Serial Administration of Clinical Concussion Assessments and Learning Effects in Healthy Young Athletes

By: Tamara C. Valovich McLeod, PhD, ATC, CSCS[*]; David H. Perrin, PhD, ATC†; Kevin M. Guskiewicz, PhD, ATC‡; Sandra J. Shultz, PhD, ATC, CSCS†; Robert Diamond, PhD§; and Bruce M. Gansneder, §

Valovich McLeod, T.C., Perrin, D.H., Guskiewicz, K.M., Shultz, S.J., Diamond, R., Gansneder, B.M. (2004). Serial administration of clinical concussion assessments and learning effects in healthy young athletes. Clinical Journal of Sports Medicine. 14:287-295.

Made available courtesy of LIPPINCOTT WILLIAMS & WILKINS: http://www.cjsportmed.com/pt/re/cjsm/home.htm;jsessionid=J1zZGw4fnJL46vYJG1mKpYFhlsf1BCrYBFsRCftbkqW82vvQmbHP!136317464!181195628!8091!-1

This format of the article is not the final published version.

*****Note: Figures may be missing from this format of the document**

Abstract:

Objective: To determine if serial administration of the Standardized Assessment of Concussion (SAC) and Balance Error Scoring System (BESS) would elicit a learning effect in young athletes and to determine the intratester reliability of scoring the BESS.

Design: A prospective study of 50 healthy young athletes who were assigned to either the control or practice group. All subjects were administered the assessments on 2 occasions, 60 days apart. In addition, subjects in the practice group received serial administration of the assessments on 3 occasions in the week following the initial assessment.

Setting: University Sports Medicine/Athletic Training Research Laboratory.

Subjects: Fifty uninjured young athletes between 9 and 14 years of age.

Main Outcome Measured: Scores on 2 clinical concussion assessments, the SAC and the BESS.

Results: We found a significant learning effect upon serial BESS testing in the practice group. BESS error scores were significantly lower than baseline (15.0 ± 4.6) on days 5 (11.3 ± 5.33), 7 (12.4 ± 6.2), and 60 (12.6 ± 6.2). We also found a significant learning effect upon the day 60 BESS administration across all subjects. We did not find a practice or learning effect with serial

---

[*] From the Arizona School of Health Sciences, Mesa, AZ; the †University of North Carolina at Greensboro, Greensboro, NC; the ‡University of North Carolina at Chapel Hill, Chapel Hill, NC; and the §University of Virginia, Charlottesville, VA.
Received for publication April 2003; accepted January 2004.
Reprints: Tamara C. Valovich McLeod, PhD, ATC, CSCS, Assistant Professor, Department of Sports Health Care, Arizona School of Health Sciences, 5850 E. Still Circle, Mesa, AZ 85206 (e-mail: tmcleod@atsu.edu).

SAC test administration. The intratester reliability of the investigator's ability to score repeated observations of the same BESS test ranged from 0.87 to 0.98.

Conclusions: Our results demonstrated that serial administration of the BESS elicited a learning effect, which was more prominent during the tandem conditions. Clinicians utilizing the BESS as a measure of postural stability should be aware of the potential for improvement with repeated testing. Clinicians should not expect improvement with the SAC, as scores remained relatively stable across all trials.

Keywords: mild head injury, SAC, BESS, practice effects, children

Article:

Sports-related mild head injury (MHI), or concussion, has been a prominent issue in the sports medicine literature of the late 1990s and early into the new millennium. It is an important public health issue because of the large number of athletes sustaining concussive injuries, the relatively young age of the individuals at the time of injury, and the potential for cumulative effects of repeated injuries. 1 The Centers for Disease Control and Prevention estimate that 300,000 sports-related head injuries occur each year in the United States. 2 The rates of traumatic brain injury reported by the Centers for Disease Control and Prevention's traumatic brain injury surveillance system demonstrated an overall injury rate of 2.6/100,000. However, the incidence of MHI was substantially greater for males in the ages groups 5 to 14 years old (7.5/100,000) and 15 to 24 years old (10/100,000). 1

The incidence and prevalence of injuries at the youth sports level are often overlooked or not reported. In light of the lack of population-based data, Kelly et al 3 examined the incidence of sport and recreational head injuries treated in the emergency department of a large Canadian city over a period of 1 year. Sports-related head injury accounted for 3% of all sport-related injuries and 24% of all serious head injuries. 3 Specifically, sport-related head injury represented a substantial percentage of total head injuries in populations <10 years old (18.2%), 10 to 14 years old (53.4%), and 15 to 19 years old (42.9%). 3 The percentage of these sport-related head injuries out of all sports-related injuries were 2.8% for children <10 years, 3.7% for children 10 to 14 years, and 4.2% for children 15 to 19 years. 3 Another study reported that 15% of the children (8.34 ± 5.31 years) who were admitted to hospitals following MHI had a sports-related mechanism of injury. 4 These findings demonstrate that sport-related injuries, specifically head injuries, are commonly seen in emergency departments and make up a significant portion of injuries to children.

Athletes participating at all levels of athletic competition are at risk for sustaining a concussion. Due to the potential catastrophic, 2,5 cumulative, 6 and long-term sequelae 7,8 following MHI, it is important for those evaluating concussed athletes to have objective assessments to follow the athlete's recovery adequately and avoid releasing the athlete to activity prematurely. Recent developments in the area of sports-related MHI have provided clinicians with several tools for assessments of MHI. Assessment of cognitive function by means of neuropsychological testing 9,10 and the Standardized Assessment of Concussion (SAC) 11–13 are becoming commonplace. Moreover, postural stability assessments using force platforms 10,14 and, more recently, clinical balance tests 15 have been found to be sensitive to changes detected following MHI in college and high school populations.

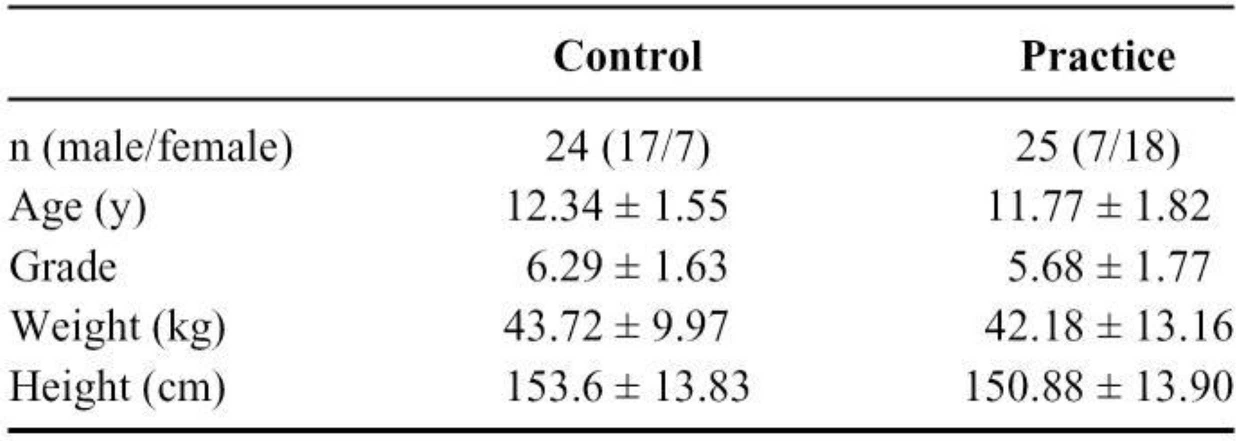
In clinical practice, these assessments are used on numerous days following MHI to assess the recovery of the athlete and make decisions regarding when an athlete can safely return to activity. 10,11,13–18 Of consideration is that previous investigations with neuropsychological test batteries 16–19 and measures of postural stability 20–24 have revealed a practice effect in which the athlete's scores improve because of familiarization with the assessment. Practice effects are typically defined as some improvement in performance between concurrent test sessions based on familiarity with the procedures and/or previous exposure to the assessment, 25 whereas learning effects relate to the retention of the improvement over a period of time. Both practice and learning effects can be a confounding factor in the interpretation of test scores. 25

