

[Athletic Training]



The Balance Error Scoring System Learned Response Among Young Adults

Ivan J. Mulligan, PT, DSc,[†] Mark A. Boland, PT, MPT,[†] and Carol V. McIlhenny, MHS, BSN, RN^{*†}

Background: Concussion management practices are important for athlete safety. Baseline testing provides a benchmark to which post-injury assessments are compared. Yet few neurophysical concussion assessment studies have examined learned response. The Balance Error Scoring System (BESS) measures postural stability through 6 conditions by counting the errors committed during each condition. In a study examining the performance of high school-aged athletes on the BESS, the learned response extinguished in 3 weeks. However, this phenomenon has not been studied in the college-aged population.

Hypothesis: College-aged adults performing the BESS will have a learned response at 1 and 2 weeks but would have no change from baseline at or after 3 weeks, as found previously in high school-aged subjects.

Study Design: Randomized controlled clinical trial.

Methods: Three groups of college-aged adults ages 18 to 26 years were tested using the BESS at scheduled intervals. Each subject was randomly assigned into 1 of 3 groups to determine learned response at weeks 1, 2, and 4. Changes in pretest and posttest BESS scores were compared using the paired *t* test for each group at week 4 and other intervals. Differences among groups were compared using analysis of variance for means or the chi-square test for proportions.

Results: After 4 weeks, participants exhibited a mean (95% confidence interval) change from pretest baseline of -2.30 ($-4.75, 0.16$) in the control group ($P = 0.065$), -3.13 ($-4.84, -1.41$) in Group 1 ($P = 0.001$), and -2.57 ($-5.28, 0.15$) in Group 2 ($P = 0.063$). There were no statistically significant differences between the 3 groups for week 4 BESS score ($P = 0.291$) or changes from baseline to week 4 BESS scores ($P = 0.868$). Overall, participant score changes from baseline to the 4-week follow-up still showed a statistically significant or close to significant reduction across the 3 groups, indicating the learned response did not extinguish after 4 weeks.

Conclusion: Repeated BESS testing results in a learned effect in college-aged adults did not extinguish after 4 weeks. These results question the ability of the BESS to assess an athlete's balance deficits following a concussion.

Clinical Relevance: Given learned response did not extinguish in this sample and the BESS has a minimal detectable change/reliable change index of 7 or greater, the effectiveness of the BESS to assess balance may be limited.

Keywords: concussion; learned behavior; practice effect; traumatic brain injury

Concussion or mild traumatic brain injury (MTBI) is a pathophysiologic process affecting the brain induced by direct or indirect biomechanical forces.¹³ Currently, the estimated number of concussions in the United States ranges between 1.6 and 3.8 million cases per year.^{3,22} Many concussions are undiagnosed, underdiagnosed, or not reported and remain untreated. There is no "reference standard" diagnostic tool for concussion.^{2,3} Imaging tools (computed tomography

scan and magnetic resonance imaging) lack sensitivity to diagnose concussion as well.¹⁷ Without the proper diagnosis and treatment of concussion, a patient may suffer from more severe symptoms and can even sustain further brain injury if subsequent injury follows the initial incident.^{4-6,14-16,20,23,24}

A wide variety of symptoms exists for patients with concussion. Subjective symptoms include headache, dizziness, nausea or vomiting, vision changes, sensitivity to light or noise,

From [†]Saint Francis University, Loretto, Pennsylvania

*Address correspondence to Carol V. McIlhenny, MHS, BSN, RN, Former Telehealth Research Specialist, Center of Excellence for Remote and Medically Under-Served Areas (CERMUSA), Saint Francis University, 445 Dickinson Street, Johnstown, PA 15905 (e-mail: Swatie2@msn.com).

The following authors declared a potential conflict of interest: Ivan J. Mulligan, Mark A. Boland and Carol V. McIlhenny's institution received a grant from the US Army Medical Research Command Telemedicine and Advanced Technology Research Center.

This project is supported by Saint Francis University's Center of Excellence for Remote and Medically Under-Served Areas (CERMUSA) in Loretto, Pennsylvania, and funded by the US Army Medical Research and Materiel Command (USAMRMC) Telemedicine and Advanced Technology Research Center (TATRC), Fort Detrick, Maryland—Contract Number W81XWH-10-2-0107.

This research study was approved by the Saint Francis University Institutional Review Board as well as the IRB at the Telemedicine and Advanced Technology Research Center (TATRC).

DOI: 10.1177/1941738112467755

Table 1. Schedule of BESS testing by group.

