

**Long Term Test Retest Reliability of the BESS Test Using C3 Logix Platform in High School Athletes**

Hango MK, Baker ET, Decoster LC, Valovich McLeod TC, Bay RC, Hollingworth AT: New Hampshire Musculoskeletal Institute, Manchester, NH, and A.T. Still University, Mesa, AZ

**Context:** The estimated risk of concussion in high school athletics is 230 concussions per 100,000 athlete exposures. Current best practices recommend performing an annual baseline examination of motor control using the Balance Error Scoring System (BESS) test; however, in practice, limited resources in secondary schools often extend this interval closer to 2 years. Previously, test retest reliability of the BESS test has been evaluated at shorter intervals ranging from 23 minutes up to 1 year. We sought to evaluate test retest reliability of the BESS test over a 2 year interval in high school athletes using the C3 Logix platform. We hypothesized that the test retest reliability would be strong for ellipse volume, but poor for error count. **Methods:** This retrospective study

used data collected in athletic training facilities and classrooms during routine mass baseline concussion testing. Participants were 390 high school athletes (223 males, 167 females, ages  $14.51 \pm 2.34$  years) from 8 high schools. Each athlete had 2 balance baseline assessments typically at the start of freshman and junior years (630 days  $\pm$  152). All participants had no previous history of concussion prior to the first assessment and no intervening concussion prior to the second assessment. Data was collected using the C3 Logix application on an Apple iPad Air 2. Balance errors were manually counted and recorded by trained assessors and ellipse volume was collected and recorded by the application. With IRB approval, de identified data was exported and analyzed. SPSS was used to calculate 2 way random consistency intraclass coefficients (ICC) for total number of balance errors at each baseline and ellipse volume and errors per each stance. **Results:** There was fair reliability (ICC=0.40-0.59) for ellipse volume in single leg stance on firm surface, tandem stance on firm surface, and double leg stance foam surface (Table). There was poor reliability (ICC<0.40) for total error count, error count per stance, and ellipse volume in the remaining 3

stances. **Conclusions:** This suggests that 2 years between baseline balance assessments may be too long and that ellipse volume is a more reliable measurement than error scores. This may be because it is an objective measurement versus subjective, as suggested by previous researchers. Future research should be performed prospectively to determine long term test retest reliability. Understanding test metrics is imperative to providing appropriate clinical guidance. These results question whether we can rely on older baseline assessments when determining concussion recovery.

**Table: Test-Retest Reliability for Error Count and Ellipse Volume**

Stance	Error Count			Ellipse Volume		
	Baseline 1 (SD)	Baseline 2 (SD)	ICC	Baseline 1 (SD)	Baseline 2 (SD)	ICC
Double Leg-Firm	.020 (.182)	.030 (.174)	.229	-4.92 (1.05)	-4.75 (1.14)	.317
Single Leg-Firm	2.59 (1.95)	2.97 (2.30)	.253	.202 (1.61)	-.171 (1.59)	.435
Tandem Stance-Firm	.770 (1.24)	.890 (1.27)	.272	-1.55 (1.57)	-1.50 (1.67)	.420
Double Leg-Foam	.120 (.404)	.180 (.729)	.101	-1.95 (1.10)	-2.02 (1.05)	.495
Single Leg-Foam	6.56 (2.01)	6.78 (2.16)	.185	2.25 (1.52)	1.84 (1.38)	.377
Tandem Stance-Foam	3.07 (2.33)	3.26 (2.30)	.215	1.35 (1.73)	1.46 (1.66)	.302
All Stances Combined	13.1 (5.19)	14.1 (5.88)	.328			