It is unclear whether practice and/or learning effects exist with repeat administration of the SAC or Balance Error Scoring System (BESS) in youth sports participants. The lack of data in this younger population with respect to normative values, as well as following injury and during recovery, warrants research regarding various concussion assessment tools in younger athletes. Therefore, the purpose of this investigation was to determine the magnitude of learning with repeated administration of these assessments in healthy young athletes. The secondary purposes of this investigation were to establish intratester reliability of scoring the BESS and to establish normative data for the SAC and the BESS in a population of young athletes. We hypothesized that we would find a learning effect with repeat administration of the SAC and the BESS, but that scores at day 60 would return to baseline. We also expected to find good intratester reliability of scoring the BESS errors.

METHODS

Subjects

Fifty healthy youth sports participants were recruited from the local community to participate in this study and assigned to either the control (n = 25) or practice (n = 25) group. The day 60 data are excluded for 1 of the control subjects since he suffered a MHI during the course of the investigation. Subject characteristics for the 49 complete subjects are presented in Table 1. Male and female participants were selected on the following general criteria, through self-report and parent report: (1) participation in recreational or competitive athletics, (2) no lower extremity musculoskeletal injuries in the past 6 months, (3) no history of head injury, (4) no existing visual, vestibular, or balance disorders, and (5) absence of attention deficit disorder or learning difficulty. Prior to participation, the parent/guardian of the participants and the participants read and signed an informed consent form approved by the University's Institutional Review Board for the Protection of Human Subjects.



	Control	Practice
n (male/female)	24 (17/7)	25 (7/18)
Age (y)	12.34 ± 1.55	11.77 ± 1.82
Grade	6.29 ± 1.63	5.68 ± 1.77
Weight (kg)	43.72 ± 9.97	42.18 ± 13.16
Height (cm)	153.6 ± 13.83	150.88 ± 13.90

TABLE 1. Subject Characteristics

Instrumentation

Balance Error Scoring System

The BESS consisted of 6 separate 20-second balance tests that the subjects performed in different stances and on different surfaces. A 16-in × 16-in piece of medium density foam (Exertools, Novato, CA) was used to create an unstable surface for the subjects. The test consisted of 3 stance conditions (double leg, single leg, and tandem stance) and 2 different surfaces (firm and foam). Errors were recorded as the quantitative measurement of postural stability under different testing conditions. These included subjects opening their eyes; stepping, stumbling, or falling from the test position; removing their hands from their hips; moving their hip into >30° of flexion or abduction; lifting their toes or heels from the test surface; or remaining out of the test position for >5 seconds.

Standardized Assessment of Concussion

The SAC measured orientation, immediate memory, concentration, and delayed recall. 13 We used the 3 alternate forms (A, B, and C) of the SAC, which were counterbalanced between subjects for each test session and across all test sessions for each subject. For each of the test sessions, 1/3 of the subjects were administered form A, 1/3 form B, and the final third form C. The test forms were also counterbalanced across the serial test sessions. Scores for each of the SAC subtests and total score were recorded.

Video Camera

A digital video camera (Sony DV Handycam; New York, NY) was used to record the BESS performance. These videos were used to establish intratester reliability in the scoring of BESS errors.

Procedures

Testing was conducted in a university sports medicine/athletic training research laboratory or at the child’s school or home. For each subject, testing was performed at the same location for all test administrations. Testing for the control group consisted of 2 test sessions, 60 days apart. The practice group was tested on 5 occasions (baseline and days 3, 5, 7, and 60). A single investigator (TVM) performed all test administrations to ensure optimal consistency of procedures.

The SAC was administered first and took approximately 5 to 7 minutes to administer. It was administered as described by McCrea et al.13,26 The 3 alternative forms of the SAC were used and counterbalanced among subjects and test sessions.

Balance Error Scoring System testing took approximately 10 minutes per subject. The order of trials followed the standard BESS format, which progressively increased the demands placed on the sensory systems; i.e., double-leg, single-leg, and tandem on firm, then foam. Subjects were asked to assume the required stance by placing their hands on their iliac crests, and upon eye closure, the test was begun. During the single leg stances, the subjects were asked to maintain the contralateral limb in 20° of hip flexion and 40° of knee flexion. Additionally, subjects were asked to stand quietly and as motionless as possible in the stance position, keeping their hands on their iliac crests and eyes closed. Subjects were told that upon losing their balance, they should make any necessary adjustments and return to the stance position as quickly as possible. Performance was assessed by individual trial scores and by adding the error points for each of the 6 trials. Trials were considered incomplete if the subject could not sustain the stance position for longer than 5 seconds. In these instances, they were then assigned a standard maximum score for that stance.23

To ensure that tester bias was not introduced into the investigation, a video camera recorded the performance of the BESS for 20 subjects during the baseline test session. The primary investigator initially scored BESS errors as the participants performed the test and then again for a second time by watching the videotape of the BESS performance. Trials on the videotapes were viewed in a randomized manner approximately 1 week after the live test session.

Statistical Analysis

All data were analyzed using SPSS 10.0 software (SPSS, Chicago, IL). To determine if a learning effect was present across the serial test days in the practice group, separate repeated-measures analyses of variance (ANOVAs) with 1 within (day) were used for total score and each subtest of the SAC. To correct for the multiple analyses, a Bonferroni adjustment was made, and significance was set at $P < 0.01$. A separate repeated-measures ANOVA with 3 within (day, surface, stance) was used for the BESS. Significance level was set a priori ($P < 0.05$). Significant differences were examined further with the Dunnett post hoc analysis, since all measures were compared with their baseline measure.27

Due to the concussion suffered by 1 of the control subjects, the group analyses presented reflect 24 subjects in the control group and 25 subjects in the practice group. To determine if differences existed between baseline and the day 60 scores between groups, 2 separate repeated-measures ANOVAs were used for the SAC and the BESS. A mixed-model repeated-measures ANOVA with 1 within (day) and 1 between (group) was used to determine significance for total SAC score. A second mixed-model repeated measures ANOVA with 3 within (day, surface, stance) and 1 between (group) was used for the total BESS score. The Tukey post hoc test was used to examine significant differences further. Level of significance ($P < 0.05$) was set a priori.

Intraclass correlation coefficients (ICCs) were used to determine the reliability and consistency of the BESS scoring between the live and videotaped performance of the participants. ICC (2, 1) was used to evaluate each test condition across the 2 modes (live and video) of scoring. Standard error of the measure (SEM) was also recorded.

