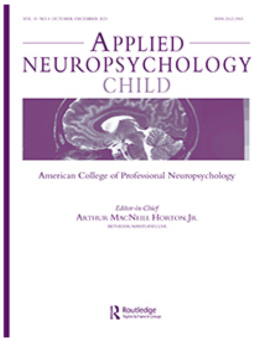

Applied Neuropsychology: Child

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hapc20

# Baseline preseason ImPACT® testing in Mandarin with adolescent student-athletes in the United States

Grant L. Iverson, Justin E. Karr, Yue Hong, Chi-Cheng Yang, Bruce Maxwell & Paul D. Berkner

**To cite this article:** Grant L. Iverson, Justin E. Karr, Yue Hong, Chi-Cheng Yang, Bruce Maxwell & Paul D. Berkner (2021): Baseline preseason ImPACT® testing in Mandarin with adolescent student-athletes in the United States, Applied Neuropsychology: Child, DOI: 10.1080/21622965.2021.1881897

**To link to this article:** https://doi.org/10.1080/21622965.2021.1881897

Published online: 14 Feb 2021.

APPLIED NEUROPSYCHOLOGY: CHILD
https://doi.org/10.1080/21622965.2021.1881897



# Baseline preseason ImPACT® testing in Mandarin with adolescent student-athletes in the United States

Grant L. Iverson[a,b], Justin E. Karr[c], Yue Hong[d,e], Chi-Cheng Yang[f,g], Bruce Maxwell[h], and Paul D. Berkner[i]

[a]Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital and Spaulding Research Institute, Harvard Medical School, Charlestown, Massachusetts, USA; [b]MassGeneral Hospital for Children Sports Concussion Program, Boston, Massachusetts, USA; [c]Department of Psychology, University of Kentucky, Lexington, Kentucky, USA; [d]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts, USA; [e]Home Base, A Red Sox Foundation and Massachusetts General Hospital Program, Charlestown, Massachusetts, USA; [f]Department of Psychology, National Chengchi University, Taipei, Taiwan; [g]Holistic Mental Health Center, Taipei City Hospital, Taipei, Taiwan [h]Department of Computer Science, Colby College, Waterville, Maine, USA; [i]Health Services, University of New England, Biddeford, Maine, USA

**ABSTRACT**

Researchers have examined differences on ImPACT® in baseline symptom reporting and neurocognitive performances based on the language of administration and racial/ethnic identity. This is the first study to examine differences between student-athletes tested in Mandarin versus English on ImPACT® during preseason baseline assessments conducted in high schools in the United States. Participants included 252 adolescent student-athletes who completed ImPACT® testing in the state of Maine in Mandarin and 252 participants who completed testing in English, matched on age, gender, and health and academic history. Participants were compared on neurocognitive composite scores and symptom ratings. Boys tested in Mandarin, but not girls, had modestly better neurocognitive performance on one of four composite scores (i.e., Visual Motor Speed, $p < .001$, $d = .45$). Although language groups did not differ in total symptom severity, boys tested in Mandarin endorsed multiple physical symptoms at higher rates than boys tested in English. These results suggest that the current ImPACT® neurocognitive normative data are reasonably appropriate for use with adolescents evaluated in Mandarin. There were some differences in the reporting of physical symptoms, with greater rates of symptom endorsement by boys tested in Mandarin than boys tested in English; but overall symptom severity ratings were comparable between the language groups.

**KEYWORDS**
Adolescent; brain concussion; cross-cultural comparison; neuropsychology

It is relatively common to undergo baseline preseason health care evaluations prior to participating in sports. A commonly used battery of tests for this purpose is included within the computerized program called the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT®) battery (ImPACT Applications Inc, 2011). ImPACT® contains a health history questionnaire, several neurocognitive tests, and a questionnaire that measures physical, cognitive, and psychological symptoms. Performance differences on the neurocognitive tests, at baseline, have been reported in association with age (Covassin et al., 2012), gender (Brown et al., 2015), socioeconomic status (Houck et al., 2020), motivation for testing (Bailey et al., 2006), and having a personal history of attention-deficit/hyperactivity disorder (ADHD) or learning disability (Cook et al., 2017; Peltonen et al., 2019). Greater baseline symptom reporting has been noted in association with a variety of participant characteristics (Iverson et al., 2015), such as female gender (Brown et al., 2015), having ADHD or a learning disability (Zuckerman et al., 2013), having a history of multiple prior concussions (Brooks et al., 2016, 2018), having a personal history of treatment for migraines or mental health problems (Iverson et al., 2015), or getting insufficient sleep prior to baseline testing (Silverberg et al., 2016).

Prior studies have also documented the impact of cultural and linguistic characteristics on baseline preseason neurocognitive performance and symptom reporting using ImPACT®, which is important because ImPACT® is administered to student-athletes of diverse backgrounds and has been translated into 21 different languages. Specifically, African American athletes have performed worse than White athletes at baseline on Visual Motor Speed and Reaction Time (Wallace et al., 2018), and Hispanic Spanish-speaking athletes have scored lower than English-speaking athletes at baseline on all composites when tested in Spanish and on three composites when tested in English (i.e., Verbal Memory, Visual Memory, and Visual Motor Speed) (Ott et al., 2014). Another study on professional baseball players found that players whose first language was Spanish performed worse on all neurocognitive composites at baseline compared to players whose first language was English (Jones

**CONTACT** Grant L. Iverson giverson@mgh.harvard.edu Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital and Spaulding Research Institute, Harvard Medical School, 79/96 Thirteenth Street, Charlestown Navy Yard, Charlestown, MA 02129, USA.

et al., 2014). In contrast, other studies have found equivalence in baseline neurocognitive performances between native and non-native English speakers (Tsushima et al., 2017), White and African American student-athletes in the United States (Kontos et al., 2010), and American football and South African rugby players (Shuttleworth-Edwards et al., 2009). The results of these studies are mixed, but they evidence some differences in performances related to racial/ethnic identity and language of administration. Impacts of cultural and/or linguistic characteristics on baseline evaluation have not been explored in student-athletes tested in Mandarin. This is a critical gap given that greater than 4 million individuals residing in the United States identify as Chinese or Chinese American, and 2.9 million people reported speaking Chinese languages at home (Ryan, 2013).

ImPACT® has the potential to meet the demand of a multilingual population because the test is completely computerized and has been translated into many languages, including Mandarin. However, translation alone does not render a test appropriate for its intended population. A test requires not only translational equivalence, but also conceptual equivalence and metric equivalence before implementation (Sue & Sue, 2000). Translational equivalence refers to how well descriptors and ability measures survive translation and back translation by independent, bilingual translators. Conceptual equivalence refers to whether the same underlying constructs are being measured, but not necessarily the specific behaviors or thoughts that are being assessed to measure the construct. For example, an aspect of irritability in one culture may be a tendency to avoid family members, whereas in another culture it may be to seek out family members and verbally express irritation; thus, the underlying construct, irritability, would be measured by quantifying different behaviors in different cultures. Metric equivalence refers to whether a score of 30 on a certain test is equivalent to a score of 30 on the translated version of the same test. To our knowledge, there has never been a study that examined the equivalence of the Mandarin version of the ImPACT® to the widely used English version.

In this study, we assessed the equivalence of the Mandarin-language and English-language ImPACT® batteries, by examining (a) the translational and conceptual equivalence of the test by performing a translation and back translation procedure, and (b) the metric equivalence of the test by examining whether there are cognitive performance or symptom reporting differences on ImPACT® testing between matched adolescents tested in English or Mandarin on baseline preseason evaluations prior to their participation in high school sports. In terms of symptom reporting, Asian populations may have a tendency to endorse greater somatic symptoms, which could reflect underlying psychological distress (Bagayogo et al., 2013; Grover & Ghosh, 2014), and we hypothesized that there would be modestly higher symptom severity reported by athletes tested in Mandarin, especially in regard to physical symptoms. Although we are unaware of published comparisons between computerized Mandarin- and English-language tests, previous research on computerized cognitive assessment has identified no consistent or meaningful group differences between White and Asian examinees (Casaletto et al., 2015); and we hypothesized that there would be no meaningful differences between adolescents tested in Mandarin versus English on baseline ImPACT® neurocognitive testing.

