3	0.3%	1.1% ^a	0.0%	0.0%	0.3%	2.1%	0.1%	1.5%	0.7%	2.0% ^b	0.4%	1.9%
4+	0.1%	0.9%	0.0%	0.0%	0.1%	0.6%	0.1%	0.0%	0.6%	1.6%	0.2%	0.0%

Note. Multivariate base rates are presented as a percentage of the subsample that displayed a particular number of invalid embedded validity indicators. Frequencies presented are based on the failure of standard ImPACT validity criteria.
 ND = neurodevelopmental history including LD, ADHD/LD, or special education history with no other neurodevelopmental disorder reported (ADHD is presented separately in Table 6). NV IF = number of validity indicators failed. In some cases, for the neurodevelopmental groups, cells were empty or values did not conform to the expectation of decreasing frequency with an increasing number of validity indicators. In these cases, the data was smoothed and cell values were replaced with estimates based on the average decrease expected based on adjacent cells.
^a 0.3 replaced with 1.1.
^b 0.0 replaced with 2.0.
^c 0.0 replaced with 3.8.

Table 6. Multivariate base rates of invalid cut score failure for ADHD group

NVIF	Aged 13–14				Aged 15–16				Aged 17–18			
	Male		Female		Male		Female		Male		Female	
	Healthy (n = 7831)	ADHD (n = 442)	Healthy (n = 6689)	ADHD (n = 132)	Healthy (n = 6990)	ADHD (n = 429)	Healthy (n = 5740)	ADHD (n = 115)	Healthy (n = 2375)	ADHD (n = 124)	Healthy (n = 1589)	ADHD (n = 29)
1	5.0%	5.9%	4.8%	5.3%	5.8%	7.0%	5.6%	5.2%	6.0%	7.3%	5.1%	10.3%
2	0.8%	2.3%	0.7%	1.5%	1.1%	3.3%	1.0%	1.7%	1.5%	2.4%	1.0%	6.9% ^a
3	0.3%	0.2%	0.0%	0.0%	0.3%	0.9%	0.1%	1.7%	0.7%	0.8%	0.4%	3.4%
4+	0.1%	0.2%	0.0%	0.0%	0.1%	0.0%	0.1%	0.0%	0.6%	0.8%	0.2%	0.0%

Note. Multivariate base rates are presented as a percentage of the subsample that displayed a particular number of invalid embedded validity indicators. Frequencies presented are based on the failure of standard ImpACT validity criteria. NV IF = number of validity indicators failed. In some cases, for the neurodevelopmental groups, cells were empty or values did not conform to the expectation of decreasing frequency with an increasing number of validity indicators. In these cases, the data was smoothed and cell values were replaced with estimates based on the average decrease expected based on adjacent cells.
^a 0.0 replaced with 6.9.

ND history (*n* = 54) and ADHD-only history (*n* = 29) had small sample sizes. Due to this, the multivariate base rates for these groups are reported but should be interpreted with caution. In some cases, for the neurodevelopmental groups, cells were empty or values did not conform to the expectation of decreasing frequency with an increasing number of validity indicators. In these cases, the data was smoothed and cell values were replaced with estimates based on the average decrease expected based on adjacent cells.

Table 5 presents the multivariate base rates for the failure of standard embedded ImpACT validity indicators stratified by age, gender (where appropriate), and non-ADHD ND history (separate multivariate base rates are presented for the athlete with ADHD only in Table 6). For healthy athletes across all age stratifications and genders, the percentage of failed validity cut scores ranged from 0.1% (four or more cut scores failed) to 6.0% (one cut score failed). For athletes with LD, ADHD/LD, or SpED history across the groups, the percentage of failed validity cut scores ranged from 0.0% (four or more cut scores failed) to 12.4% (one cut score failed). For athletes with ADHD, the percentage of failed validity cut scores ranged from 0.0% (four or more cut scores failed) to 10.3% (one cut score failed). On average (across all NDs), having a neurodevelopmental diagnosis was associated with a 146% increase in the frequency of failed cut scores compared to healthy athletes.

