Repeat administration elicits a practice effect with the balance error scoring system but not with the standardized assessment of concussion in high school athletes

By: Tamara C Valovich, David H Perrin, Bruce M Gansneder.


Made available courtesy of NATIONAL ATHLETIC TRAINERS ASSOCIATION: [http://www.nata.org/jat/](http://www.nata.org/jat/)

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

**Abstract:**

**Objective:** To assess whether repeated administration of the Standardized Assessment of Concussion (SAC) and Balance Error Scoring System (BESS) demonstrates a practice effect in high school athletes.

**Design and Setting:** Subjects were assigned to either a control or practice group. Subjects in the control group were assessed twice, 30 days apart. Subjects in the practice group were assessed on 5 occasions. Four assessments were performed within a 7-day period and 1 more assessment 30 days after the initial test.

**Subjects:** Thirty-two uninjured high school athletes participated in this investigation. Sixteen were randomly assigned to a control group and 16 to a practice group.

**Measurements:** We measured performance on the SAC and on 6 test conditions of the BESS.

**Results:** We found a significant time-by-group interaction on BESS performance on the day-30 test session. A significant practice effect of the BESS was found during the course of repeated administrations in the practice group. After repeated testing, the number of BESS errors decreased with each test session, and error scores on day 5 (10.94 + or - 2.17) and day 7 (9.44 + or - 3.32) were significantly lower than the baseline score (12.88 + or - 3.34). We did not find group differences or a practice effect on the SAC (baseline score = 26.16).

**Conclusions:** Our results revealed no practice effect with the SAC and a slight practice effect with repeated administrations of the BESS, especially with the single-leg stance on foam. Clinicians must acknowledge the potential for practice effects when readministering these concussion assessments to track recovery of an athlete or as a guide in return-to-play decision making.

**Key Words:** postural stability, serial assessments, cognitive assessment

**Article:**
Sport-related mild head injury (MHI) is becoming more frequent in high school athletes. The incidence of injury to the head, neck, and spine has been as high as 13.3% of reported injuries in high school football, whereas the rates for the same in other sports such as baseball and wrestling range from 1.9% to 9.5%, respectively.1 The rate of sport-related traumatic brain injuries reported by the Centers for Disease Control and Prevention was 7.5/100000 for 5- to 14-year-old boys and 10/100000 for 15- to 24-year-old male subjects.2 These age groups represented the highest number of sport-related brain injuries among males, and this is of concern to those participating in high school athletics. In a prospective study of MHI in high school and collegiate football players, the greatest incidence was found at the high school level (5.6%), which was significantly higher than that reported by National Collegiate Athletic Association Divisions I and II but not by Division III.3

Of additional concern to athletic trainers is that football players who sustain an MHI are 3 times more likely to sustain another during the same season.3 Therefore, adequate assessment of MHI is important for both monitoring the athlete and aiding in return-to-play decision making. In recent years, both cognitive testing4-14 and postural-stability13-16 assessments have become widely used in the evaluation of sport-related MHI.

In an attempt to make these assessments more accessible to clinicians, field tests for both cognitive and postural-stability assessments have recently been developed. The Standardized Assessment of Concussion (SAC) is a mental-status test that takes 5 to 7 minutes to administer. It tests the domains commonly affected by MHI, orientation, immediate and delayed memory, and concentration, while allowing for neurologic examination and clinical evaluation of symptoms with exertion, coordination, strength, and sensation as part of the assessment. The SAC is both valid and reliable in the evaluation of early mental-status deficits after MHI; however, its test-retest reliability is relatively low.17 McCrea et al18 found that immediately after MHI in a group of high school and college football athletes, scores on each SAC domain and the total score were significantly lower in the athletes with concussions than in the uninjured controls. More recently, investigators17,19,20 have also found decreased mental status immediately after MHI on the SAC.

<table>
<thead>
<tr>
<th>Subject Characteristics</th>
<th>Control</th>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (men/women)</td>
<td>16 (16/0)</td>
<td>16 (14/2)</td>
</tr>
<tr>
<td>Age (y)</td>
<td>16.63 ± 1.41</td>
<td>17.25 ± 1.69</td>
</tr>
<tr>
<td>Grade</td>
<td>10.94 ± 1.34</td>
<td>11.06 ± 1.06</td>
</tr>
<tr>
<td>Mass (kg)</td>
<td>67.39 ± 6.55</td>
<td>77.56 ± 12.20*</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>175.9 ± 8.16</td>
<td>178.28 ± 9.47</td>
</tr>
</tbody>
</table>

*Significantly different from control group.