## Methods

### *Participants*

The participants for this study were derived from a database of 47,606 student-athletes aged 13–18 years old from across the state of Maine, USA who completed ImPACT® as a part of their baseline, preparticipation evaluations between 2009 and 2015. In this sample, 263 participants (0.55%) completed test administration in Mandarin. Of these individuals, nine were removed for invalid test performances based on embedded performance validity criteria on ImPACT® testing, one was removed for reporting a history of treatment for seizures, and one was removed for reporting a history of treatment for meningitis. This resulted in a sample of 252 participants. For comparison, matched participants were drawn from a sample of student-athletes who completed the test in English with valid test performances and no personal self-reported history of treatment for seizures, brain surgery, or meningitis ($N = 45{,}067$). Matching was conducted using the case-control matching procedure in SPSS v.25, which took participants tested in Mandarin and identified participants tested in English that matched participants exactly on pre-specified criteria, which included age, gender, concussion history, self-reported diagnosis of attention-deficit disorder (ADD) or ADHD, diagnosis of learning disability or dyslexia, history of special education, and previous treatment for a psychiatric condition (e.g., depression, anxiety), substance use disorder, headaches, or migraines. This case-control matching procedure found exact matches tested in English for all participants tested in Mandarin. The final sample included 252 participants tested in Mandarin and 252 participants tested in English. The average age of both the student-athlete groups was 16.07 years ($SD = 1.18$). The gender distribution for both groups was 62.7% boys and 37.3% girls. Data regarding race and ethnicity were not available. Information pertaining to language proficiency or bilingualism was not available (e.g., the level of English proficiency of student-athletes tested in Mandarin is unknown, and it is not known whether some bilingual students chose to do baseline testing in English).

### *Procedure*

All health history information in this study is retrospective and based on adolescent self-report. The health history survey includes questions relating to whether the student has "been diagnosed with ADD/ADHD" or "had problems with ADD/hyperactivity," been "diagnosed with Dyslexia" or been "diagnosed with a learning disability," attended special education classes, or repeated one or more years of school. These questions require a yes or no response. The health history survey also asks about the number of times the student has been diagnosed with a concussion as well as whether the student has had past treatment for headaches, migraines, epilepsy/seizures, brain surgery, meningitis, substance/alcohol use, or a psychiatric condition (e.g., anxiety or depression). Student-athletes in our sample completed baseline testing with ImPACT® prior to participating in sports for a given season. ImPACT® was administered in a group setting in high school computer laboratories by the local athletic training staff. There was no information on group size available in the database. These data were collected and merged into a statewide dataset. Institutional Review Board approval for use of this de-identified database was obtained from Colby College and an exemption was obtained from Spaulding Rehabilitation Hospital for secondary use.

### Measures

ImPACT® is a brief, computer-administered neuropsychological test battery consisting of six individual test modules that contribute to four neurocognitive composite scores (i.e., Verbal Memory, Visual Memory, Visual Motor Speed, and Reaction Time) and an Impulse Control score used as a validity indicator. The ImPACT® neurocognitive composite scores have adequate test-retest reliability across many, but not all, studies (Brett et al., 2016; Elbin et al., 2011; Iverson & Schatz, 2015; Mason et al., 2020; Nakayama et al., 2014; Schatz, 2010; Schatz & Ferris, 2013; Womble et al., 2016). Baseline testing in schools is usually done in group settings, and there is some evidence that scores are lower when tested in groups compared to when tested individually (Moser et al., 2011). In research and clinical practice, it is customary to interpret the two memory composite scores and the two speed composite scores. This battery has been widely used, for many years, as part of preseason baseline testing protocols in high schools and colleges (Houck et al., 2019; Katz et al., 2018). The battery is used following an injury to measure cognitive functioning and symptom reporting (McClincy et al., 2006; Schatz et al., 2006). Normative data for the neurocognitive test scores are adjusted for age and gender, and the results are presented as percentile ranks for clinical use. Factor analytic studies have revealed a two-factor solution, memory and speed (Brett et al., 2018; Gerrard et al., 2017), but those factors are not included by the company on the printout or in the downloadable database. ImPACT® also contains a Post-Concussion Symptom Scale (PCSS) consisting of 22 commonly reported symptoms following concussion (e.g., headache, dizziness, fatigue, etc.) that are rated from 0 (none) to 6 (severe). The total symptom severity score is derived by summing the 22 items (range of possible scores: 0–132).

### Translational and conceptual equivalence

To understand the translational equivalence of the Mandarin-language ImPACT® compared to the original English version, two of the authors, as well as three researchers blinded to this study, who are all bilingual in English and Mandarin conducted double-translation and reconciliation of the PCSS used within ImPACT®. All five translators are native Mandarin speakers and native to Chinese culture, which is recommended as one method to ensure conceptual equivalence in translation (International Test Commission, 2017). Specifically, two of the authors forward translated the English-language PCSS independently and compared their translations to the Mandarin-language PCSS currently used in ImPACT®. Discrepancies between the two authors and the current Mandarin ImPACT® were identified and forwarded to a panel of three blinded bilingual researchers for backward translation. Backward translation results were then reconciled into a single recommended translation of the symptom. This exercise was done to understand the equivalence between the English-language PCSS items and the Mandarin-language PCSS items. The translations were not incorporated into the PCSS for use by the student-athletes. Participants completed the PCSS that was included in the ImPACT® software.

### Statistical analyses

Student-athletes tested in Mandarin versus English were compared using a 2 (gender) × 2 (language group) multivariate analysis of variance (MANOVA) for the four neurocognitive composite raw scores (i.e., Verbal Memory, Visual Memory, Visual Motor Speed, and Reaction Time). An analogous 2 × 2 univariate analysis of variance (ANOVA) was conducted for the Impulse Control score. For these analyses, $p < .05$ was set as the threshold for statistical significance. Follow-up univariate independent-samples *t*-tests compared the language groups on the four ImPACT® neurocognitive composite raw scores and the Impulse Control score, with stratifications by gender. Due to a non-normal distribution, Mann–Whitney *U* tests were conducted for the PCSS total symptom score to compare language groups, with stratifications by gender due to known differences in baseline symptom reporting between boys and girls (Brown et al., 2015). For these univariate analyses, a Bonferroni correction was used to control for the familywise type I error rate, using an adjustment of $p < .05/18$ based on 18 univariate analyses (i.e., six ImPACT® scores compared across language groups for the total sample, and boys and girls separately).

Sample means and standard deviations were used to calculate Cohen's *d* for parametric comparisons, interpreted according to conventional guidelines: $d = .20$ is interpreted as small, $d = .50$ is interpreted as medium, and $d = .80$ is interpreted as large (Cohen, 1988). The *z*-statistic from the Mann–Whitney *U* tests was used to calculate the non-parametric effect size *r*, calculated as $r = \frac{Z}{\sqrt{N}}$ (Fritz et al., 2012). The *r* effect size was interpreted according to conventional guidelines: $r = .10$ is interpreted as small, $r = .30$ is interpreted as medium, and $r = .50$ is interpreted as large (Cohen, 1988). Odds ratios (ORs) were computed to examine group differences in individual symptom endorsement rates. The ORs can be interpreted as effect sizes based on a

Table 1. Summary of health and academic history and sport affiliation.