Test Group	Day 0	Day 7	Day 14	Day 28
Control	X			X
Group 1	X	X		X
Group 2	X		X	X

BESS, Balance Error Scoring System.

difficulty concentrating, tinnitus, drowsiness, feeling “foggy,” sadness, and hallucinations.^{8,21} Individuals can also exhibit objective signs: loss of consciousness, amnesia, disorientation, appearing dazed or acting confused, forgetfulness, poor coordination, inappropriate displays of emotion, personality change, seizure, and slowed speech.^{8,22}

Another objective sign exhibited by individuals with concussion is decreased postural control.^{11,23} Postural control markedly decreases in concussion when the eyes are closed, eliminating visual referencing.²⁶ Athletes sustaining a concussion exhibit sensory interaction problems with an inability to process altered sensory information.¹¹ This results in decreased postural stability, which can last up to 3 days after the injury.^{3,11,26}

To assess a change in postural control, an individual baseline reference score is required. A memorandum, issued in 2010 by the National Collegiate Athletic Association (NCAA), indicates that baseline testing should be performed on athletes in most sports prior to preseason training to determine individual baseline scores. Minimum baseline assessments should include a balance assessment.⁷ The post-injury clinical evaluation should also include the same assessments. Comparison of post-injury results against the baseline scores, along with a physical assessment, is required for further treatment or return-to-play decisions.³

The National Athletic Trainers Association and the NCAA both recommend that the Balance Error Scoring System (BESS) be used when performing an on-field assessment following a suspected concussion.^{7,12} The Defense and Veterans Brain Injury Center Working Group has also recommended using the BESS for soldiers.¹⁴ The BESS is a static postural stability screen based on the Romberg test and other balance theories.^{24,25} The sensitivity of the BESS for concussion, however, is 0.34 immediately following the injury. Sensitivity further decreases to 0.16 to 0.24 at 1 to 3 days, while specificity remains ≥ 0.91 through 7 days.¹⁸

When comparing post-injury BESS scores, the difference must be weighed against the known reliable change index (RCI) or minimum detectable change (MDC) values. Finnoff et al⁹ reported that the interrater and intrarater MDC in a normal population were 9.4 and 7.3, respectively, with a 95% confidence interval (CI). Valovich McLeod et al²⁹ established an RCI of +3 or -7 with a 70% confidence interval. That study,

however, was performed on children between the ages of 9 and 14 years.

Adult participants can learn to improve their performance on balance equipment with repeated trials.³¹ Balance scores (BESS) return to baseline 3 weeks after the last assessment in high school athletes.³⁰ This study did not cite MDC or RCI parameters.

If the athlete retains a practice effect, it could render the balance assessment inaccurate, thus potentially allowing the athlete to return to play before recovery is complete. Since there is a paucity of data determining the length of learned response in collegiate athletes, the authors questioned whether the length of time required for scores to return to baseline would be similar to that for high school athletes. The authors hypothesized that the college-aged adult participants performing the BESS would have a learned response at 1 to 2 weeks but would have no change from baseline at 3 weeks from the last assessment.³⁰

METHODS

A convenience sample of 84 participants was recruited by word-of-mouth for this study. The target population was young adults, aged 18 to 26 years old. Exclusion criteria included pregnancy and history of head injury within the past 6 months since these factors had the potential to introduce confounding effects. Corresponding Institutional Review Boards approved the study. Prior to participation, each subject read and signed an informed consent form.

The participants were randomly assigned to 1 of 3 groups in a consecutive, fixed order allocation algorithm with a block size of 3 (Table 1). Each participant underwent BESS testing according to their group assignment, allowing the research to be conducted at various intervals to determine the length of learned response in the collegiate athlete population.^{10,26} Participants assigned to control group underwent testing at baseline and week 4 only. Group 1 and 2 participants underwent testing at week 1 and 2, respectively, in addition to baseline and week 4 testing.

Before each trial, the participant completed a questionnaire, which included demographics, history of head injury, disease, inner ear or balance problems, cold or flu symptoms, alcohol or drug use, medication usage, and sleep habits—variables that potentially could affect balance scores.^{10,27,28,33} The scorers were all familiar with BESS scoring and did not require any review.

Table 2. Participant demographic data.

	Control (n = 27)	Group 1 (n = 24)	Group 2 (n = 23)
Sex	F = 22; M = 5	F = 16; M = 8	F = 17; M = 6
Mean age, y (SD)	21.6 (1.52)	20.9 (1.41)	21.6 (1.62)
Mean height, cm (SD)	167.8 (7.64)	171.2 (10.69)	169.8 (8.94)
Mean weight, kg (SD)	67.7 (13.42)	72.4 (18.29)	68.6 (13.75)
Mean BMI (SD)	23.9 (3.90)	24.4 (4.07)	23.7 (3.90)
Mean hours slept previous night (SD)	6.8 (0.62)	6.8 (0.78)	6.5 (0.87)
Baseline BESS score ($P = 0.432$)	10.23 (6.71)	11.5 (6.19)	12.6 (6.18)

BESS, Balance Error Scoring System; BMI, Body Mass Index; SD, standard deviation; F, female; M, male.