RESULTS

Serial Administration

Balance Error Scoring System

The repeated-measures ANOVA within the practice group found significant main effects for time ($F_{4,96}$ = 5.643; $P < 0.0001$), surface ($F_{1,24}$ = 202.343; $P < 0.0001$), and stance($F_{2,48}$ = 197.594; $P < 0.0001$). BESS scores across time were 14.96 ± 4.61 at baseline and 13.96 ± 5.27, 11.28 ± 5.30, 12.410 ± 6.16, and 12.64 ± 6.20 on days 3, 5, 7, and 60, respectively. Post hoc analysis for time showed that total errors on days 5, 7, and 60 were significantly less than baseline (Fig. 1). Post hoc analysis for surface demonstrated differences between the firm and foam and for stance showed significant differences between all 3 stances. We found significant 2-way interactions for time by stance ($F_{8,192}$ = 2.095; $P = 0.038$) and surface by stance ($F_{2,48}$ = 29.666; $P < 0.0001$). Post hoc analyses for the interactions revealed that the day 7 tandem stance score was significantly less than the tandem baseline score for the time by stance interaction (Fig. 2), and scores on the foam surface were significantly greater than the firm for the single-leg and tandem stances for the surface by stance interaction.

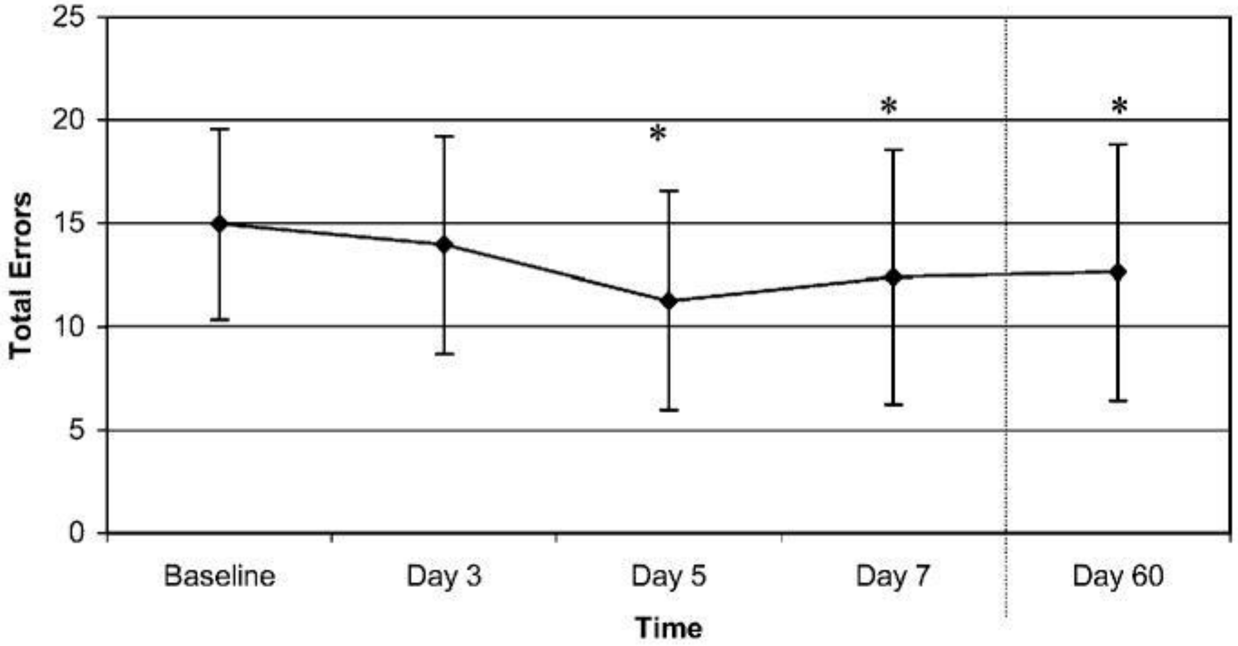


FIGURE 1. Total BESS errors (mean and SD) across the test sessions in the practice group. *Significantly less than baseline ($P < 0.05$).

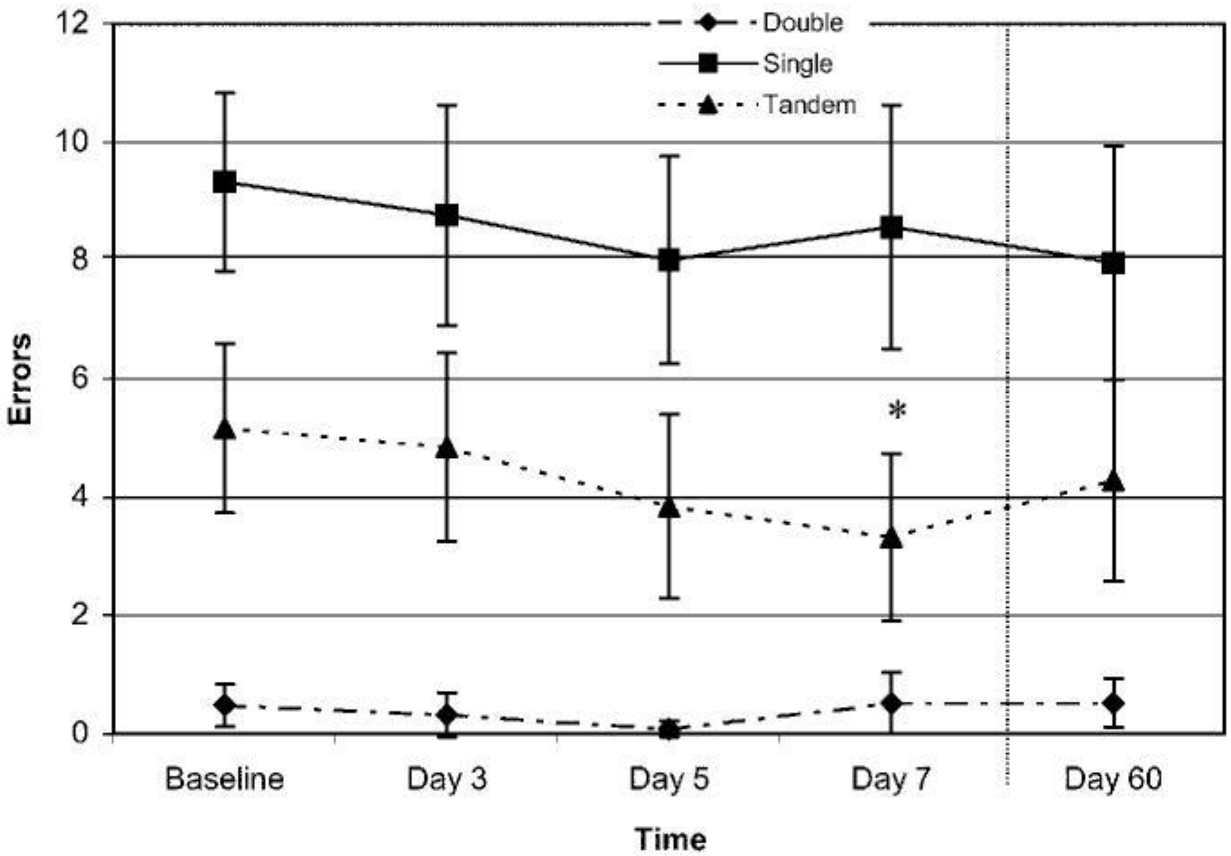


FIGURE 2. Time by stance interaction for the BESS across test sessions in the practice group. *Significantly less than tandem baseline ($P < 0.05$).