| | Tested in Mandarin | | | | | | Tested in English | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total sample | | Boys | | Girls | | Total sample | | Boys | | Girls | |
| | (*n* = 252) | | (*n* = 158) | | (*n* = 94) | | (*n* = 252) | | (*n* = 158) | | (*n* = 94) | |
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| Health and academic history | | | | | | | | | | | | |
| Received speech therapy[a] | 7 | 2.8 | 7 | 4.4 | 0 | 0.0 | 10 | 4.0 | 8 | 5.1 | 2 | 2.1 |
| Learning problems[a,b] | 17 | 6.8 | 10 | 6.3 | 7 | 7.5 | 22 | 8.7 | 14 | 8.9 | 8 | 8.5 |
| ADD/ADHD | 4 | 1.6 | 4 | 2.5 | 0 | 0.0 | 4 | 1.6 | 4 | 2.5 | 0 | 0.0 |
| Treatment for headaches | 15 | 6.0 | 8 | 5.1 | 7 | 7.4 | 15 | 6.0 | 8 | 5.1 | 7 | 7.4 |
| Treatment for migraines | 11 | 4.4 | 10 | 6.3 | 1 | 1.1 | 11 | 4.4 | 10 | 6.3 | 1 | 1.1 |
| Treatment for substance/alcohol use | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 |
| Treatment for a psychiatric condition | 4 | 1.6 | 3 | 1.9 | 1 | 1.1 | 4 | 1.6 | 3 | 1.9 | 1 | 1.1 |
| ≥Prior concussion | 11 | 4.4 | 8 | 5.1 | 3 | 3.2 | 11 | 4.4 | 8 | 5.1 | 3 | 3.2 |
| Sport affiliation | | | | | | | | | | | | |
| Football (any level) | 42 | 16.7 | 26 | 16.5 | 16 | 17.1 | 36 | 14.3 | 36 | 22.8 | 0 | 0.0 |
| Basketball | 35 | 13.9 | 27 | 17.1 | 8 | 8.5 | 31 | 12.3 | 23 | 14.6 | 8 | 8.5 |
| Rugby | 28 | 11.1 | 28 | 17.7 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 |
| Soccer | 28 | 11.1 | 25 | 15.8 | 3 | 3.2 | 47 | 18.7 | 34 | 21.5 | 13 | 13.8 |
| Archery | 27 | 10.7 | 10 | 6.3 | 17 | 18.1 | 8 | 3.2 | 5 | 3.2 | 3 | 3.2 |
| Swimming | 20 | 7.9 | 7 | 4.4 | 13 | 13.8 | 5 | 2.0 | 1 | 0.6 | 4 | 4.3 |
| Tennis | 14 | 5.6 | 10 | 6.3 | 4 | 4.3 | 9 | 3.6 | 4 | 2.5 | 5 | 5.3 |
| Track and field | 12 | 4.8 | 8 | 5.1 | 4 | 4.3 | 13 | 5.2 | 9 | 5.7 | 4 | 4.3 |
| Cheerleading | 6 | 2.4 | 0 | 0.0 | 6 | 6.4 | 13 | 5.2 | 0 | 0.0 | 13 | 13.8 |
| Other[c] | 40 | 15.9 | 17 | 10.6 | 23 | 24.5 | 90 | 35.7 | 46 | 29.1 | 44 | 46.9 |

*Note.* ADD/ADHD: Attention deficit disorder or Attention-deficit/hyperactivity disorder. The data were gathered as part of the Maine Concussion Management Initiative (MCMI) under the direction of the principal investigator Dr. Paul Berkner. The authors thank the Maine Athletic Trainers Association for their collaboration with the MCMI.

[a]These numbers do not match exactly across Mandarin-speaking and English-speaking groups because participants were not matched based on special education history or repeating an academic grade.

[b]Learning difficulties include formal diagnoses of learning disability/dyslexia, history of attending special education classes, or history of repeating one of more years of school.

[c]Other sports include lacrosse, field hockey, ice hockey, baseball, softball, volleyball, skiing, snowboarding, cross country, skateboarding, golf, equestrian sports, wrestling, martial arts, and biking.

conversion method reported by Chinn (2000) and interpreted according to widely used criteria (Cohen, 1988): an OR between 1.20 and 1.71 is interpreted as small, between 1.72 and 2.40 is interpreted as medium, and greater than 2.40 is interpreted as large in magnitude. All ORs are oriented so that values above 1.00 are indicative of greater odds of endorsing the symptom by the language group with a higher endorsement rate. Data were analyzed using SPSS v.25.

## Results

### *Characteristics of student-athletes*

Information regarding academic and health history is provided in Table 1, including history of speech therapy, learning problems (i.e., learning disability or dyslexia, special education, repeated grades in school), ADHD diagnosis, prior concussion history, prior treatment for different health conditions (i.e., headaches, migraines, substance or alcohol use, or psychiatric conditions), and sport affiliation at the time of baseline testing. The student-athletes tested in Mandarin participated in many different sports at the time of baseline assessment, such as American football (16.7%), basketball (13.9%), rugby (11.1%), and soccer (11.1%). These percentages may be biased or even invalid, however, by potential misunderstandings through English-to-Mandarin translation. For instance, football was translated into "American soccer" and rugby was translated into "Irish soccer," both of which could be mistakenly selected for "soccer." The most common sport affiliations reported for student-athletes tested in English included soccer (18.7%), American football (14.3%), and basketball (12.3%).

### *Translational and conceptual equivalence*

For the Mandarin versions of the neurocognitive tests, there were no differences in the translated instructions or stimuli that the authors believed may interfere with participants' understanding of the test requirements or their capacity to accurately complete the tests. Regarding symptom reporting, a double-translation procedure was performed to determine the translational and conceptual equivalence between the English- and Mandarin-language versions of the PCSS items. Overall, 19 of 22 items were forward translated to create results that are equivalent to the currently used Mandarin-language ImPACT® PCSS. Three items on the Mandarin-language PCSS, however, were identified to represent a more severe symptom than the English item from which they were translated, which may in turn impact the conceptual equivalence between the Mandarin and English versions of the PCSS. Specifically, the item 眩晕 was back translated to *vertigo* rather than *dizziness*, 嗜睡 was back translated to *hypersomnia* rather than *drowsiness*, and 神志模糊 was back translated to *dazed and confused* rather than *feeling mentally "foggy"*. Alternative translations from the forward translation procedures may provide better translational equivalence. Specifically, alternative translations for these three items could be as follows: dizziness = 头晕, which back translated

**Table 2.** Summary of group differences on ImPACT® neurocognitive composite scores and total symptom scores.

| | Tested in Mandarin ($N = 252$; Girls: $n = 94$; Boys: $n = 158$) Age: $M = 16.07 \pm 1.18$ years, range: 13–18 | | | | Tested in English ($N = 252$; Girls: $n = 94$; Boys: $n = 158$) Age: $M = 16.07 \pm 1.18$ years, range: 13–18 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ImPACT® component | Median | Mean | SD | Range | Median | Mean | SD | Range | *t/U* | *p* | *d/r* |
| Total sample | | | | | | | | | | | |
| Verbal memory | 87.00 | 84.73 | 10.12 | 56.00–100 | 85.00 | 84.00 | 9.53 | 54.00–100 | $t = -.83$ | .410 | $d = .07$ |
| Visual memory | 77.00 | 75.69 | 12.77 | 41.00–100 | 74.00 | 72.88 | 12.90 | 38.00–99.00 | $t = -2.46$ | .014 | $d = .22$ |
| Visual motor | 38.40 | 38.18 | 8.09 | 11.30–53.63 | 35.37 | 35.41 | 6.82 | 19.30–51.20 | $t = -4.15$ | <.001 | $d = .37$ |
| Reaction time° | 0.61 | 0.63 | 0.10 | 0.45–1.31 | 0.60 | 0.61 | 0.09 | 0.45–1.02 | $t = -1.89$ | .060 | $d = .17$ |
| Impulse control° | 5 | 6.16 | 4.31 | 0–25 | 5 | 6.37 | 5.01 | 0–30 | $t = -.51$ | .614 | $d = .05$ |
| Total symptom score° | 2 | 6.12 | 8.76 | 0–54 | 2 | 4.97 | 7.60 | 0–60 | $U = 29{,}598.5$ | .177 | $r = .06$ |
| Girls only | | | | | | | | | | | |
| Verbal memory | 89.00 | 87.60 | 9.19 | 56.00–100 | 86.00 | 84.30 | 9.58 | 54.00–100 | $t = -2.41$ | .017 | $d = .35$ |
| Visual memory | 78.00 | 76.46 | 13.05 | 41.00–100 | 73.00 | 72.04 | 11.92 | 41.00–96.00 | $t = -2.42$ | .017 | $d = .35$ |
| Visual motor | 37.32 | 37.32 | 7.56 | 13.50–53.25 | 35.94 | 35.66 | 6.71 | 20.90–50.38 | $t = -1.59$ | .114 | $d = .23$ |
| Reaction time° | 0.61 | 0.62 | 0.09 | 0.45–0.87 | 0.60 | 0.61 | 0.08 | 0.45–0.89 | $t = -1.22$ | .225 | $d = .18$ |
| Impulse control° | 5 | 5.83 | 4.17 | 1–25 | 5 | 6.31 | 5.32 | 0–30 | $t = .69$ | .493 | $d = .10$ |
| Total symptom score° | 3 | 6.59 | 8.95 | 0–54 | 3 | 5.54 | 6.89 | 0–28 | $U = 4{,}287.5$ | .723 | $r = .02$ |
| Boys only | | | | | | | | | | | |
| Verbal memory | 84.00 | 83.02 | 10.29 | 57.00–100 | 85.00 | 83.83 | 9.52 | 56.00–100 | $t = .73$ | .468 | $d = .08$ |
| Visual memory | 75.50 | 75.23 | 12.62 | 42.00–98.00 | 75.00 | 73.38 | 13.46 | 38.00–99.00 | $t = -1.26$ | .207 | $d = .14$ |
| Visual motor | 39.13 | 38.69 | 8.36 | 11.30–53.63 | 35.17 | 35.26 | 6.89 | 19.30–51.20 | $t = -3.97$ | <.001 | $d = .45$ |
| Reaction time° | 0.61 | 0.63 | 0.11 | 0.45–1.31 | 0.61 | 0.61 | 0.09 | 0.45–1.02 | $t = -1.46$ | .146 | $d = .16$ |
| Impulse control° | 5 | 6.36 | 4.39 | 0–22 | 5 | 6.41 | 4.84 | 0–25 | $t = .10$ | .922 | $d = .01$ |
| Total symptom score° | 2 | 5.85 | 8.66 | 0–43 | 1 | 4.63 | 7.99 | 0–60 | $U = 11{,}391.5$ | .165 | $r = .06$ |