DISCUSSION

Results of the current study demonstrate that athletes who have a history of ADHD, LD, comorbid ADHD/LD, or SpED history are significantly more likely than healthy athletes to obtain ImpACT baseline scores that are flagged as invalid by the standard ImpACT validity indicators as well as two novel indicators proposed in the literature (Higgins et al., 2017; Manderino et al., 2019). Athletes with ADHD have more baselines flagged as invalid compared to healthy controls, but fewer when compared to those with LD, ADHD/LD, or SpED history, suggesting that there may be factors associated with ADHD that resulted in somewhat spared performance across all of the validity indicators we examined. For example, individuals with ADHD commonly take medications to ameliorate their cognitive and behavioral symptoms, which may result in fewer interfering effects on ImpACT validity indicators when compared to the cognitive deficits (e.g., reading impacts) for the other neurodevelopmental groups. Prior studies report mixed findings regarding medication use’s effects on ImpACT performance (Cook et al., 2019; Gardner et al., 2017; Polysophon & Rao, 2018). Although medication data was not available in this study, it is expected that most athletes who reported a history of ADHD were also taking medications to improve attention, which contrasts with the untreated cognitive deficits associated with LD. These deficits may have suppressed overall performance on the ImpACT, which caused the increased frequency of baseline performance to be flagged as invalid in the

ADHD/LD and LD groups. It is also important to note that we did not know the type of learning disability experienced by the athletes in this study, although language-based LDs are the most common in the general population (American Psychiatric Association, 2013) and in children diagnosed with ADHD (Mayes & Calhoun, 2006; Parke et al., 2015). Future research comparing athletes with different types of confirmed LDs would provide more definitive evidence regarding the influence of specific deficits on rates of invalid baseline performance.

While the overall sample in this study had an invalid baseline rate of 7.1%, when assessed by diagnostic group, athletes with NDs were between 1.2 and 2.7 times as likely as healthy athletes to have their baseline performance flagged as invalid, with between 9.7% and 54.3% flagged. Comorbid ADHD/LD consistently demonstrated the highest rates of invalid baselines across the standard ImPACT, Higgins, and Manderino indicators. The large discrepancies found between the validity indices in this study are consistent with a prior investigation that have found the Higgins indicators to identify substantially more baselines as invalid than the standard ImPACT indicators in naturalistic samples (Abeare et al., 2018). Because the standard ImPACT and Higgins indicators were created using healthy athletes, it may not be particularly surprising that athletes with NDs are more likely to have performance that is flagged as invalid, not necessarily because of performance validity issues, but possibly due to underlying attention and reading requirements of ImPACT (Lovell, 2018; Schatz et al., 2012). Given that the Manderino indicators were created specifically for people with ADHD and those with academic difficulty, it is surprising, however, that athletes with these problems are still more likely than healthy athletes to have baseline performance flagged as invalid. Although these indicators (Higgins et al., 2017; Manderino et al., 2019) have demonstrated higher positive predictive value than the standard ImPACT embedded validity indicators (Higgins et al., 2017; Manderino & Gunstad 2018b), it is concerning that the baselines for athletes with ND are still disproportionately identified as invalid. These results are consistent with recent literature suggesting high rates of invalid baselines among even athletes without ND in naturalistic and coached (i.e., when athletes are instructed to “fake bad” in laboratory settings) samples. Some of these studies report that between 28% and 83% of athletes demonstrate baselines performance that is flagged as invalid (Abeare et al., 2018; Gaudet & Weyandt, 2017; Raab et al., 2020). It has also been suggested that current standard embedded validity indicators may miss up to 20% or more of invalid baselines, which is obviously much higher than is ideal for effective management of sport concussions and well-founded return-to-play decisions (Gaudet & Weyandt, 2017; Raab et al., 2020). These rates are concerning given the potential complicating implications for clinical practice (including the need to reassess athletes who have an initial invalid performance; Schatz et al., 2014) because they diminish the utility of baseline comparison scores. The current study, in conjunction with previous literature (Manderino & Gunstad, 2018a; Manderino et al., 2019), suggests that the current method of identifying invalid baselines may not be appropriate for athletes

with ND history. It is also possible that rates of invalid performance may be inherently increased in ND populations; however, additional research is needed using coached samples to examine the root of these issues.

Indeed, while athletes with ND are more likely to have performance flagged as invalid, the factors that contribute to this increase are not well understood. Our examination of key demographic factors indicated that age, sex, sport, and ND were predictive of invalid performance in a binomial logistic regression model. While significant predictors, the combination of these demographic factors accounted for a very small portion of the variance (1.2%) of invalid baseline performance suggesting that there are other more substantive factors to consider that have not yet been identified. One of the most obvious factors that have not been systematically investigated is the influence of cognitive deficits associated with ADHD and ND on increased rates of invalid baseline performance. Other investigators have suggested that performance validity is a multifaceted phenomenon and that factors such as overall cognitive abilities, gender, fatigue, level of supervision during testing, group *versus* individual testing, return-to-play incentives, and sport season can all impact baseline performances (Alsalaheen et al., 2016). Even given this, it is still concerning that neurodevelopmental history was predictive of invalid baseline performance.