Standardized Assessment of Concussion

The repeated-measures ANOVA for total SAC score within the practice group revealed no significant main effect for time ($F_{4,96} = 2.517$; $P = 0.046$; [beta]= .695;Fig. 3). SAC scores across the 5 test days were 27.00 ± 2.06, 26.00 ± 2.08, 26.08 ± 1.91, 26.96 ± 2.11, and 26.96 ± 1.81. Results of the analysis for each subtest of the SAC demonstrated no significant differences for time for the orientation ($F_{4,96} = .302$; $P = 0.876$; [beta] = .115), immediate memory ($F_{4,96} = 2.026$; $P = 0.097$; [beta] = .587), concentration ($F_{4,96} = 1.498$; $P = 0.209$; = .449), or delayed recall ($F_{4,96} = 3.396$; $P = 0.012$; [beta] = .835) subtests.

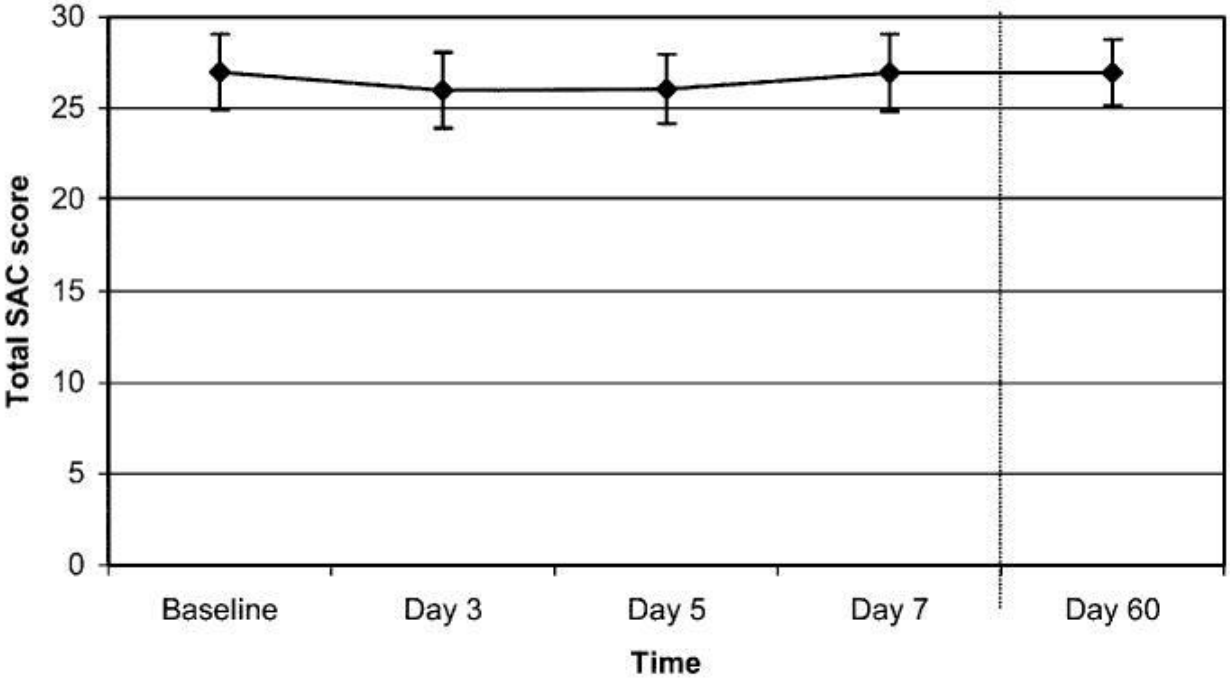


FIGURE 3. Total SAC score across the test sessions in the practice group.

Group Comparisons

Balance Error Scoring System

We found a significant time by stance by group interaction ($F_{2,94} = 3.110$; $P = 0.049$) and surface by stance interaction ($F_{2,94} = 27.858$; $P < 0.0001$). The significant 3-way interaction resulted from a group by time interaction on errors for the double-leg stance that was not found for the single-leg or tandem stances. Tukey post hoc tests for the 2-way interaction for the double-leg stance indicated that the control group made more errors than the practice group at baseline, but there was no difference between groups at day 60. We also found a significant main effect for time ($F_{1,47} = 8.274$; $P = 0.006$), surface ($F_{1,47} = 192.402$; $P < 0.0001$), and stance ($F_{2,94} = 298.884$; $P <0.0001$). We did not find a significant main effect for group($F_{1,47} = .614$; $P = 0.437$; [beta] = .120) or a significant time by group interaction ($F_{1,47} = .614$; $P = 0.437$; [beta] = .120;Fig. 4).Post hoc analysis for time revealed that all subjects combined scored fewer errors on the day 60 test compared with baseline. In addition, fewer errors were scored on the firm surfaces when compared with the foam. A significant difference in the number of errors scored was also found between all 3 of the stances. Subjects scored the fewest errors on the double-leg stances, followed by the single-leg stances, and scored the most errors on the tandem stances.

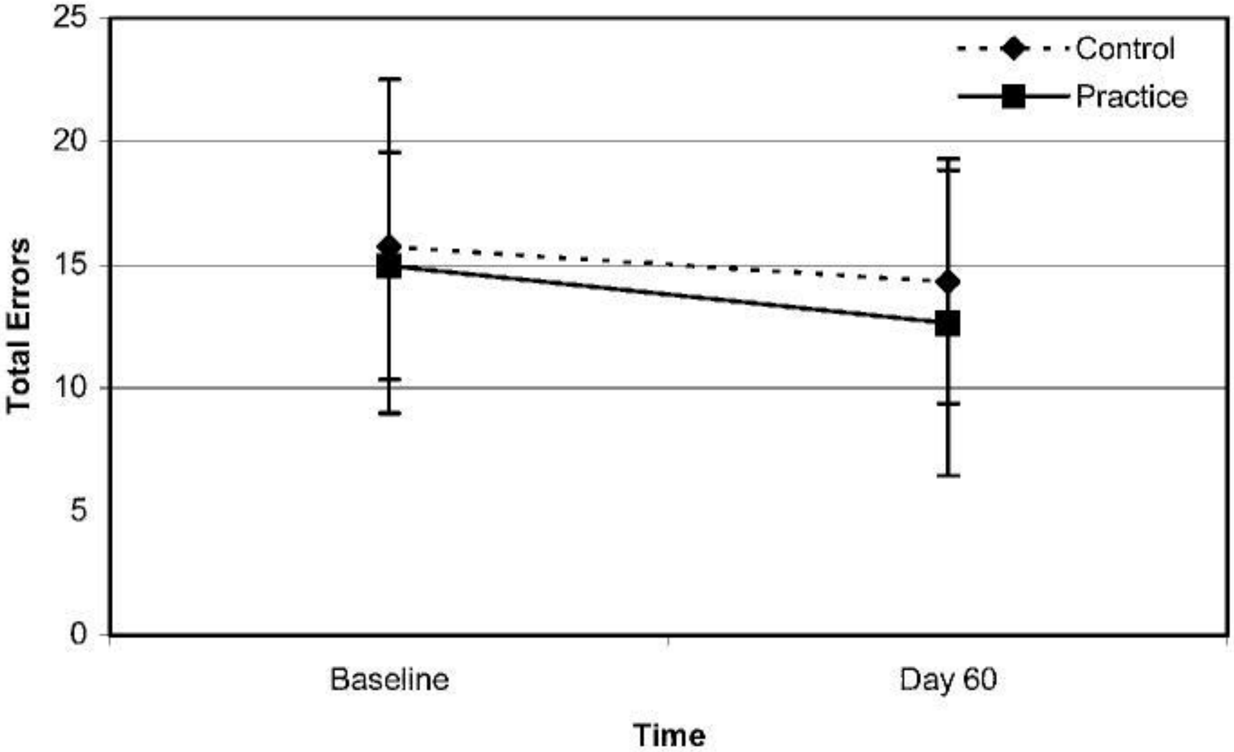


FIGURE 4. BESS group comparisons at baseline and day 60.