The Balance Error Scoring System (BESS) is a clinical test battery that uses modified Romberg stances on different surfaces to assess the postural stability of the athlete after MHI. Previous researchers using the BESS to evaluate postural stability after MHI in athletes found decreased
stability on the foam surfaces through day 3 after injury, indicating that the BESS is sensitive to
the postural-stability alterations after MHI.16

In most cases, however, athletic trainers readminister these assessments on several occasions
after MHI to assess the recovery of the athlete and to aid in return-to-play decision
making.6,13,14,16,18,21,22 One caveat of repeated performance of concussion assessments is
the inability of the clinician to "tease out" performance improvements from true neurologic
recovery versus learning or practice effects. Previous investigations with neuropsychological test
batteries have revealed practice effects or improvement in performance as a function of having
previously taken the test or performed the task.6,21-26 The neuropsychological tests are not the
only assessments potentially confounded by practice effects; practice effects have also been
found in other assessments of postural stability.13,27-32

It is unclear whether practice or learning effects occur with repeated administration of the SAC
or BESS; yet, it is imperative to understand whether improvement on the SAC and BESS is the
result of the recovery of the athlete or familiarization with the assessments. Therefore, the
purpose of this study was to investigate the effects of repeated administration of the SAC and
BESS in high school athletes. We hypothesized that performance on the SAC and BESS would
improve in a practice group that was exposed to the test on repeated occasions and that the scores
of the assessments in both groups on day 30 would not differ from baseline.

METHODS

Subjects
Thirty-two uninjured high school student-athletes (age = 16.94 + or - 1.56 years, grade = 11 + or
- 1.19, body mass = 72.47 + or - 10.93 kg, height = 177.09 + or - 8.78 cm) volunteered to
participate in this study (Table). Subjects were randomly assigned to either the control (n = 16)
or practice (n = 16) group. All subjects participated in either recreational or interscholastic
athletics on the soccer, lacrosse, track, cheerleading, tennis, basketball, or baseball teams.
Subjects were excluded if they had sustained an MHI or a lower extremity injury in the last 6
months or suffered from any visual, vestibular, or balance disorders. All subjects read and signed
an informed consent form approved by the university's human investigations committee, which
also approved the study. The parents of all minors also read and signed an informed consent
form before their children participated.

Procedures
All subjects were administered the SAC and BESS at an initial test session, which served as each
participant's baseline score. All subjects were tested again on day 30. Participants assigned to the
practice group also received repeated administrations on days 3, 5, and 7 after baseline.
The SAC was administered to each participant as previously described by McCrea et al.18,33,34
Scores for each domain of the SAC were recorded along with the subjects' total scores of a
possible 30. Forms A, B, and C of the SAC were counterbalanced among subjects and among
repeated administrations. Previous investigators33 have shown the 3 forms of the SAC to be
equivalent.

The BESS consists of six 20-second balance assessments on both firm and foam surfaces.
Subjects performed a double-leg, a single-leg, and a tandem (heel-to-toe) stance on the floor and
again on a 46- x 43- x 13-cm$^3$ block of medium-density foam. Subjects were instructed to stand in the test position with the hands on the iliac crests, head up, and eyes closed. Foot placement was changed according to the particular balance test. Subjects were told to remain in the test position until the investigator asked them to relax. Errors were counted by the investigator if the subject (1) opened the eyes, (2) stepped, stumbled, or fell out of the test position, (3) removed the hands from the hips, (4) moved the hip to more than 30[degrees] of flexion or abduction, (5) lifted the toes or heels from the test surface, or (6) remained out of the test position for longer than 5 seconds. The total number of errors for each of the 6 tests was calculated and summed for the subject's total BESS score. The subjects performed the stances in the following order: double leg, single leg, and tandem. The firm and foam surfaces were counterbalanced across subjects and test days.

**Data Analysis**

To determine if group differences existed between the baseline and day-30 scores, 2 separate, repeated-measures analyses of variance (ANOVA) were calculated for the SAC and BESS. Using a mixed-model, repeated-measures ANOVA with 1 within-group factor (day) and 1 between-group factor (group), we determined the significance for total SAC score. A second mixed-model, repeated-measures ANOVA with 1 within-group factor (day) and 1 between-group factor (group) was used for the total BESS score. Significant differences were examined further with the Tukey Honestly Significant Difference post hoc analysis. Level of significance (P < .05) was set a priori. Data were analyzed using SPSS software (version 10.0, SPSS Inc, Chicago, IL).