The BESS assesses balance during 6 separate 20-second trials. These trials include 3 different stances (double leg, single leg, and tandem) on each of 2 different surfaces (firm and foam). The foam creates an unstable surface. Subjects are asked to stand as still as possible with hands on iliac crests and eyes closed. Testing begins when the subject assumes the proper stance.

The BESS is scored by counting the errors the individual commits during the testing conditions.³² Errors include opening the eyes, removing hands from hips, lifting the heel or forefoot from the testing surface, moving hips more than 30° of flexion or abduction, stepping, stumbling or falling, or staying out of the testing position for more than 5 seconds. Multiple errors occurring at once score 1 point. A lower score indicates fewer errors were committed. "The BESS has established good test-retest reliability and good concurrent validity when compared with laboratory force plate measures."¹²

Data analysis was also performed via SAS version 9.2 (SAS Institute, Cary, North Carolina) and Minitab 16 (Minitab Statistical Software, State College, Pennsylvania). To examine participant comparability among the 3 groups, the chi-square test compared differences for categorical variables such as sex. Analysis of variance (ANOVA) compared differences among groups for quantitative demographic and baseline variables. Since the primary objective was to determine if the learned response remained at week 4, a paired *t* test compared changes in pretest and posttest BESS score for each group. Paired *t* test was applied at each follow-up testing. ANOVA was also applied to test differences in BESS scores at week 4 and changes from baseline to week 4 among the 3 groups. Analysis of covariance was used to compare changes in BESS score from baseline to week 4 adjusting for baseline BESS score. Proportions of subjects with changes reaching clinically significant levels (determined by MDC) were compared among groups using the chi-square test. A *P* value of < 0.05 was used to determine the statistical significance in 2-sided testing.

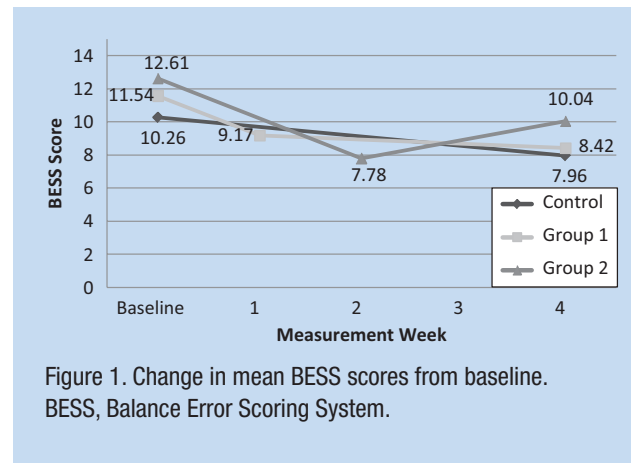


Figure 1. Change in mean BESS scores from baseline. BESS, Balance Error Scoring System.

There was no prespecified multiple testing adjustment for the pairwise comparisons among groups or over time.

RESULTS

A total of 74 participants (88.1%) completed follow-up testing (Table 2). There was no statistically significant difference among the 3 groups in baseline BESS scores (Table 3).

Reductions in BESS score were statistically significant at week 1 and week 2, indicating a learned response (Table 3 and Figure 1). While Group 2 showed an increase in scores between 2 and 4 weeks, the week 4 score remained lower than baseline ($P = 0.063$). Using the paired *t* test for each group, the analysis identified statistically significant changes in BESS scores from baseline to week 4 in participants of Group 1. The reductions from baseline to week 4, in the control group and Group 2 approached statistical significance also, even with a relatively small sample size. There were no statistically significant differences between 3 groups for week 4 BESS score ($P = 0.291$) or changes from baseline to week 4 BESS scores

Table 3. Change in BESS Scores.

Time Interval	Mean Change	95% Confidence Interval
0 to 4 weeks – Control	–2.30	–4.75 to –0.16*
0 to 4 weeks – Group 1	–3.12	–4.84 to –1.41**
0 to 4 weeks – Group 2	–2.57	–5.28 to –0.15*
0 to 1 week – Group 1	–2.38	–4.12 to –0.06**
0 to 2 weeks – Group 2	–4.83	–7.32 to –2.33**
1 to 4 weeks – Group 1	–0.75	–2.30 to 0.80
2 to 4 weeks – Group 2	2.26	0.20 to –4.32

BESS, Balance Error Scoring System.