*Note*. ImPACT®: Immediate Post-Concussion Assessment and Cognitive Testing; SD: standard deviation; independent-samples *t*-tests were used for group comparisons on the four ImPACT® neurocognitive test composite raw scores and the Impulse Control score, with Cohen's *d* calculated as the effect size. Mann–Whitney *U* tests were used for group comparisons on the Total Symptom Score due to non-normal distributions, with a non-parametric effect size (*r*) calculated as $r = \frac{Z}{\sqrt{N}}$. For comparison, a previously published sample of 17,290 boys and 14,668 girls from the state of Maine, using an earlier version of the database used in the present study, revealed the following for the Total Symptom Score: Boys M = 4.5 (7.9), Md = 1, IQR = 0–9 and Girls M = 6.5 (9.9) Md = 3, IQR = 0–9 (Iverson et al., 2015).

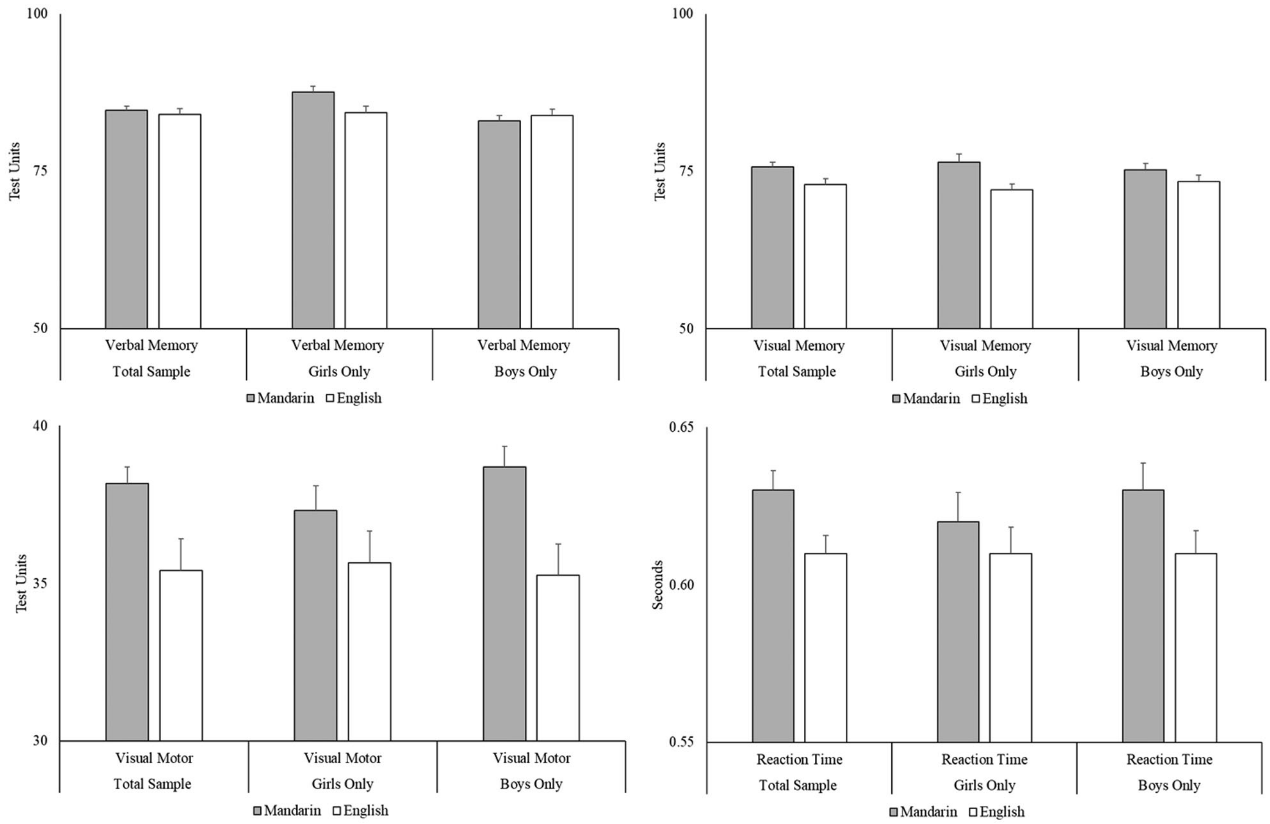


**Figure 1.** Average ImPACT® neurocognitive composite scores, with standard error bars, by language of administration and gender.

to dizziness by all three independent researchers; drowsiness = 昏沉, which back translated to drowsiness by two independent researchers and to lethargy by one; and feeling mentally "foggy" = 感觉脑子"一团浆糊", which back translated to feeling fuzzy headed, mentally blurred, and head in a fog by three independent researchers, respectively.

## Comparisons on neurocognitive composites and impulse control

A 2 × 2 MANOVA with the four neurocognitive composite scores as dependent variables indicated significant differences in performances between student-athletes tested in Mandarin versus English, $F(4,497) = 8.83$, $p < .001$, $\Lambda = .93$, ${\eta_p}^2 = .07$, significant differences between genders, $F(4,497) = 3.28$, $p = .011$, $\Lambda = .97$, ${\eta_p}^2 = .03$, and an interaction between language group and gender, $F(4,497) = 2.54$, $p = .039$, $\Lambda = .98$, ${\eta_p}^2 = .02$. A 2 × 2 ANOVA for the Impulse Control composite score did not result in significant differences based on language, $F(1,500) = 0.38$, $p = .539$, ${\eta_p}^2 < .01$, or gender, $F(1,500) = 0.54$, $p = .462$, ${\eta_p}^2 < .01$, and found no interaction between these variables either, $F(1,500) = 0.25$, $p = .620$, ${\eta_p}^2 < .01$.