Currently, invalidity on ImPACT is based on the failure of one or more of the embedded validity indicators (for both the standard and novel criteria). In the broader literature on performance validity, there has been extensive discussion about the number of failures that should constitute invalid performance (for a very comprehensive review, the interested reader is referred to Lippa, 2018). A number of authors have suggested that failure of one or more tests of performance validity is sufficient to deem a case as invalid (Inman & Berry, 2002; Iverson & Franzen, 1996; Vickery et al., 2004), while other authors argue that one failure is not enough and that two or more failures are required for determination of invalid responding (Binder et al., 2009; Brooks et al., 2013). Indeed, the positive predictive values and specificity associated with utilizing two or more failures result in greater confidence in the designation of a profile as invalid (Chafetz, 2011; Larrabee 2014; Lippa, 2018). Due to normal intraindividual variability in test performance, it has been demonstrated that the more tests performed, the more likely it is that an individual will have one low score based on chance alone, and so utilization of only one cut score to indicate failure would likely result in false-positive errors (Binder et al., 2009; Brooks et al., 2013). In this way, a criterion of two or more performance validity test failures (rather than a single failure alone) certainly makes more sense for identifying an invalid profile (Binder et al., 2009; Brooks et al., 2013).

Results of the current study also coincide with the recommendation that two or more failures may be necessary for invalid performance designation. To our knowledge, our use of multivariate base rates of invalid classification is the first reported in the literature for athletes with ND. Our results demonstrated that (for standard ImPACT indicators) between

4.9% and 5.8% of healthy controls had one score that fell outside the range of cut-off thresholds (see Table 5). When athletes with ADHD only were considered (see Table 6), these rates were between 5.2% and 10.3%, and when other non-ADHD ND groups (i.e., LD, ADHD/LD, or SpED history) were considered, these rates rose to between 5.6% and 12.4%. Multivariate base rates can be used in clinical settings to protect against overinterpretation of a single low score, which commonly occurs in unimpaired individuals within a battery of tests (Brooks et al., 2013; Cook et al., 2019; Holdnack, 2019). As an example, if a 15-year-old male athlete who has comorbid ADHD/LD takes ImPACT and his profile is designated as invalid based on one of the standard indicators, a neuropsychologist could use Table 5 to see that 10.4% of athletes in this demographic group had one invalid indicator score, which suggests that this is a somewhat common phenomenon based on the results of the current study that the standard embedded validity indicators are biased towards identifying this group as invalid more frequently. In this case, it is not clear whether this profile is truly invalid. On the other hand, if this 15-year-old male with comorbid ADHD/LD obtained two invalid indicators (base rate; 4.0%), it is much more likely that his performance is invalid due to this level of failure occurring infrequently. Using a two or more failure criteria for ImPACT scores, in conjunction with the current multivariate base rates, particularly for those with neurodevelopmental conditions, allows for increased clinical confidence in invalid profile designation given that, in most cases, less than 5% of the sample had more than two validity indicators flagged. The use of these two or more criteria (vs. the current one or more failure designation) still needs to be tested in a rigorous experimental investigation so that sensitivity, specificity, and positive predictive values can be provided. However, at this point, it is certainly clear that athletes with neurodevelopmental conditions are much more likely to be flagged as invalid by the current ImPACT validity indicators, which may be more of a testament to the underlying reading and/or attentional difficulties in this population rather than true performance validity concerns. Our multivariate base rate analyses utilized a combined ND sample in order to maximize sample size and increase the stability of the rates for athletes with ND. Given differences in invalid baselines within the ND groups, particularly the ADHD group compared to the other groups, future research with large samples would be quite useful in determining whether MVBRs should be separately calculated for each diagnostic category.