To determine if a practice effect was present across the first 4 test days in the practice group, separate, repeated-measures ANOVA with 1 within-group factor (day) were calculated for total score and for each of the 4 domains of the SAC. To correct for the multiple analyses, a Bonferroni adjustment was made, and significance was set at P = .01. Significant differences were further analyzed with the Dunnett post hoc analysis to compare subsequent measures with baseline. A separate, repeated-measures ANOVA with 3 within-group factors (day, surface, and stance) was used for the BESS. Significance level was set a priori (P < .05). Significant differences were examined further with Dunnett post hoc analysis because all measures were compared with their baseline measures.

**RESULTS**

**Standardized Assessment of Concussion**

We found no significant differences between the control and practice groups on performance of the SAC between baseline and day 30 (time) ($F^{sub} 1,30^{=}= 2.217$, $P = .147$, $[beta] = .302$), group ($F^{sub} 1,30^{=}= 0.342$, $P = .563$, $[beta] = .087$), or time-by-group interaction ($F^{sub} 1,30^{=}= 1.381$, $P = .249$, $[beta] = .207$). Mean SAC scores for the control group were 26.13 + or - 1.93 and 27.19 + or - 2.26 and for the practice group were 26.19 + or - 2.61 and 26.31 + or - 2.21 at baseline and on day 30, respectively.

The repeated-measures ANOVA for total SAC score within the practice group revealed a significant main effect for test days (time) ($F^{sub} 3,45^{=}= 4.934$, $P = .005$, $[beta] = .886$). The Dunnett post hoc analysis showed that days 3, 5, and 7 were not significantly different from baseline. Mean SAC scores across the 4 test sessions were 26.19 + or - 2.61, 24.81 + or - 1.60,
25.38 + or - 2.13, and 27.19 + or - 1.97, respectively. Results for each domain of SAC (Figure 1) demonstrated significant differences for the orientation ($F^\text{sub} 3,45^ = 7.069, P = .001, \beta = .971$) and delayed-recall ($F^\text{sub} 3,45^ = 6.883, P = .001, \beta = .967$) domains; scores on day 3 were significantly lower than the baseline score. We found no differences for immediate memory ($F^\text{sub} 3,45^ = 0.707, P = .553, \beta = .188$) or concentration ($F^\text{sub} 3,45^ = 0.952, P = .424, \beta = .243$) across the test sessions.

**Balance Error Scoring System**

For total BESS errors, we did not find a significant main effect for pretest to day 30 (time) ($F^\text{sub} 1,30^ = 0.078, P = .782, \beta = .058$) or group ($F^\text{sub} 1,30^ = 1.148, P = .292, \beta = .179$), but we did find a significant time-by-group interaction ($F^\text{sub} 1,30^ = 5.770, P = .023, \beta = .642$). Post hoc analysis revealed that the practice group scored fewer errors than did the control group on the day-30 test; however, neither group's scores were different from its baseline score (Figure 2). Mean BESS scores were 12.88 + or - 3.34 and 14.25 + or - 3.15 for the control group and 12.88 + or - 3.34 and 11.69 + or - 3.63 for the practice group at baseline and on day 30, respectively.

Using the repeated-measures ANOVA within the practice group for the 4 test days, we found significant main effects for test day (time) ($F^\text{sub} 3,45^ = 7.134, P = .001, \beta = .972$), surface ($F^\text{sub} 1,45^ = 63.210, P < .0001, \beta = 1.0$), and stance ($F^\text{sub} 2,30^ = 237.114, P < .0001, \beta = 1.0$). We noted significant 2-way interactions for time by surface ($F^\text{sub} 3,45^ = 4.464, P = .008, \beta = .8949$), time by stance ($F^\text{sub} 6,90^ = 2.888, P = .013, \beta = .873$), and surface by stance ($F^\text{sub} 6,90^ = 12.325, P < .0001, \beta = .992$). Post hoc analysis for time showed that day-5 and day-7 total errors were significantly lower than baseline (Figure 3). BESS errors across the 4 test sessions were 12.88 + or - 3.34, 12.38 + or - 1.89, 10.94 + or - 1.71, and 9.44 + or - 3.22, respectively. Post hoc analysis for surface demonstrated differences between the firm and foam surfaces, and for stance, it showed significant differences among 3 stances. Post hoc analyses for the time-by-surface interaction revealed that the day-7 score on the foam surface was significantly less than baseline (Figure 4). The time-by-stance interaction showed significantly fewer errors on the day-7 single-leg stance when compared with baseline (Figure 5).
Figure 1. Standard Assessment of Concussion scores for each domain across the repeated testing days in the practice group. *Orientation scores significantly less than baseline. †Delayed-recall scores significantly less than baseline.
Figure 2. Total Balance Error Scoring System score between groups on the baseline and day-30 tests. *Significantly different between groups on day 30.