The language groups differed significantly on one of the four ImPACT® neurocognitive composites, in that the group tested in Mandarin obtained higher scores on the Visual Motor Speed composite (see Table 2), with a small-to-medium effect size ($p < .001$, $d = .37$). For Visual Motor Speed, a higher score is indicative of a better performance. When divided by gender, there were no statistically significant group difference on any of the neurocognitive composites for girls, but boys tested in Mandarin scored statistically higher on the Visual Motor Speed than boys tested in English, with a medium effect size ($p < .001$, $d = .45$). The average performances for the participants tested in Mandarin and English, with gender stratifications, are displayed visually in Figure 1. The percentiles for the student-athletes tested in Mandarin, stratified by gender, are provided in Table 3.

## Symptom reporting

The language groups did not differ significantly on the PCSS total score, either for the total sample or when separated by gender. We examined gender differences within each language group as a post hoc analysis. There was no statistically significant difference in the total symptom score between boys and girls tested in Mandarin, $U = 6782.0$, $p = .24$, $r = .07$. In contrast, there was a marginally significant difference in the total symptom score between boys and girls tested in English, $U = 6{,}270.0$, $p = .03$, such that girls endorsed a greater severity of symptoms, albeit with a small effect size ($r = .13$).

We examined individual symptom endorsement rates among the student-athletes, stratified by language of administration and gender (see Table 4). Girls tested in Mandarin endorsed 2 of the 22 individual symptoms at a greater rate than girls tested in English, with large effect sizes: *sensitivity to noise* (OR = 5.31) and *visual problems* (OR = 3.65). Girls tested in English endorsed the symptom of *feeling more*

**Table 3.** Percentiles for ImPACT® neurocognitive composite scores for student-athletes tested in Mandarin.

| | Total ($n = 252$) | | | | Girls ($n = 94$) | | | | Boys ($n = 158$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Percentile | Verbal memory | Visual memory | VMS | RT | Verbal memory | Visual memory | VMS | RT | Verbal memory | Visual memory | VMS | RT |
| 2nd | 61.0 | 45.0 | 14.5 | 0.48 | 60.0 | 47.0 | 14.5 | 0.46 | 61.0 | 45.0 | 14.1 | 0.48 |
| 5th | 65.0 | 51.0 | 25.1 | 0.50 | 73.0 | 53.0 | 23.7 | 0.52 | 65.0 | 50.0 | 25.1 | 0.50 |
| 9th | 70.0 | 57.0 | 27.6 | 0.53 | 74.0 | 57.0 | 26.1 | 0.53 | 68.0 | 58.0 | 28.2 | 0.52 |
| 16th | 74.0 | 64.0 | 31.4 | 0.54 | 79.0 | 64.0 | 31.7 | 0.54 | 71.0 | 64.0 | 31.2 | 0.54 |
| 25th | 78.0 | 67.5 | 33.1 | 0.56 | 84.0 | 67.0 | 33.0 | 0.56 | 76.0 | 68.0 | 33.5 | 0.56 |
| 75th | 92.0 | 85.0 | 44.3 | 0.67 | 93.0 | 86.0 | 42.5 | 0.68 | 91.0 | 84.0 | 45.1 | 0.67 |
| 84th | 95.0 | 90.0 | 46.4 | 0.72 | 96.0 | 90.0 | 44.9 | 0.71 | 93.0 | 90.0 | 48.0 | 0.72 |
| 91st | 97.0 | 92.0 | 48.6 | 0.76 | 99.0 | 93.0 | 46.3 | 0.75 | 96.0 | 92.0 | 49.0 | 0.76 |
| 95th | 99.0 | 94.0 | 49.4 | 0.81 | 100.0 | 95.0 | 48.6 | 0.80 | 99.0 | 94.0 | 49.7 | 0.82 |
| 98th | 100.0 | 97.0 | 50.8 | 0.88 | 100.0 | 100.0 | 52.7 | 0.83 | 100.0 | 97.0 | 50.8 | 0.91 |

*Note.* VMS: Visual Motor Speed composite; RT: reaction time.

*emotional* at a greater rate than girls tested in Mandarin, with a medium effect size (OR = 2.38). Boys tested in Mandarin endorsed 7 of the 22 symptoms at a greater rate. Five of these group differences were associated with large effect sizes, including *nausea* (OR = 4.64), *dizziness* (OR = 2.51), *fatigue* (OR = 2.54), *sleeping less than usual* (OR = 2.48), and *visual problems* (OR = 3.33); and two of these group differences were associated with medium effect sizes, including *nervousness* (OR = 2.39) and *difficulty concentrating* (OR = 2.34). Boys tested in English endorsed *trouble falling asleep* at a greater rate than the boys tested in Mandarin, with a medium effect size (OR = 1.84). For the symptom *balance problems*, the OR could not be calculated for boys because 0% of boys tested in Mandarin endorsed this symptom.

## Discussion

To our knowledge, this is the first empirical study to examine whether there are differences on baseline preseason neurocognitive performances or symptom reporting on ImPACT® between adolescents residing in the United States who completed ImPACT® in Mandarin versus English. Overall, our results showed no significant differences on three of the four neurocognitive composites and the Impulse Control score at baseline between the two groups. This aligns, in part, with our initial hypothesis; however, boys tested in Mandarin scored significantly higher than boys tested in English on the domain of Visual Motor Speed, with a small-to-medium effect size ($d = .37$). Although nonsignificant after controlling for multiple comparisons, there was a small group difference on the Visual Memory composite as well ($d = .22$), with participants tested in Mandarin again performing slightly better than participants tested in English. Cross-cultural group differences have been observed previously on ImPACT®, including differences based on language, and race and ethnicity, although these latter variables were not available for analysis in the current study. African American athletes have obtained lower scores compared to White athletes at baseline on the neurocognitive composite scores (Wallace et al., 2018); Hispanic Spanish-speaking athletes have scored lower than English-speaking athletes at baseline on all composites when tested in Spanish and on three composites when tested in English (i.e., Verbal Memory, Visual Memory, and Visual Motor Speed) (Ott et al., 2014); and Finnish athletes have scored lower than English-speaking athletes at baseline on Visual Motor Speed and Reaction Time (Vartiainen et al., 2021). In contrast, another study found equivalence in baseline neurocognitive performances between White and African American student-athletes (Kontos et al., 2010). The current findings are unique in their focus on student-athletes tested in Mandarin, and the findings fit somewhere in between previous cross-cultural studies, with some modest differences in neurocognitive performances based on language of administration.

Regarding total symptom reporting, girls tested in English endorsed a greater severity of symptoms than boys tested in English, which is consistent with prior studies (Brown et al., 2015). In contrast, there was no significant gender difference in total symptom reporting for the participants tested in Mandarin. This is not due to an overall underreporting of symptoms by adolescents tested in Mandarin, because total symptom scores were not statistically different between the language groups. The absence of a gender-related difference in the sample tested in Mandarin seemed to be driven by slightly greater symptom reporting by boys evaluated in Mandarin than boys evaluated in English, without a difference in symptom reporting by girls. As with previous research on neurocognitive findings, results have been mixed regarding cross-cultural differences in symptom reporting. Spanish-speaking student-athletes have reported a modestly greater baseline symptom severity on the Spanish-language PCSS than Spanish-speaking or English-speaking student-athletes on the English-language PCSS (Ott et al., 2014). African American student-athletes have reported modestly greater symptom severity than White athletes at baseline in one study (Wallace et al., 2018), but no difference was observed in another study (Kontos et al., 2010).