Given that those with NDs make up a large minority of the population of athletes around the country who are tested with ImPACT, it is critical that measurement and interpretation of baseline performance accurately portray cognitive functioning and performance validity in this population. The current study sheds light on the difficulties of invalidity classification in this population; however, future research is still needed to elucidate the differences in cognitive performance in athletes with neurodevelopmental conditions and to update clinical recommendations for baseline interpretation in this

population. As is the case with the provision of separate ImPACT normative data for athletes with LD and ADHD, separate cut scores may need to be developed for more accurate classification of performance validity in groups with these and other NDs. The current study, along with others (Elbin et al., 2013; Gardner et al., 2017; Manderino et al., 2019; Manderino & Gunstad, 2018a; Zuckerman et al., 2013), serve to spark future endeavors that can continue to enhance the measurement and outcomes for athletes with neurodevelopmental conditions worldwide.

ACKNOWLEDGMENTS

No funding was received for this research.

CONFLICT OF INTEREST

The authors have nothing to disclose.

REFERENCES

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association. Retrieved from <https://doi.org/10.1176/appi.books.9780890425596>
- Abear, C.A., Messa, I., Zuccato, B.S., Merker, B., & Erdodi, L. (2018). Prevalence of invalid performance on baseline testing for sport-related concussion by age and validity indicator. *JAMA Neurology*, 75(6), 697–703. <https://doi.org/10.1001/jamaneurol.2018.0031>
- Alsalaheen, B., Stockdale, K., Pechumer, D., & Broglio, S.P. (2016). Validity of the immediate post concussion assessment and cognitive testing (ImPACT). *Sports Medicine*, 46(10), 1487–1501. <https://doi.org/10.1007/s40279-016-0532-y>
- Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57(1), 289–300.
- Binder, L.M., Iverson, G.L., & Brooks, B.L. (2009). To err is human: “Abnormal” neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology*, 24(1), 31–46. <https://doi.org/10.1093/arclin/acn001>
- Brett, B.L. & Solomon, G.S. (2017). Comparison of neurocognitive performance in contact and noncontact nonconcussed high school athletes across a two-year interval. *Developmental Neuropsychology*, 42(2), 70–82. <https://doi.org/10.1080/87565641.2016.1243114>
- Brooks, B.L., Iverson, G.L., & Holdnack, J.A. (2013). Understanding and using multivariate base rates with the WAIS-IV/WMS-IV. In J.A. Holdnack, L.W. Drozdzick, L.G. Weiss, & G.L. Iverson (Eds.), *WAIS-IV/WMS-IV/ACS: Advanced clinical interpretations* (pp. 75–102). Berlin: Elsevier Science.
- Castile, L., Collins, C.L., McIlvain, N.M., & Comstock, R.D. (2012). The epidemiology of new versus recurrent sports concussions among high school athletes, 2005–2010. *British Journal of Sports Medicine*, 46(8), 603–610. <https://doi.org/10.1136/bjsports-2011-090115>
- Chafetz, M. (2011). Reducing the probability of false positives in malingering detection of social security disability claimants.