* $P \leq 0.05$.

** $P \leq 0.01$.

Table 4. Change from baseline in score.

Test Group	Change in BESS Score, %		
	< 0	≤ 7	≤ 9
Group 1 at 4 weeks	66.67	20.83	8.33
Group 2 at 4 weeks	78.26	26.09	13.04
Control group at 4 weeks	44.44	22.22	18.52
Group 1 at 1 week	66.67	20.83	8.33
Group 2 at 2 weeks	78.26	30.43	13.04

BESS, Balance Error Scoring System.

($P = 0.868$). No statistical difference was found among the 3 groups at week 4 and changes from baseline to week 4 BESS scores with adjustment for baseline score through analysis of covariance. Overall, participant score changes from baseline to 4-week follow-up still showed statistically significant or close to significant reduction across the 3 groups, indicating the learned response persisted.

The number and percentage of participants who achieved a reduction in scores of 7 or greater at each time interval was analyzed to determine the proportion of participants that had score changes greater than the MDC or RCI, representing a clinical difference in assessment. There was a similar proportion of participants from each group who demonstrated changes in BESS scores of clinical significance based on MDC of 7 or greater (Table 4). Similar proportions were observed in reduction from baseline at week 1 in Group 1 (20.8%) and at week 2 in Group 2 (30.4%).

DISCUSSION

The learned response, evidenced by the statistically significant reduction in follow-up scores, did not extinguish after the 4-week study period in college-aged adults. Although the results in this study represent a statistically significant change, it does not represent a clinically significant change, defined as greater than the MDC or RCI. Because of the +3- or RCI at 70% CI or ± 7.3 - to 9.4-point MDC at 95% CI, statistically significant changes from baseline scores may not make the test clinically useful.^{9,29} Because of these large ranges, on a 60-point assessment, the BESS test may not necessarily be sensitive enough to pick up subtle changes in postural control. MDC values of the BESS may need to be reevaluated. Re-examining MDC at each time interval following “practice” may be needed to determine whether the MDC changes with participant experience.

The specificity of the BESS at different intervals has ranged from 0.91 to 0.96.¹⁸ Using these data, the positive likelihood ratios range from 1.43 to 7.75, indicating, at best, only a small to moderate shift in probability. Negative likelihood ratios range from 0.72 to 0.97, indicating the test is not significant in altering probability.

When examining the 3 groups in this study, 22.2% of the control group, 20.8% of Group 1, and 26.1% of Group 2 demonstrated a learned response that exceeded MDC and did not return to baseline by the end of the 4-week testing period (Table 4). These findings suggest that the BESS has a potential for a high false negative rate and may have limited usefulness as a screening tool.

If the athlete sustains a concussion before the learned response has extinguished, they may demonstrate a score that is within the MDC of baseline during post-injury testing. Recently, vestibular limitations were noted in athletes who are recovering from concussion.¹ Two functions of the vestibular system are postural control and visual fixation while moving the head.¹⁹ The BESS is a static test that does not appear to fully assess the visual and dynamic aspects of postural control.

Study limitations include a relatively small sample size (< 30/group) of healthy subjects and did not include a sample size estimate. In addition, the study did not determine the time required to extinguish the learned response.

CONCLUSION

This study identified a learned effect of the BESS that does not extinguish after 4 weeks in college-aged individuals. The BESS score reduction from baseline was still statistically significant at week 4 for Group 1. The results demonstrate a potential limitation when using the BESS assessment in collegiate-aged athletes.