In addition to total symptom severity, group differences were also examined in regard to frequencies at which participants endorsed individual symptoms. Girls tested in Mandarin were more likely to endorse two sensory symptoms, *sensitivity to noise* and *visual problems,* than their peers tested in English, and they were less likely to endorse an affective symptom of *feeling more emotional*. Boys tested in Mandarin, compared to their peers tested in English, were more likely to endorse various somatic symptoms, including *nausea, dizziness, fatigue, sleeping less than usual,*

**Table 4.** Symptom reporting of student-athletes tested in Mandarin and English stratified by gender.

| Symptom | Girls | | | Boys | | |
|---|---|---|---|---|---|---|
| | Mandarin ($n = 94$) (%) | English ($n = 94$) (%) | OR (95% CI) | Mandarin ($n = 158$) (%) | English ($n = 158$) (%) | OR (95% CI) |
| Headache | 22.3 | 20.2 | 1.14 (.56–2.29) | 17.7 | 19.6 | 1.13 (.64–2.00) |
| Vomiting | 5.3 | 4.3 | 1.26 (.33–4.86) | 5.7 | 4.4 | 1.30 (.47–3.59) |
| Nausea | 9.6 | 4.3 | 2.38 (.71–8.03) | **10.8** | **2.5** | 4.64 (1.53–14.13) |
| Balance problems | 2.1 | 3.2 | 1.52 (.25–9.29) | 0.0 | 3.8 | – |
| Dizziness | 16.0 | 7.4 | 2.36 (.92–6.09) | **15.8** | **7.0** | 2.51 (1.19–5.30) |
| Trouble falling asleep | 21.3 | 22.3 | 1.07 (.53–2.14) | **16.5** | **26.6** | 1.84 (1.06–3.18) |
| Fatigue | 38.3 | 29.8 | 1.46 (.80–2.68) | **39.2** | **20.3** | 2.54 (1.54–4.20) |
| Sleeping more than usual | 8.5 | 5.3 | 1.66(.52–5.26) | 7.6 | 6.3 | 1.22 (.51–2.90) |
| Sleeping less than usual | 43.6 | 33.0 | 1.57 (.87–2.84) | **36.7** | **19.0** | 2.48 (1.48–4.13) |
| Drowsiness | 24.5 | 13.8 | 2.02 (.95–4.28) | 18.4 | 12.0 | 1.65 (.88–3.08) |
| Sensitivity to light | 17.0 | 16.0 | 1.08 (.50–2.34) | 7.0 | 13.3 | 2.05 (.95–4.41) |
| Sensitivity to noise | **14.9** | **3.2** | 5.31 (1.47–19.14) | 7.0 | 5.1 | 1.40 (.55–3.59) |
| Irritability | 16.0 | 13.8 | 1.18 (.53–2.65) | 14.6 | 9.5 | 1.62 (.81–3.24) |
| Sadness | 20.2 | 20.2 | 1.00 (.49–2.04) | 16.5 | 10.1 | 1.75 (.90–3.40) |
| Nervousness | 27.7 | 22.3 | 1.33 (.69–2.58) | **27.8** | **13.9** | 2.39 (1.35–4.22) |
| Feeling more emotional | **13.8** | **27.7** | 2.38 (1.14–4.99) | 13.3 | 8.2 | 1.71 (.82–3.55) |
| Numbness or tingling | 5.3 | 2.1 | 2.58 (.49–13.67) | 5.7 | 3.8 | 1.53 (.53–4.41) |
| Feeling slowed down | 4.3 | 5.3 | 1.21 (.36–4.12) | 6.3 | 5.7 | 1.12 (.44–2.83) |
| Feeling mentally "foggy" | 7.4 | 6.4 | 1.18 (.38–3.65) | 3.8 | 5.1 | 1.35 (.46–3.99) |
| Difficulty concentrating | 26.6 | 19.1 | 1.53 (.77–3.04) | **25.3** | **12.7** | 2.34 (1.30–4.22) |
| Difficulty remembering | 9.6 | 13.8 | 1.52 (.62–3.74) | 7.6 | 6.3 | 1.22 (.51–2.90) |
| Visual problems | **17.0** | **5.3** | 3.65 (1.28–10.43) | **18.4** | **6.3** | 3.33 (1.56–7.09) |

*Note*. CI: Confident Interval; OR: Odds ratio. Bold values correspond to ORs with CIs that do not overlap 1.00. All ORs are oriented so that values above 1.00 are indicative of greater odds of endorsing the symptom by the language group with a higher endorsement rate. For the symptom *Balance Problems*, an OR could not be calculated for boys because 0% of Mandarin speakers endorsed this symptom.

and *visual* problems, an affective symptom (i.e., *nervousness*), and a cognitive symptom (i.e., *difficulty concentrating*). Boys tested in English were more likely to endorse *trouble falling asleep*. These group differences indicate that student-athletes tested in Mandarin, particularly boys, are more likely to endorse primarily physical symptoms than their peers tested in English during preseason baseline evaluations, which aligns in part with our initial hypothesis.

Of the 22 PCSS items, only three symptoms were not consistent in meaning between the English-language and Mandarin-language versions (i.e., *dizziness*, *drowsiness*, and *feeling mentally "foggy"*). Adolescents differed in frequency of endorsement on only one of these items based on language of administration: the item *dizziness*, which was translated to *vertigo* in Mandarin, was endorsed at a significantly higher rate by boys tested in Mandarin than its equivalent symptom in English. It is unclear whether or how this comparison would change if the alternative translation were provided.

Aside from considerations of test equivalence, another important factor to consider when interpreting group differences in individual symptom endorsements is the general mental health of the student-athletes. Over half of the items on the PCSS overlap with common symptoms of depressive and anxiety disorders, such as sadness, nervousness, fatigue, sleeping more or less than usual, and difficulty concentrating. Particularly, evidence suggests that individuals of Asian descent are more likely to describe their mental health concerns in somatic terms (Akutsu & Chu, 2006). There was no way for us to determine if the symptoms endorsed on the PCSS by adolescents tested in Mandarin were, in fact, related to psychological distress. The Mandarin language group in this current study is a unique sample, in that these adolescents reside and function in the state of Maine, in a predominantly English-speaking environment, yet chose to complete the assessment in Mandarin. These individuals may have a lower level of English proficiency, which has been linked with a higher level of acculturative stress (Kuo & Roysircar, 2004) that may put them at greater risk for developing symptoms of anxiety and depression (Sirin et al., 2013). We cannot deduce whether any student-athletes met criteria for a current mental disorder, however. Although few student-athletes tested in Mandarin in our sample reported a history of psychiatric treatment (i.e., 1.6%), this value could indicate an absence of access to, or willingness to pursue, evaluation and treatment. Preexisting mental health problems are associated with prolonged recovery following sport-related concussion (for a review, see Iverson et al., 2015); and it is, therefore, important to assess current psychological functioning at the time of baseline and post-concussion evaluations, and appreciate that mental health symptoms may be expressed differently for distinct cultural groups.

The strengths of this study include the large sample size and exact match of participants on key demographic variables that are known to be associated with ImPACT® performances and symptom reporting. There are also important limitations to note. First, the circumstances under which the participants chose to complete the evaluation in Mandarin were not documented. We do not know the country of origin of these participants, their generational status in the United States, or why they chose to complete ImPACT® in Mandarin versus English. We do not have information on the race or ethnicity of these participants. A lower level of English proficiency and the potential role of acculturative stress are a natural and important consideration, but they were not directly measured. Researchers have found an association between acculturation and neuropsychological test performances (Boone et al., 2007), and have stressed the importance of considering degree of acculturation when working with patients of Asian descent (Wong & Fujii, 2004). Acculturation, although not directly measured, could have been related to test performances in our sample. Second, we do not have access to data from Mandarin-speaking adolescents who complete the evaluation in a predominantly Mandarin-speaking environment for comparison, rendering interpretation of these data less clear. Third, we do not have data on whether the adolescents were monolingual, bilingual, or multilingual; we only know they chose to complete ImPACT® testing in Mandarin or English. It is possible that some of these athletes spoke other first or second languages, but this was not documented during data collection. As a result, we do not have an identifiable sample of bilingual adolescents who speak Mandarin but elected to complete ImPACT® testing in English, and we do not know if those athletes who completed ImPACT® in Mandarin had sufficient English language proficiency to complete ImPACT® in English. Future research studies could collect information on current psychological functioning, direct measurement of acculturation, as well as recruiting a comparison group of Mandarin-speaking adolescents who function in a predominantly Mandarin-speaking environment and Mandarin-speaking adolescents who elected to complete ImPACT® in English.