Figure 3. Total Balance Error Scoring System scores across the repeated testing days in the practice group. *Significantly less than baseline.
Figure 4. Time-by-surface interaction across the repeated test sessions in the practice group. *Significantly different from baseline.
DISCUSSION

**Standardized Assessment of Concussion**

Our main finding with regard to the SAC was that there was no practice effect with repeated administration of this assessment tool. Our results differ from the results of McCrea, who noted slight improvement in normal controls from baseline to 48 hours after baseline (with 1 test in between) on the SAC total score and the immediate memory domain. Our results are also contrary to the practice effects found upon repeated administration of more complex cognitive test batteries in athletes. Of the literature pertaining to the use of cognitive assessments in athletes, the lack of a practice effect was seen only by Oliaro et al on the Hopkins Verbal Learning Test, when subjects' performance actually decreased from the first to the third test day. The authors attributed this to a possible lack of motivation on the part of the subjects. One difference between the Hopkins Test and other neuropsychological assessments is the use of alternative forms, which attempt to account for a practice effect. The test we used, the SAC, also has 3 forms that have been previously shown to be equivalent. These 3 forms are used in assessments to attempt to alleviate problems associated with practice effects and may be an explanation for the lack of learning in our high school athletes.

Clinicians should note that after MHI and with repeated follow-up administrations of the SAC, scores should return to the athlete's baseline or normative values, and if they do not, one should assume that the athlete is not fully recovered. In addition, our findings may be helpful to

---

**Figure 5. Time-by-stance interaction across the 5 test sessions in the practice group.** Single-leg day-7 score significantly less than baseline.
researchers in that the control group within MHI investigations is expected to demonstrate no practice effect with repeated administration of the SAC.

We also found no differences between the groups at the baseline and day-30 test sessions. Thus, repeat administration of the SAC in high school athletes does not affect long-term retention of the assessment or reveal developmental improvements in cognitive efficiency. After MHI, athletic trainers can comfortably readminister the SAC at day 30 to obtain new baseline values without having to worry about practice effects this long after the repeat administration. Although we used a day-30 assessment point and showed no differences compared with baseline, clinicians and researchers using a longer follow-up interval in adolescent athletes need to be aware of the potential for improvements in cognitive function as a result of brain maturation.35 In general, clinicians should perform new baseline assessments on athletes after an MHI and, in the case of younger athletes, at the start of a new academic year.

**Balance Error Scoring System**

After repeated administration of the BESS, the practice group displayed significantly fewer errors on days 5 and 7 compared with baseline. These results differ from those of Riemann et al,32 who did not find a practice effect in collegiate athletes using the BESS over 3 test sessions. However, the sample size and number of test sessions in their investigation were small. Our findings are more consistent with other investigations on repeated testing using various measures of postural stability.28-31 The total number of errors decreased during each test session and was significantly different from baseline by day 5. During this time, subjects were tested with a very short interval between test sessions (1 day). This short time interval is likely a factor in the performance improvements noted in our investigation. In a study investigating the influences of time intervals between repeated test sessions using stabilometric testing, Nordahl et al31 found that subjects who were tested with the shortest time interval between test sessions (11 days) had significantly greater learning on various postural sway measures than did groups that had test-retest time intervals of 17, 31, and 115 days.

Although we noted a significant practice effect using the total number of BESS errors, our results differ slightly when we look at the separate conditions performed by the subjects. Not every condition resulted in the same extent of improvement. Our significant time-by-surface interaction showed significant differences on the foam surface, when the score on day 7 was significantly less than baseline. Our time-by-stance interaction also demonstrated that the day-7 test on the single-leg stance was significantly less than the baseline score on that stance. The scores for the double-leg and tandem stances remained relatively constant throughout the test sessions. The foam-surface and the single-leg stance conditions represented the surface and stance on which the subjects performed the poorest and on which the greatest improvement was found.