REFERENCES

- Alsalaheen BA, Mucha A, Morris LO, et al. Vestibular rehabilitation for dizziness and balance disorders after concussion. *J Neurol Phys Ther.* 2010;34(2):87-93.
- Broglio SP, Guskiewicz K. Concussion in sports: the sideline assessment. *Sports Health.* 2009(1):361-369.
- Broglio SP, Macciocchi SN, Ferrara MS. Sensitivity of the concussion assessment battery. *Neurosurgery.* 2007;60(6):1050-1057; discussion, 1057-1058.
- Cantu R. Second-impact syndrome. *Clin Sports Med.* 1998;17:37-44.
- Cantu R, ed. *Classification Concussion: Ongoing Controversy.* New York, NY: Springer Science + Business Media Inc; 2006.
- Cantu R, Voy R. Second impact syndrome: a risk in any contact sport. *Phys Sportsmed.* 1995;23:27-34.
- Concussion Management Plan [memorandum]. Indianapolis, IN: National Collegiate Athletic Association; 2010.
- Field M, Collins M, Lovell M, Maroon J. Does age play a role in recovery from sports related concussion? A comparison of high school and collegiate athletes. *J Pediatr.* 2003;142:546-559.
- Finnoff JT, Peterson VJ, Hollman JH, Smith J. Intrarater and interrater reliability of the Balance Error Scoring System (BESS). *PM R.* 2009;1(1):50-54.
- Guskiewicz K. Balance assessment in the management of sport-related concussion. *Clin Sports Med.* 2011;30:89-102.
- Guskiewicz K, Riemann BL, Perrin DH. Alternative approaches to the assessment of mild head injury in athletes. *Med Sci Sports Exerc.* 1997;29:S213-221.
- Guskiewicz KM, Bruce SL, Cantu RC, et al. National Athletic Trainers' Association Position Statement: management of sport-related concussion. *J Athl Train.* 2004;39(3):280-297.
- Herring S, Bergfeld J, Boland A. Concussion (mild traumatic brain injury) and the team physician: a consensus statement. *Med Sci Sports Exerc.* 2006;38:280-297.
- Le Roux PD, Choudhri H, Andrews BT, eds. *Cerebral Concussion and Diffuse Brain Injury.* 4th ed. New York, NY: McGraw Hill Companies Inc; 2000.
- Lew HL, Thomander D, Chew KT, Bleiberg J. Review of sports-related concussion: potential for application in military settings. *J Rehabil Res Dev.* 2007;44(7):963-974.
- Lovell MR, Collins MW, Iverson GL, Johnston KM, Bradley JP. Grade 1 or "ding" concussions in high school athletes. *Am J Sports Med.* 2004;32(1):47-54.
- Mann MD. *The Nervous System in Action.* <http://www.unmc.edu/physiology/Mann/mann9.html>. Accessed August 24, 2011.
- McCrea M, Barr WB, Guskiewicz K, et al. Standard regression-based methods for measuring recovery after sport-related concussion. *J Int Neuropsychol Soc.* 2005;11(1):58-69.
- McCrory P, Meeuwisse W, Johnston K, et al. Consensus Statement on Concussion in Sport 3rd International Conference on Concussion in Sport held in Zurich, November 2008. *Clin J Sport Med.* 2009;19(3):185-200.
- McCrory PR, Berkovic SF. Second impact syndrome. *Neurology.* 1998;50(3):677-683.
- Meehan WP III, Bachur RG. Sport-related concussion. *Pediatrics.* 2009;123(1):114-123.
- Mihalik JP, Ondrak KS, Guskiewicz KM, McMurray RG. The effects of menstrual cycle phase on clinical measures of concussion in healthy college-aged females. *J Sci Med Sport.* 2009;12(3):383-387.
- Peterson CL, Ferrara MS, Mrazik M, Piland S, Elliott R. Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. *Clin J Sport Med.* 2003;13(4):230-237.
- Randolph C, McCrea M, Barr WB. Is neuropsychological testing useful in the management of sport-related concussion? *J Athl Train.* 2005;40(3):139-152.
- Riemann BL, Guskiewicz K, Shields EW. Relationship between clinical and forceplate measure of postural stability. *J Sport Rehabil.* 1999;8:71-82.
- Riemann BL, Guskiewicz KM. Effects of mild head injury on postural stability as measured through clinical balance testing. *J Athl Train.* 2000;35(1):19-25.
- Schmidt PM, Giordani AM, Rossi AG, Coser PL. Balance assessment in alcohol subjects. *Braz J Otorhinolaryngol.* 2010;76(2):148-155.
- Trato J, Johnson EG. Differential diagnosis and management of a patient with peripheral vestibular and central nervous: a case study. *J Man Manip Ther.* 2010;18(3):159-165.
- Valovich McLeod TC, Barr WB, McCrea M, Guskiewicz KM. Psychometric and measurement properties of concussion assessment tools in youth sports. *J Athl Train.* 2006;41(4):399-408.
- 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(1):51-56.
- Weli-Santos L. *Topology of Learning and Correction in Dynamic Balance.* Cambridge, MA: Massachusetts Institute of Technology; 2004.
- Wilkins JC, Valovich McLeod TC, Perrin DH, Gansneder BM. Performance on the Balance Error Scoring System decreases after fatigue. *J Athl Train.* 2004;39(2):156-161.
- Yardley L, Owen N, Nazareth I, Luxon L. Prevalence and presentation of dizziness in a general practice community sample of working age people. *Br J Gen Pract.* 1998;48:1131-1135.