In summary, adolescent student-athletes who completed baseline preseason evaluation with ImPACT® in Mandarin had slightly better performance on the Visual Motor Speed composite score and otherwise largely comparable neurocognitive performances compared to their peers who completed ImPACT® in English. Their total scores on the PCSS were not statistically different from their peers tested in English. However, individual symptoms endorsement analyses revealed higher endorsement rates for several physical symptoms in boys tested in Mandarin that might be associated with somatic expressions of psychological distress. The results of this study highlight three important considerations when conducting preseason baseline assessments for student-athletes who chose to complete the evaluation in a non-English language. First, group differences were modest overall, and fit within a mixed body of literature on the presence or absence of cross-cultural differences in ImPACT® neurocognitive performances and symptom reporting (Kontos et al., 2010; Ott et al., 2014; Wallace et al., 2018). From these results, we infer that the current normative data are reasonably appropriate for use with adolescents who are evaluated in Mandarin at baseline assessments, but future research must evaluate whether group differences are present at post-concussion assessments. Second, it is important to understand translational, conceptual, and metric equivalence of the tests and consider any potential effects on scores (total and item-level). There were some differences in symptom reporting, more so in the boys tested in Mandarin than in the girls tested in Mandarin. And third, it is important to consider students' psychological functioning at the time of baseline and post-concussion evaluations, because students who do not identify with the dominant culture of their sporting environment might experience acculturative stress that could affect symptom endorsement.

## Authors' contributions

Grant L. Iverson: Dr. Iverson conceptualized the study, conceptualized the statistical analyses, assisted with the literature review, wrote portions of the manuscript, edited the manuscript, and approved the final manuscript.

Justin E. Karr: Dr. Karr conceptualized the study, conceptualized the statistical analyses, assisted with the literature review, conducted the statistical analyses, wrote portions of the manuscript, edited the manuscript, and approved the final manuscript.

Yue Hong: Dr. Hong edited the manuscript, assisted with the literature review, assisted with running the statistical analyses, wrote portions of the manuscript, and approved the final manuscript.

Chi-Cheng Yang: Dr. Yang reviewed the Mandarin translation of the battery of tests, edited the manuscript, and approved the final manuscript.

Bruce Maxwell: Dr. Maxwell helped design and coordinate data collection, managed the database of participants, investigated database questions relating to language of administration, and approved the final manuscript.

Paul Berkner: Dr. Berkner helped design and coordinate data collection, wrote the IRB, conceptualized the overall project, investigated database questions relating to language of administration, and edited/approved the final manuscript.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

## Acknowledgements

The data were gathered as part of the Maine Concussion Management Initiative (MCMI) under the direction of the principal investigator Dr. Paul Berkner. The authors thank the Maine Athletic Trainers Association for their collaboration with the MCMI. Dr. Grant Iverson acknowledges unrestricted philanthropic support from the Mooney-Reed Charitable Foundation, National Rugby League, ImPACT® Applications, Inc., and the Spaulding Research Institute.

## Disclosure statement

Dr. Grant Iverson has been reimbursed by the government, professional scientific bodies, and commercial organizations for discussing or presenting research relating to MTBI and sport-related concussion at meetings, scientific conferences, and symposiums. He has a clinical practice in forensic neuropsychology, including expert testimony,

involving individuals who have sustained mild TBIs (including athletes). He has received honorariums for serving on research panels that provide scientific peer review of programs. He is a co-investigator, collaborator, or consultant on grants relating to mild TBI funded by the federal government and other organizations. He has received research support from test publishing companies in the past, including ImPACT® Applications Systems, Psychological Assessment Resources, and CNS Vital Signs. He has received research support from the Harvard Integrated Program to Protect and Improve the Health of NFLPA Members, and a grant from the National Football League. He serves as a scientific advisor for NanoDx™ (formerly BioDirection, Inc.), Sway Medical, Inc., and Highmark, Inc.

## ORCID

Grant L. Iverson http://orcid.org/0000-0001-7348-9570
Justin E. Karr http://orcid.org/0000-0003-3653-332X

## References

Akutsu, P. D., & Chu, J. P. (2006). Clinical problems that initiate professional help-seeking behaviors from Asian Americans. *Professional Psychology: Research and Practice*, *37*(4), 407–415. https://doi.org/10.1037/0735-7028.37.4.407

Bagayogo, I. P., Interian, A., & Escobar, J. I. (2013). Transcultural aspects of somatic symptoms in the context of depressive disorders. In R. D. Alarcón (Ed.), *Cultural psychiatry* (Vol. 33, pp. 64–74). Karger Publishers.

Bailey, C. M., Echemendia, R. J., & Arnett, P. A. (2006). The impact of motivation on neuropsychological performance in sports-related mild traumatic brain injury. *Journal of the International Neuropsychological Society*, *12*(4), 475–484. https://doi.org/10.1017/s1355617706060619

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, *22*(3), 355–365. https://doi.org/10.1016/j.acn.2007.01.010

Brett, B. L., Smyk, N., Solomon, G., Baughman, B. C., & Schatz, P. (2016). Long-term stability and reliability of baseline cognitive assessments in high school athletes using ImPACT at 1-, 2-, and 3-year Test-Retest Intervals. *Archives of Clinical Neuropsychology*, *31*(8), 904–914. https://doi.org/10.1093/arclin/acw055

Brett, B. L., Solomon, G. S., Hill, J., & Schatz, P. (2018). Two-year Test-Retest Reliability in High School Athletes Using the Four- and Two-Factor ImPACT Composite Structures: The Effects of Learning Disorders and Headache/Migraine Treatment History. *Archives of Clinical Neuropsychology*, *33*(2), 216–226. https://doi.org/10.1093/arclin/acx059

Brooks, B. L., Mannix, R., Maxwell, B., Zafonte, R., Berkner, P. D., & Iverson, G. L. (2016). Multiple past concussions in high school football players: Are there differences in cognitive functioning and symptom reporting? *The American Journal of Sports Medicine*, *44*(12), 3243–3251. https://doi.org/10.1177/0363546516655095

Brooks, B. L., Silverberg, N., Maxwell, B., Mannix, R., Zafonte, R., Berkner, P. D., & Iverson, G. L. (2018). Investigating effects of sex differences and prior concussions on symptom reporting and cognition among adolescent soccer players. *The American Journal of Sports Medicine*, *46*(4), 961–968. https://doi.org/10.1177/0363546517749588

Brown, D. A., Elsass, J. A., Miller, A. J., Reed, L. E., & Reneker, J. C. (2015). Differences in symptom reporting between males and females at baseline and after a sports-related concussion: A systematic review and meta-analysis. *Sports Medicine*, *45*(7), 1027–1040. https://doi.org/10.1007/s40279-015-0335-6

Casaletto, K. B., Umlauf, A., Beaumont, J., Gershon, R., Slotkin, J., Akshoomoff, N., & Heaton, R. K. (2015). Demographically corrected normative standards for the English version of the NIH Toolbox Cognition Battery. *Journal of the International Neuropsychological Society*, *21*(5), 378–391. https://doi.org/10.1017/S1355617715000351

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, *19*(22), 3127–3131. https://doi.org/10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cook, N. E., Huang, D. S., Silverberg, N. D., Brooks, B. L., Maxwell, B., Zafonte, R., Berkner, P. D., & Iverson, G. L. (2017). Baseline cognitive test performance and concussion-like symptoms among adolescent athletes with ADHD: Examining differences based on medication use. *The Clinical Neuropsychologist*, *31*(8), 1341–1352. https://doi.org/10.1080/13854046.2017.1317031

Covassin, T., Elbin, R. J. I., Larson, E., & Kontos, A. P. (2012). Sex and age differences in depression and baseline sport-related concussion neurocognitive performance and symptoms. *Clinical Journal of Sport Medicine*, *22*(2), 98–104. https://doi.org/10.1097/JSM.0b013e31823403d2

Elbin, R. J., Schatz, P., & Covassin, T. (2011). One-year test-retest reliability of the online version of ImPACT in high school athletes. *The American Journal of Sports Medicine*, *39*(11), 2319–2324. https://doi.org/10.1177/0363546511417173

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology. General*, *141*(1), 2–18. https://doi.org/10.1037/a0024338

Gerrard, P. B., Iverson, G. L., Atkins, J. E., Maxwell, B. A., Zafonte, R., Schatz, P., & Berkner, P. D. (2017). Factor structure of ImPACT® in adolescent student athletes. *Archives of Clinical Neuropsychology*, *32*(1), 117–122. https://doi.org/10.1093/arclin/acw097