Standardized Assessment of Concussion

We found no significant differences between the control and practice groups on performance of the SAC for time ($F_{1,47}$ = .714; P = 0.402; [beta] = .131), group ($F_{1,47}$ = .028; P =0.868; [beta] = .053), or time by group interaction ($F_{1,347}$ = .960; P = 0.332; [beta] = .160;Fig. 5).

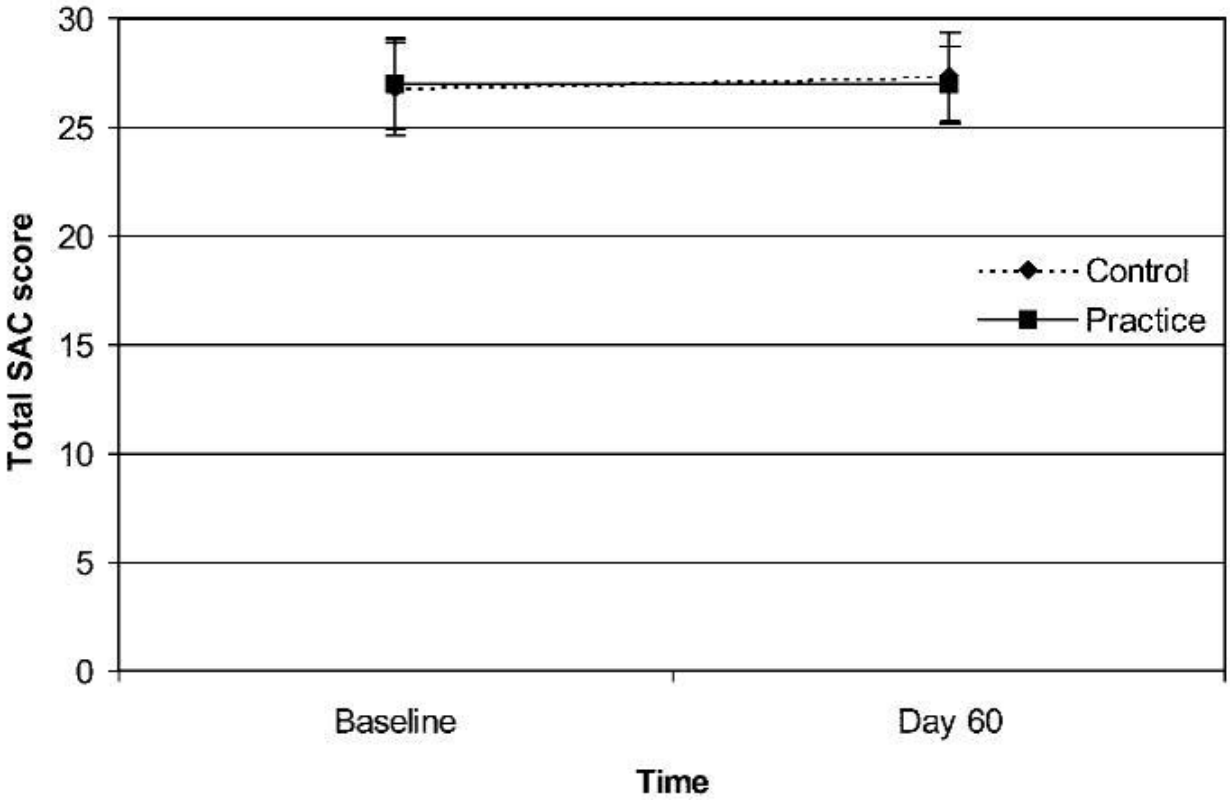


FIGURE 5. SAC group comparisons at baseline and day 60.

Intratester Reliability

The ICCs and SEMs are presented in Table 2. High ICCs were found for all of the test conditions and ranged from 0.87 to 0.98. Intratester reliability could not be calculated for the double firm condition since all subjects scored 0 (no errors) on both the live and video formats. The SEMs ranged from 0.28 to 0.77 for the 6 conditions, and the SEM was 1.01 for total BESS score.

Condition	ICC	SEM
Double-firm	—	—
Single-firm	.95	.51
Tandem-firm	.94	.28
Double-foam	.91	.29
Single-foam	.87	.77
Tandem-foam	.94	.49
Total BESS Score	.98	1.01

TABLE 2. Intratester Reliability ICCs (2,1) for the 6 Conditions and BESS Total Score

DISCUSSION

Serial Administration

Balance Error Scoring System

Our main finding with serial administration of the BESS was a learning effect in which subjects scored fewer errors on days 5 and 7 compared with baseline. These findings agree with previous BESS studies in both high school 20 and collegiate 24 athletes (Fig. 6). In addition, our findings resemble those of other methods of postural stability assessment in children in which a learning effect was evident following repeat administration. 28–30 The improvement that we found is likely due in part to the time interval between test sessions and the nature of the balancing task that we had the children perform.