- The Clinical Neuropsychologist*, 25(7), 1239–1252. <https://doi.org/10.1080/13854046.2011.586785>
- Chrisman, S.P., Rivara, F.P., Schiff, M.A., Zhou, C., & Comstock, R.D. (2013). Risk factors for concussive symptoms 1 week or longer in high school athletes. *Brain Injury*, 27(1), 1–9. <https://doi.org/10.3109/02699052.2012.722251>
- Cook, N.E., Karr, J.E., Brooks, B.L., Garcia-Barrera, M.A., Holdnack, J.A., & Iverson, G.L. (2019). Multivariate base rates for the assessment of executive functioning among children and adolescents. *Child Neuropsychology*, 25(6), 836–858. <https://doi.org/10.1080/09297049.2018.1543389>
- Covassin, T., Stearne, D., Elbin, R. (2008). Concussion history and postconcussion neurocognitive performance and symptoms in collegiate athletes. *Journal of Athletic Training*, 43(2), 119–124. <https://doi.org/10.4085/1062-6050-43.2.119>
- Elbin, R.J., Kontos, A.P., Kegel, E., Johnson, E., Burkhart, S., & Schatz, P. (2013). Individual and combined effects of LD and ADHD on computerized neurocognitive concussion test performance: Evidence for separate norms. *Archives of Clinical Neuropsychology*, 28(5), 476–484. <https://doi.org/10.1093/arclin/act024>
- Elbin, R.J., Schatz, P., & Covassin, T. (2011). One-year test-retest reliability of the online version of ImPACT in high school athletes. *American Journal of Sports and Medicine*, 39(11), 2319–2324. <https://doi.org/10.1177/0363546511417173>
- Erdal, K. (2012). Neuropsychological testing for sports-related concussion: How athletes can sandbag their baseline testing without detection. *Archives of Clinical Neuropsychology*, 27(5), 473–479. <https://doi.org/10.1093/arclin/acs050>
- Gardner, R.M., Yengo-Kahn, A., Bonfield, C.M., & Solomon, G.S. (2017). Comparison of baseline and post-concussion ImPACT test scores in young athletes with stimulant-treated and untreated ADHD. *The Physician and Sportsmedicine*, 45(1), 1–10. <https://doi.org/10.1080/00913847.2017.1248221>
- Gaudet, C.E. & Weyandt, L.L. (2017). Immediate post-concussion and cognitive testing (ImPACT): A systematic review of the prevalence and assessment of invalid performance. *The Clinical Neuropsychologist*, 31(1), 43–58. <https://doi.org/10.1080/13854046.2016.1220622>
- Higgins, K.L., Denney, R.L., Maerlender, A. (2017). Sandbagging on the immediate post-concussion assessment and cognitive testing (ImPACT) in a high school athlete population. *Archives of Clinical Neuropsychology*, 32(3), 259–266. <https://doi.org/10.1093/arclin/acw108>
- Holdnack, J.A. (2019). The development, expansion, and future of the WAIS-IV as a cornerstone in comprehensive cognitive assessments. In G. Goldstein, D.N. Allen, & J. DeLuca (Eds.), *Handbook of psychological assessment* (pp. 103–139). Berlin: Elsevier Academic Press. <https://doi.org/10.1016/B978-0-12-802203-0.00004-3>
- Inman, T.H. & Berry, D.T. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology*, 17(1), 1–23.
- Iverson, G.L. & Franzen, M.D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology*, 18(1), 38–51. <https://doi.org/10.1080/01688639608408260>
- Larrabee, G.J. (2014). False-positive rates associated with the use of multiple performance and symptom validity tests. *Archives of Clinical Neuropsychology*, 29(4), 364–373. <https://doi.org/10.1093/arclin/acu019>
- Lippa, S.M. (2018). Performance validity testing in neuropsychology: A clinical guide, critical review, and update on a rapidly evolving literature. *The Clinical Neuropsychologist*, 32(3), 391–421. <https://doi.org/10.1080/13854046.2017.1406146>
- Lovell, M.R. (2018). *ImPACT Administration and Interpretation Manual* [Measurement Instrument]. Retrieved from <https://impacttest.com/>
- Kaye, S., Sundman, M.H., Hall, E.E., Williams, E., Patel, K., & Ketcham, C.J. (2019). Baseline neurocognitive performance and symptoms in those with attention deficit hyperactivity disorders and history of concussion with previous loss of consciousness. *Frontiers in Neurology*, 10, 396. <https://doi.org/10.3389/fneur.2019.00396>
- Kerr, Z.Y., Snook, E.M., Lynall, R.C., Dompier, T.P., Sales, L., Parsons, J.T., & Hainline, B. (2015). Concussion-related protocols and preparticipation assessments used for incoming student-athletes in National Collegiate Athletic Association member institutions. *Journal of Athletic Training*, 50(11), 1174–1181.
- Manderino, L.M. & Gunstad, J. (2018a). Collegiate student athletes with history of ADHD or academic difficulties are more likely to produce an invalid protocol on baseline ImPACT testing. *Clinical Journal of Sports Medicine*, 28(2), 111–116. <https://doi.org/10.1097/JSM.0000000000000433>
- Manderino, L.M. & Gunstad, J. (2018b). Performance of the immediate post-concussion assessment and cognitive testing protocol validity indices. *Archives of Clinical Neuropsychology*, 33(5), 596–605. <https://doi.org/10.1093/arclin/acx102>
- Manderino, L.M., Zachman, A.M., & Gunstad, J. (2019). Novel ImPACT validity indices in collegiate student-athletes with and without histories of ADHD or academic difficulties. *The Clinical Neuropsychologist*, 33(8), 1455–1466. <https://doi.org/10.1080/13854046.2018.1539191>
- Mayes, S.D. & Calhoun, S.L. (2006). Frequency of reading, math, and writing disabilities in children with clinical disorders. *Learning and Individual Differences*, 16, 145–157.
- McClure, D.J., Zuckerman, S.L., Kutscher, S.J., Gregory, A.J., Solomon, G.S. (2013). Baseline neurocognitive testing in sports-related concussions: The importance of a prior night's sleep. *The American Journal of Sports Medicine*, 42(2), 472–478. <https://doi.org/10.1177/0363546513510389>
- Parke, E.M., Thaler, N.S., Etcoff, L.M., & Allen, D.N. (2015). Intellectual profiles in children with ADHD and comorbid learning and motor disorders. *Journal of Attention Disorders*, 1–10. <https://doi.org/10.1177/1087054715576343>
- Peltonen, K., Vartiainen, M., Laitala-Leinonen, T., Koskinen, S., Luoto, T., Pertab, J., & Hokkanen, L. (2019). Adolescent athletes with learning disability display atypical maturational trajectories on concussion baseline testing: Implications based on a Finnish sample. *Child Neuropsychology*, 25(3), 336–351. <https://doi.org/10.1080/09297049.2018.1474865>
- Poysofphon, P. & Rao, A.L. (2018). Neurocognitive deficits associated with ADHD in athletes: A systematic review. *Sports Health – A Multidisciplinary Approach*, 10(4), 317–326. <https://doi.org/10.1177/1941738117751387>
- Raab, C.A., Peak, A.S., & Knoderer, C. (2020). Half of purposeful baseline sandbaggers undetected by impact's embedded invalidity indicators. *Archives of Clinical Neuropsychology*, 35(3), 283–290. <https://doi.org/10.1093/arclin/acz001>
- Rabinowitz, A.R., Merritt, V.C., & Arnett, P.A. (2015). The return-to-play incentive and the effect of motivation on