Grover, S., & Ghosh, A. (2014). Somatic symptom and related disorders in Asians and Asian Americans. *Asian Journal of Psychiatry*, *7*(1), 77–79. https://doi.org/10.1016/j.ajp.2013.11.014

Houck, Z. M., Asken, B. M., Bauer, R. M., Caccese, J. B., Buckley, T. A., McCrea, M. A., McAllister, T. W., Broglio, S. P., & Clugston, J. R. (2020). Academic aptitude mediates the relationship between socioeconomic status and race in predicting ImPACT scores in college athletes. *The Clinical Neuropsychologist*, *34*(3), 561–579. https://doi.org/10.1080/13854046.2019.1666923

Houck, Z. M., Asken, B. M., Bauer, R. M., Kontos, A. P., McCrea, M. A., McAllister, T. W., Broglio, S. P., & Clugston, J. R. (2019). Multivariate base rates of low scores and reliable decline on ImPACT in healthy collegiate athletes using Care Consortium norms. *Journal of the International Neuropsychological Society*, *25*(09), 961–971. https://doi.org/10.1017/S1355617719000729

ImPACT Applications Inc. (2011). *Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT®) Test: Technical manual*. ImPACT Applications, Inc.

International Test Commission. (2017). *The ITC guidelines for translating and adapting tests* (2nd ed.). www.InTestCom.org

Iverson, G. L., & Schatz, P. (2015). Advanced topics in neuropsychological assessment following sport-related concussion. *Brain Injury*, *29*(2), 263–275. https://doi.org/10.3109/02699052.2014.965214

Iverson, G. L., Silverberg, N. D., Mannix, R., Maxwell, B. A., Atkins, J. E., Zafonte, R., & Berkner, P. D. (2015). Factors associated with concussion-like symptom reporting in high school athletes. *JAMA Pediatrics*, *169*(12), 1132–1140. https://doi.org/10.1001/jamapediatrics.2015.2374

Jones, N. S., Walter, K. D., Caplinger, R., Wright, D., Raasch, W. G., & Young, C. (2014). Effect of education and language on baseline concussion screening tests in professional baseball players. *Clinical Journal of Sport Medicine*, *24*(4), 284–288. https://doi.org/10.1097/JSM.0000000000000031

Katz, B. P., Kudela, M., Harezlak, J., McCrea, M., McAllister, T., & Broglio, S. P. (2018). Baseline performance of NCAA athletes on a concussion assessment battery: A report from the CARE Consortium. *Sports Medicine*, *48*(8), 1971–1985. https://doi.org/10.1007/s40279-018-0875-7

Kontos, A. P., Elbin, R. J., Covassin, T., & Larson, E. (2010). Exploring differences in computerized neurocognitive concussion testing between African American and White athletes. *Archives of Clinical Neuropsychology*, *25*(8), 734–744. https://doi.org/10.1093/arclin/acq068

Kuo, B. C. H., & Roysircar, G. (2004). Predictors of acculturation for Chinese adolescents in Canada: Age of arrival, length of stay, social class, and English reading ability. *Journal of Multicultural Counseling and Development*, *32*(3), 143–154. https://doi.org/10.1002/j.2161-1912.2004.tb00367.x

Mason, S. J., Davidson, B. S., Lehto, M., Ledreux, A., Granholm, C., & Gorgens, K. A. (2020). A cohort study of the temporal stability of ImPACT scores among NCAA Division I collegiate athletes: Clinical implications of test-retest reliability for enhancing student athlete safety. *Archives of Clinical Neuropsychology*, *35*(7), 1131–1144. https://doi.org/10.1093/arclin/acaa047

McClincy, M. P., Lovell, M. R., Pardini, J., Collins, M. W., & Spore, M. K. (2006). Recovery from sports concussion in high school and collegiate athletes. *Brain Injury*, *20*(1), 33–39. https://doi.org/10.1080/02699050500309817

Moser, R. S., Schatz, P., Neidzwski, K., & Ott, S. D. (2011). Group versus individual administration affects baseline neurocognitive test performance. *The American Journal of Sports Medicine*, *39*(11), 2325–2330. https://doi.org/10.1177/0363546511417114

Nakayama, Y., Covassin, T., Schatz, P., Nogle, S., & Kovan, J. (2014). Examination of the test-retest reliability of a computerized neurocognitive test battery. *The American Journal of Sports Medicine*, *42*(8), 2000–2005. https://doi.org/10.1177/0363546514535901

Ott, S., Schatz, P., Solomon, G., & Ryan, J. J. (2014). Neurocognitive performance and symptom profiles of Spanish-speaking Hispanic athletes on the ImPACT test. *Archives of Clinical Neuropsychology*, *29*(2), 152–163. https://doi.org/10.1093/arclin/act091

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

Ryan, C. (2013). *Language use in the United States: 2011* [American Community Survey Reports]. US Census Bureau.

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

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

Schatz, P., Pardini, J. E., Lovell, M. R., Collins, M. W., & Podell, K. (2006). Sensitivity and specificity of the ImPACT Test Battery for concussion in athletes. *Archives of Clinical Neuropsychology*, *21*(1), 91–99. https://doi.org/10.1016/j.acn.2005.08.001

Shuttleworth-Edwards, A. B., Whitefield-Alexander, V. J., Radloff, S. E., Taylor, A. M., & Lovell, M. R. (2009). Computerized neuropsychological profiles of South African versus U.S. athletes: A basis for commentary on cross-cultural norming issues in the sports concussion arena. *The Physician and Sportsmedicine*, *37*(4), 45–52. https://doi.org/10.3810/psm.2009.12.1741

Silverberg, N. D., Berkner, P. D., Atkins, J. E., Zafonte, R., & Iverson, G. L. (2016). Relationship between short sleep duration and preseason concussion testing. *Clinical Journal of Sport Medicine*, *26*(3), 226–231. https://doi.org/10.1097/JSM.0000000000000241

Sirin, S. R., Ryce, P., Gupta, T., & Rogers-Sirin, L. (2013). The role of acculturative stress on mental health symptoms for immigrant adolescents: A longitudinal investigation. *Developmental Psychology*, *49*(4), 736–748. https://doi.org/10.1037/a0028398

Sue, S., & Sue, D. W. (2000). *Conducting psychological research with the Asian American/Pacific Islander population*. Guidelines for Research in Ethnic Minority Communities. Council of National Psychological Associations for the Advancement of Ethnic Minority Interests (Author). American Psychological Association.

Tsushima, W. T., Tsushima, V. G., Oshiro, R. O., & Murata, N. M. (2017). Role of Native Language in Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) of youth athletes. *Archives of Clinical Neuropsychology*, *32*(4), 450–455. https://doi.org/10.1093/arclin/acx025

Vartiainen, M. V., Peltonen, K., Holm, A., Koskinen, S., Iverson, G. L., & Hokkanen, L. (2021). Preliminary normative study of ImPACT® in Finnish professional male ice hockey players. *Applied Neuropsychology: Adult*, *28*(1), 53–59. Advance Online Publication. https://doi.org/10.1080/23279095.2019.1592175

Wallace, J., Covassin, T., Moran, R., & Deitrick, J. M. (2018). Factors contributing to disparities in baseline neurocognitive performance and concussion symptom scores between black and white collegiate athletes. *Journal of Racial and Ethnic Health Disparities*, *5*(4), 894–900. https://doi.org/10.1007/s40615-017-0437-y

Womble, M. N., Reynolds, E., Schatz, P., Shah, K. M., & Kontos, A. P. (2016). Test-retest reliability of computerized neurocognitive testing in youth ice hockey players. *Archives of Clinical Neuropsychology*, *31*(4), 305–312. https://doi.org/10.1093/arclin/acw011

Wong, T. M., & Fujii, D. E. (2004). Neuropsychological assessment of Asian Americans: Demographic factors, cultural diversity, and practical guidelines. *Applied Neuropsychology*, *11*(1), 23–36. https://doi.org/10.1207/s15324826an1101_4

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