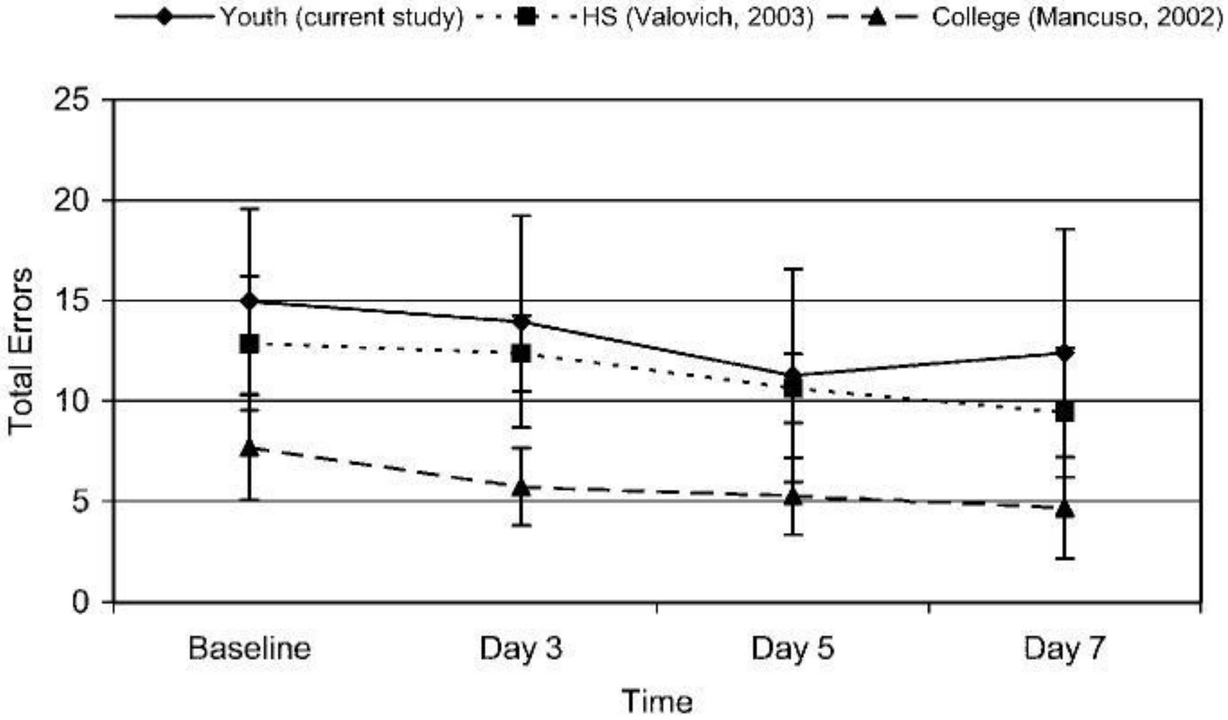


FIGURE 6. BESS scores within the practice group of the current investigation plotted against scores from a high school 20 and collegiate 24 populations.

This investigation used a 1-day interval between subsequent test administrations, which might have increased the potential for a learning effect. In the literature regarding learning, there tends to be a more robust effect with shorter time intervals regardless of test paradigm. 20,22,30 Our earlier investigation of high school athletes using the same serial assessment protocol found decreased BESS errors with each subsequent administration and significantly fewer errors on days 5 and 7 compared with baseline. 20 Shorter time intervals have also been implicated in performance improvements in healthy adults using stabilimetric testing 22 and static and dynamic measures of postural stability. 31 Therefore, with serial assessments of the BESS to monitor recovery following MHI, clinicians need to be aware of the test-retest interval and should expect increased improvement in healthy individuals or under nonpathologic conditions when this interval is short.

Although we reported a significant learning effect for total BESS errors, each of the separate BESS conditions did not demonstrate the same performance curve upon serial administration. Our analyses demonstrated a learning effect only for the tandem stance, with scores on the day 7 tandem stance being significantly less than baseline scores. This finding is likely the result of a tandem stance being the most challenging for children this age. Scores for the single leg stance, although not significant, showed a trend of fewer errors with subsequent test administrations, while the double leg scores remained relatively constant across serial assessments. The results from this investigation support the results found using the BESS and other measures of postural stability with serial testing, which show improved performance on tasks that are novel or challenging for the individual. 20,22,24,28,29

Our findings that the youth sports participants improved the most on the tandem stance are similar to those of previous postural stability investigations in children. 28,29 In an investigation of the Pediatric Clinical Test of Sensory Interaction for Balance, Westcott et al 29 evaluated test-retest reliability of 2 stances in children 4 to 9 years of age. For all 6 conditions of the test, children scored higher on the retest session, and this difference was more evident during the tandem stance condition compared with standing with the feet together. The magnitude of the differences between the test and retest condition were larger and learning effects were noted for 3 of the combined sensory conditions (vision absent, vision inaccurate, and somatosensory inaccurate) of the tandem stance, while no differences were found for the double-leg stance. 29 In addition, there was little variability among scores for the feet-together condition, whereas the tandem condition showed instability across time. These results, along with our findings, might indicate that the tandem stance is the most challenging to children and is therefore suspect to learning effects.

These results differ slightly from those of previous investigations of high school 20 and collegiate athletes 24 using the BESS. In the investigation of practice effects and high school athletes, the foam surface and single-leg stance conditions represented the surface and stance on which the greatest improvement was found. Scores on the foam surface as well as the single-leg stance were significantly less than baseline by day 7. 20 Utilizing a 7-session protocol, Mancuso et al 24 found significant learning on the single leg-firm condition by day 3, the single leg-foam condition between days 3 and 5, the single-leg tremor box between days 1 and 2, and on the tandem tremor box condition between baseline and day 1. Findings across these studies have been consistent in that learning effects are most evident with more challenging conditions that involve an increased skill level.

Standardized Assessment of Concussion

We did not find a learning effect with repeat administration of the SAC. This finding agrees with previous research involving serial administration of the SAC in high school athletes 20 and in a mixed high school and collegiate population. 13 The scores of the youth sports participants during the serial assessments mirrored those of the high school athletes reported by Valovich et al, 20 as shown in Figure 7. Although the values approached significance for total SAC score and delayed recall across time, these trends were opposite our hypothesis (i.e., scores at day 3 were worse than baseline), further rejecting the notion of a practice effect.

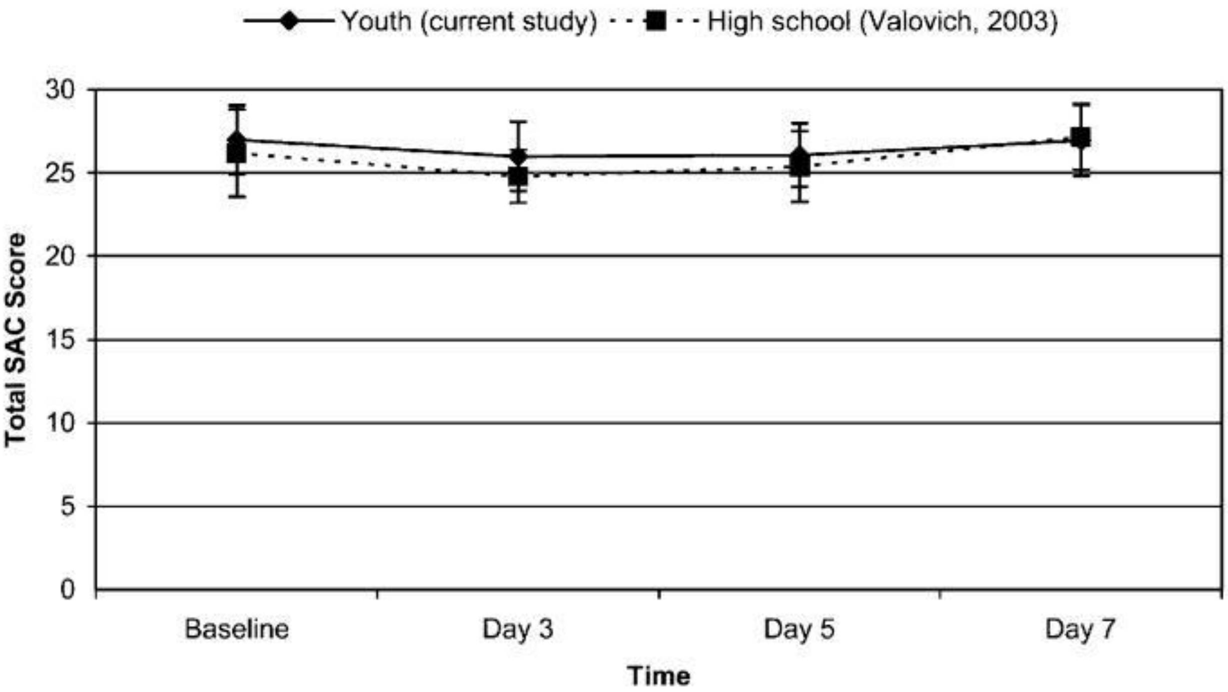


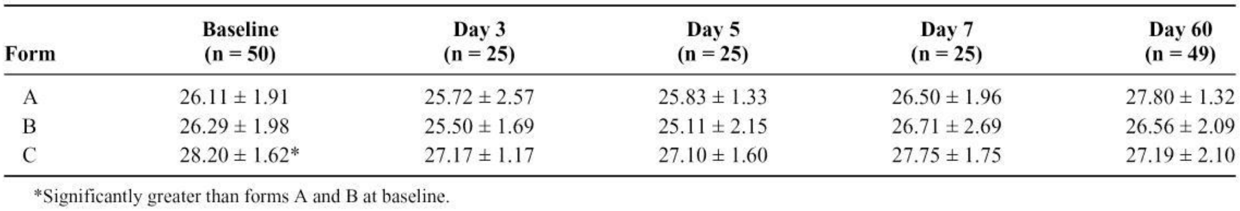
FIGURE 7. SAC scores within the practice group of the current investigation plotted against scores from a high school population. 20