Given the models set forth by other measures of postural stability and learning, it is typical that improved performance takes place in activities that are novel to or unusual for the subject. In an investigation on repeated stabilometric testing, Nordahl et al31 found that the maximum learning occurred on a foam surface with the eyes closed and with the shortest interval between test sessions. Nordahl et al31 noted no learning effects for any stance in which the eyes were open but noted high learning potential with loss of visual reference or changed proprioceptive input.
In another investigation comparing weekly versus daily repeated testing of postural stability, Hamman et al.28 showed that both methods of practice were enough to elicit decreases in several dynamic variables, including transition time, path error, and peripheral sway area. However, neither training nor practice elicited any practice effects on static tests, including eyes open or eyes closed, or stances involving feedback. Hamman et al.28 did note a significant difference between the daily and weekly groups for the eyes-closed static test. Subjects accustomed to daily testing sessions performed better with less sway than did those in the group tested weekly. In an earlier investigation, Hamman et al.29 found that practice involving dynamic postural sway did not affect eyes-open or eyes-closed static-balance measures.

The finding that a greater practice effect takes place during activities that are unusual or that require different senses or movements is supported by the work of Hertel et al.30 who investigated practice effects with a dynamic measure of balance, the Star Excursion Balance Test (SEBT). Athletes performed the SEBT on 2 days, with a time interval of 1 week. A significant practice effect was seen in 3 of the 8 directions on the left leg and 4 of the 8 directions on the right. Scores on the SEBT did not plateau until trials 7 to 9 (day 2), indicating that performance increased during the first 6 trials (day 1). Of interest is that the lateral, posterior-medial, posterior-lateral, and posterior directions of the SEBT were prone to the learning effect. These directions do not allow for easy view of the target; therefore, the practice effects may be related to coping with impaired visual feedback. As with our results on the foam surface and with the single-leg stance, the difficulty of the task seems to be a large factor in the extent of improvement that takes place with repeated administrations.

Our main finding between the groups on BESS performance was that they differed on the day-30 test session. This difference is likely due to slight increases in the control group's errors and slight decreases in the practice group's errors. The number of errors scored by either group was not significantly different from its baseline score. This is important in demonstrating that the practice group did not retain the skills needed to perform BESS when subjects were not exposed to the test for 3 weeks. The increased errors in the control group could be attributed to a lack of motivation on the follow-up test. In a clinical situation, after a time period of approximately 3 weeks during which an athlete is not exposed to the BESS, error scores should return toward and not be significantly different from baseline values.

**CLINICAL RELEVANCE**

The BESS is an effective tool for the sideline evaluation of postural stability after MHI.16 However, one should be cautious in that performing this assessment on an athlete on several occasions to monitor recovery can elicit a practice effect, especially with novel tasks (ie, single-leg stance, foam surface). Although at this time we are unable to distinguish between improved performance on the BESS due to recovery or practice in an injured population, athletic trainers should expect their athletes to improve on subsequent test administrations. If the athlete's performance does not improve, then one might suspect a lingering abnormality, which should be followed up with an in-depth clinical examination. Our results also indicate that practice effects are not retained after a 30-day period, and at that time, it would be acceptable to perform a new baseline assessment on those who were exposed to repeated administrations (ie, athletes after MHI).
Athletic trainers should be confident that repeated administration of the SAC does not elicit practice effects in healthy athletes and that it can be used to assess mental-status alterations immediately after a concussion. Additionally, previous research with this tool has demonstrated its sensitivity in finding deficits in persons suffering from an MHI. We recommend that athletic trainers administer the different forms of the SAC when performing repeated test sessions. In addition, clinicians should understand the difference between mental-status testing in the acute stages of concussion and comprehensive neuropsychological testing for lingering postconcussion symptoms.

Although the SAC and BESS are becoming more frequently used as ways of objectively grading the cognitive and postural-stability function of athletes after MHI, neither should be used as the sole basis for monitoring the recovery of athletes and aiding in return-to-play decision making. We recommend that they be used in conjunction with the clinical examination and a symptom checklist and that no athlete be allowed to return to competition until asymptomatic.

ACKNOWLEDGMENTS
We thank Kevin M. Guskiewicz, PhD, ATC, for his assistance with the interpretation of the data and Brent L. Arnold, PhD, ATC, and Sandra J. Shultz, PhD, ATC, CSCS, for providing feedback on the manuscript.

REFERENCES
Administration, Scoring, and Interpretation. 2nd ed. Waukesha, WI: CNS Inc; 2000.