- neuropsychological test-performance: Implications for baseline concussion testing. *Developmental Neuropsychology*, 40(1), 29–33. <https://doi.org/10.1080/87565641.2014.1001066>
- Rice, S.G. (2008). Medical conditions affecting sports participation. *Pediatrics*, 121(4), 841–848. <https://doi.org/10.1542/peds.2008-0080>
- Salinas, C.M., Dean, P., LoGalbo, A., Dougherty, M., Field, M., & Webbe, F.M. (2016). Attention-deficit hyperactivity disorder status and baseline neurocognitive performance in high school athletes. *Applied Neuropsychology: Child*, 5(4), 264–272. <https://doi.org/10.1080/21622965.2015.1052814>
- Schatz, P., Elbin, R.J., Anderson, M.N., Savage, J., & Covassin, T. (2017). Exploring sandbagging behaviors, effort, and perceived utility of the ImPACT baseline assessment in college athletes. *Sport, Exercise, and Performance Psychology*, 6(3), 243–251. <https://doi.org/10.1037/spy0000100>
- Schatz, P. & Glatts, C. (2013). “Sandbagging” baseline test performance on ImPACT, without detection, is more difficult than it appears. *Archives of Clinical Neuropsychology*, 28(3), 236–244. <https://doi.org/10.1093/arclin/act009s>
- Schatz, P., Kelley, T., Ott, S.D., Solomon, G.S., Elbin, R.J., Higgins, K., & Scolaro Moser, R. (2014). Utility of repeated assessment after invalid baseline neurocognitive test performance. *Journal of Athletic Training*, 49(5), 659–664. <https://doi.org/10.4085/1062-6050-49.3.37>
- Schatz, P., Scolaro Moser, R., Solomon, G.S., Ott, S.D., & Karpf, R. (2012). Prevalence of invalid computerized baseline neurocognitive test results in high school and collegiate athletes. *Journal of Athletic Training*, 47(3), 289–296. <https://doi.org/10.4085/1062-6050-47.3.14>
- Vickery, C.D., Berry, D.T., Dearth, C.S., Vagnini, V.L., Baser, R.E., Cragar, D.E., & Orey, S.A. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology*, 19(1), 37–48.
- Walton, S.R., Broshek, D.K., Freeman, J.R., Cullum, C.M., & Resch, J.E. (2018). Valid but invalid: Suboptimal ImPACT baseline performance in university athletes. *Medicine & Science in Sports & Exercise*, 50(7), 1377–1384. <https://doi.org/10.1249/MSS.0000000000001592>
- Zuckerman, S.L., Lee, Y.M., Odom, M.J., Solomon, G.S., Sills, A.K. (2013). Baseline neurocognitive scores in athletes with attention deficit-spectrum disorders and/or learning disability. *Journal of Neurosurgery Pediatrics*, 12(2), 103–109. <https://doi.org/10.3171/2013.5.PEDS12524>