These findings with serial SAC assessments differ from those of the literature involving more complex neuropsychological assessments in high school or collegiate athletes. Repeat administration of the Stroop Test, 10 Trail Making Test, 10,14,19 Wechsler Digit Span Test, 10,14 Digit Symbol Substitution Test, 19 and Paced Auditory Serial Addition Task 19 have all demonstrated learning effects.

In fact, the lack of a learning effect was reported only by Oliaro et al 17 on the Hopkins Verbal Learning Test, on which subjects scored significantly lower in the third test administration compared with the first. This decrease was attributed to a possible lack of motivation on the part of the subjects during the third test session. Another plausible explanation for their results and our current and previous findings with the SAC is the use of alternative forms, which may have negated a learning effect. These alternate forms use different words for the immediate and delayed recall memory and different numbers in the concentration subtest and were created to alleviate any practice effects from serial assessments. 26

In our investigation, we counterbalanced the SAC forms for all subjects and test administrations in an effort to replicate how this assessment is given clinically. This may be 1 explanation as to why we did not find a learning effect upon serial administration. However, post hoc analysis of the SAC scores by test form demonstrated higher SAC scores on form C. In fact, scores on form C at baseline were significantly higher than on A and B (Table 3). This could also have contributed to a lack of learning effects in that if subjects were given form C at baseline, there is little expectation that they would improve on subsequent days when they were given form A or B, the harder forms. Although previous work has shown these forms to be equivalent, 26 another investigation also found form C to have marginally significantly higher scores than forms A and

B. 13 These findings demonstrating that significant differences exist between form C of the SAC and forms A and B persist despite attempts to equalize the test stimuli on a number of factors.



Form	Baseline (n = 50)	Day 3 (n = 25)	Day 5 (n = 25)	Day 7 (n = 25)	Day 60 (n = 49)
A	26.11 ± 1.91	25.72 ± 2.57	25.83 ± 1.33	26.50 ± 1.96	27.80 ± 1.32
B	26.29 ± 1.98	25.50 ± 1.69	25.11 ± 2.15	26.71 ± 2.69	26.56 ± 2.09
C	28.20 ± 1.62*	27.17 ± 1.17	27.10 ± 1.60	27.75 ± 1.75	27.19 ± 2.10

*Significantly greater than forms A and B at baseline.

TABLE 3. SAC Scores by Form

A second possible explanation for the lack of learning effects in this investigation may be due to our healthy subject population and a ceiling effect on certain SAC subtests. Although there is room for improvement in total SAC score before subjects would hit the ceiling of 30 points, performance on each of the SAC subtests did differ in that some were easier than others. We found room for improvement in the concentration and delayed recall subtests; however, our subjects seemed to hit a ceiling on the orientation and immediate memory sections. Mean scores for these 2 sections were 4.72 ± 0.41 out of 5.0 and 14.68 ± 0.65 out of 15.0, respectively. These scores leave little room for improvement in these 2 sub-tests, therefore lessening the potential for a learning effect in overall SAC score.

Based on the analysis of our baseline through day 7 data in our youth sports participants and the findings of our previous work in high school athletes, there seems to be no practice or learning effect evident with repeat administration of the alternate forms of the SAC in healthy athletes. Clinicians should expect similar scores across test sessions in healthy controls and an improvement in performance in athletes during recovery from MHI relative to their most immediate postinjury score. If concussed athletes do not show an improvement in SAC performance following injury, one should suspect lingering pathology and refer for a more complex neuropsychological assessment. In addition, to account for the differences in the difficulty of the alternate SAC forms, it is recommended to administer form A at baseline and counterbalance forms A, B, and C during follow-up assessments.

Group Comparisons

Balance Error Scoring System

Our main finding of the group comparison analysis was that all subjects demonstrated improved BESS performance and fewer errors at the day 60 administration compared with baseline. A learning effect was found in that both the control and practice groups improved their BESS performance in a similar fashion. These findings are contrary to the investigation of high school athletes in which neither group scored a significantly different number of errors from baseline on the day 30 test session. 20 However, they do agree with the results of Mancuso et al, 24 who found that improvements were retained in the experimental group on the single-firm, single-foam, single-tremor, and tandem-tremor conditions 90 days following baseline. Our results indicate that the practice group was able to retain the learning effects gained by the serial administration of the BESS and, more surprisingly, that the control group was also able to improve their performance significantly with only 1 previous test administration. Although the time period between assessments was shorter (1 week), Westcott et al 29 found similar learning effects during the tandem stance condition in that subjects were able to improve performance in

3 conditions with only 1 previous exposure to the balance tasks. These results may indicate that children are better able to learn and retain a balance task with fewer exposures to the particular task and support the notion of dual baseline assessments, especially for this younger age group.

Intratester Reliability

Although the BESS attempts to objectify postural stability through a clinical balance assessment, there is still the potential for subjectivity with the scoring of the errors. Subtle differences in postural stability following MHI might be overlooked if the scoring of the errors is not consistent among testers or across test days with the same tester. Intertester reliability of BESS scoring was previously found to be good, and ranged between 0.78 to 0.96 for all the stances except double-firm. 23 Our investigation of intratester reliability of BESS scoring demonstrated high ICCs ranging from 0.87 to 0.98, indicating that the same tester was reliable across the test days and that the learning effects found in this investigation are not likely the result of unreliable scoring of BESS errors. Both the intertester and intratester reliability of scoring errors is important in the clinical use of this assessment tool. In assessments given following MHI, clinicians need to be confident that they will score errors consistently from 1 test administration to another. In addition, in settings in which more than 1 athletic trainer is responsible for scoring the BESS, consistency among testers is of the utmost importance to ensure safe return-to-play decision-making.

Limitations

We acknowledge that there are some limitations to our study design that could possibly affect our findings. We used only 25 subjects in our practice group, and this small number of subjects could possibly limit the generalizability of our results to all young athletes. However, previous research has produced similar findings with even smaller sample sizes. 20 In addition, we did not have the means to control for the subject’s motivation to participate in the serial testing. Although the interactions between the tester and each subject remained similar through all test sessions, the fact that the SAC resembles a school test could have affected subject motivation with the serial testing. A possible limitation to the clinical interpretation of our BESS results and the time interval between test sessions is the short interval between baseline and the hypothetical first postinjury assessment (day 3). In clinical practice, the baseline test is typically done during the preseason along with the pre-participation physical examination. If the MHI occurs at some point during the season, there is likely a longer time interval between the baseline and the serial administrations to monitor recovery. Although this baseline-to-injury interval will vary for each individual injury, the recovery timeline used clinically and in most sports-related concussion investigations is similar to the intervals we used from days 3 to 7.

Clinical Relevance

This investigation addressed the utility of 2 common clinical MHI assessment tools, the SAC and the BESS, in a youth sports population. For both of these assessments, scores for our youth sports participants closely resembled the scores reported in adult populations. These results indicate that the SAC and BESS are appropriate tests to administer to young athletes between 9 and 14 years of age to obtain baseline measures of mental status and balance ability.

Clinicians employing serial assessments to monitor an athlete following MHI should be aware of the potential for learning effects with the BESS and should expect the performance of an athlete to improve upon each subsequent administration. In addition, improvements in balance ability as scored by the BESS should be expected once a child has been exposed to the assessment through their baseline test. The use of dual baseline assessments should be incorporated to allow familiarization with the BESS and help alleviate any possible confounding factors due to previous exposure to the task. Because of the continual development of postural stability through the age of 10 years, clinicians should readminister baseline BESS assessments at the beginning of each competitive season.

We did find a significant difference in scores between the 3 SAC forms, with subjects scoring significantly higher on form C at baseline. Clinicians utilizing the SAC can feel confident that repeat administration of this instrument will not elicit a learning effect, and scores should be expected to improve as the athlete recovers from the MHI. It is recommended that clinicians administer SAC form A at baseline and counterbalance all 3 forms during serial assessments following MHI so that any decrements following injury can be attributed solely to the MHI, not the test form.

We feel that this article lays a foundation for the study of concussion assessments in younger athletes. We feel that our data begin to address the issue of concussion assessment in young athletes with respect to normative data for the SAC and BESS. Future studies should continue to address the issue of recovery following a concussion. The definition of recovery varies and is often assumed to be the return to preseason baseline or some expected percentage of improvement. Future studies could use more sophisticated statistical techniques such as the Reliable Change Index to help distinguish between learning and recovery in this population.

REFERENCES

1. Thurman DJ, Branche CM, Sniezek JE. The epidemiology of sports-related traumatic brain injuries in the United States: recent developments. J Head Trauma Rehabil. 1998;13:1–8.
2. Centers for Disease Control and Prevention. Sports-related recurrent brain injuries—United States. JAMA. 1997;277:1190–1191.
3. Kelly KD, Lissel HL, Rowe BH, et al. Sport and recreation-related head injuries treated in the emergency department. Clin J Sports Med. 2001; 11:77–81.
4. Adams J, Frumiento C, Shatney-Leach L, et al. Mandatory admission after isolated mild closed head injury in children: is it necessary?J Pediatr Surg. 2001;36:119–121.
5. Saunders RL, Harbaugh RE. The second impact in catastrophic contact-sports head trauma. JAMA. 1984;252:538–539.
6. Gaetz M, Goodman D, Weinberg H. Electrophysiological evidence for the cumulative effects of concussion. Brain Inj. 2000;14:1077–1088.
7. Alves W, Macciocchi SN, Barth JT. Postconcussive symptoms after uncomplicated mild head injury. J Head Trauma Rehabil. 1993;8:48–59.
8. Gulbrandsen GB. Neuropsychological sequelae of light head injuries in older children 6 months after trauma. J Clin Neuropsychol. 1984;6:257–268.
9. Collins MW, Grindel SH, Lovell MR, et al. Relationship between concussion and neuropsychological performance in college football players. JAMA. 1999;282:964–970.
10. Guskiewicz KM, Riemann BL, Perrin DH, et al. Alternative approaches to the assessment of mild head injury in athletes. Med Sci Sports Exerc. 1997;29:S213–S221.

11. McCrea M. Standardized mental status testing of acute concussion. Clin J Sports Med. 2001;11:176–181.
12. McCrea M. Standardized mental status testing on the sideline after sport-related concussion. J Athl Train. 2001;36:274–279.
13. McCrea M, Kelly JP, Randolph C, et al. Standardized assessment of concussion (SAC): on-site mental status evaluation of the athlete. J Head Trauma Rehabil. 1998;13:27–35.
14. Guskiewicz KM, Ross SE, Marshall SW. Postural stability and neuropsychological deficits after concussion in collegiate athletes. J Athl Train. 2001;36:263–273.
15. Riemann BL, Guskiewicz KM. Effects of mild head injury on postural stability as measured through clinical balance testing. J Athl Train. 2000; 35:19–25.
16. Macciocchi SN, Barth JT, Alves W, et al. Neuropsychological functioning and recovery after mild head injury in collegiate athletes. Neurosurgery. 1996;39:510–514.
17. Oliaro SM, Guskiewicz KM, Prentice WE. Establishment of normative data on cognitive tests for comparison with athletes sustaining mild head injury. J Athl Train. 1998;33:36–40.
18. Onate JA, Guskiewicz KM, Riemann BL, et al. A comparison of sideline versus clinical cognitive test performance in collegiate athletes. J Athl Train. 2000;35:155–160.
19. Macciocchi SN. Practice makes perfect: the retest effects in college athletes. J Clin Psychol. 1990;46:628–631.
20. Valovich TC, Perrin DH, Gansneder BM. Repeat administration elicits a practice effect with the Balance Error Scoring System but not with the Standardized Assessment of Concussion in high school athletes. J Athl Train. 2003;38:51–56.
21. Guskiewicz KM, Perrin DH, Gansneder BM. Effect of mild head injury on postural stability in athletes. J Athl Train. 1996;31:300–306.
22. Nordahl SH, Aasen T, Dyrkorn BM, et al. Static stabilometry and repeated testing in a normal population. Aviat Space Environ Med. 2000;71:889–893.
23. Riemann BL, Guskiewicz KM, Shields EW. Relationship between clinical and forceplate measure of postural stability. J Sport Rehabil. 1999;8: 71 –82.
24. Mancuso JJ, Guskiewicz KM, Onate JA, et al. An investigation of the learning effect for the Balance Error Scoring System and its clinical implications [abstract]. J Athl Train. 2002;37:S10.
25. Randolph C. Implementation of neuropsychological testing models for the high school, collegiate, and professional sport settings. J Athl Train. 2001;36:288–296.
26. McCrea M, Kelly JP, Kluge J, et al. The Standardized Assessment of Concussion: Manual for Administration, Scoring, and Interpretation. 2nd Ed., Waukesha, WI: CNS, Inc., 2000.
27. Lomax RG. Statistical Concepts. London: Lawrence Erlbaum Associates; 1992.
28. Atwater SW, Crowe TK, Deitz JC, et al. Interrater and test-retest reliability of two pediatric balance tests. Phys Ther. 1991;70:79–87.
29. Westcott SL, Crowe TK, Deitz JC. Test-retest reliability of the Pediatric Clinical Test of Sensory Interaction for Balance. Phys Occup Ther Pediatr. 1994;14:1–22.
30. Westcott SL, Lowes LP, Richardson PK. Evaluation of postural stability in children: current theories and assessment tools. Phys Ther. 1997;77: 629–645.
31. Holliday PJ, Fernie GR. Changes in the measurement of postural sway resulting from repeated testing. Agressologie. 1979;20:225–228.
32. Daniel J, Olesniewicz M, Reeves D, et al. Repeated measures of cognitive processing efficiency in adolescent athletes: implications for monitoring recovery from concussion. Neuropsychiat Neuropsychol Behav Neurol. 1999;12:167–169.