Responsiveness of the Star Excursion Balance Test on Firm and Unstable Underground

Alain Amacker[1], Jonas Meng[1], Kurt Jordan[2], Nicolas Mathieu[1], Martin Sattelmayer[1], Hans Spring[2], Roger Hilfiker[1,3,4]

[1] HES-SO Valais-Wallis, University of Applied Sciences and Arts Western Switzerland Valais, School of Health Sciences, Physiotherapy, Leukerbad, Switzerland
[2] Swiss Olympic Medical Center Leukerbad, Switzerland
[3] HES-SO Valais-Wallis, University of Applied Sciences and Arts Western Switzerland Valais, School of Health Sciences, Health & Social Work, Sion, Switzerland
[4] Institut de recherche en réadaptation et Clinique romande de réadaptation (Suvacare), Sion, Switzerland

Abstract

Objectives: To evaluate the responsiveness of the classic Star Excursion Balance Test (SEBT) and of a modified version performed on a soft surface (Airex Balance-Pad). The classic Star Excursion Balance Test is widely used in persons with chronic ankle instability. For the use in healthy athletes, the test can be made more difficult by performing it on a soft surface.
Design: Pre-post study with balance assessment before and after a five-week intervention.
Setting and participants: A convenience sample of 28 healthy adults performed both SEBT variants at baseline and after five weeks (9 sessions) of sensorimotor and strength training. To quantify the responsiveness of the classic SEBT and the SEBT performed on a soft foam (Airex Balance-Pad), effect sizes and standardized response mean were calculated for both test variants.
Results: The difference between the Standardized Response Mean of the sum of three directions in the SEBT performed on an Airex Balance-Pad and the SEBT performed on a firm ground was 0.15 in favour of the SEBT on the Balance-Pad with a 95% confidence interval of -0.51 to 0.21).
Conclusion: The SEBT performed on a soft underground is slightly more responsive than the SEBT on a firm underground.

Résumé

Objectif: Evaluer la sensibilité au changement du classique Star Excursion Balance Test (SEBT) et une version modifiée accomplie sur une surface instable (Airex Balance-Pad). Le classique SEBT est souvent utilisé pour évaluer les personnes ayant une instabilité chronique de cheville. Chez les athlètes sains, le test peut être effectué de manière plus ardue si effectué sur une surface instable. Design: étude de cohorte avec évaluation de l'équilibre avant et après une intervention de 5 semaines. Méthode et participants: un échantillon de convenance de 28 adultes sains accomplissant les deux variantes de test avant et après 5 semaines d'entraînement sensori-moteur et de force (9 séances). Pour quantifier la sensibilité au changement du classique SEBT et le SEBT accompli sur une surface instable, la taille de l'effet et la moyenne standardisée ont été calculées pour les deux variantes de tests. Résultats: La différence entre la moyenne standardisée de la somme des 3 directions du SEBT effectuées sur le Airex Balance-Pad et le SEBT classique était de 0.15 en faveur du SEBT sur la surface instable [IC 95% (-0.51 à 0.21)]. Conclusion: le SEBT effectué sur une surface instable est légèrement plus sensible au changement que le SEBT classique sur une surface stable.

Introduction

Testing static and dynamic postural stability is important as deficiencies in postural stability have been shown to be a risk factor for lower limb injuries and in particular ankle sprains [1, 2]. Furthermore, testing of postural stability is important to evaluate sensorimotor exercise programmes that are run for example to help reduce ankle sprains [3]. Ankle sprains count up to 30% of all sports injuries and the incidence rate in the general population presenting to an emergency department with an ankle sprain was 2.15 per 1000 person-years [4]. The incidence was higher in athletes and after an initial injury (recurrence up to 80%) [5].

Balance can be improved [6, 7] and sensorimotor, neuromuscular, "coordination" or balance training can reduce injuries of the lower extremity in athletes of different high risk sports [8–11].

A widely used balance test is the Star Excursion Balance Test (SEBT), consisting of a set of reaching movements of one leg while standing on the other leg. The classic SEBT is performed on a firm underground. The SEBT is reliable and valid to predict lower extremity injuries [12]. In a prevention-setting with healthy persons, the test might be too easy to detect a meaningful change (ceiling effect). The responsiveness (ability to detect change) of the test might be improved when the SEBT is performed on a soft, compliant and thus unstable underground. The SEBT performed on an Airex Balance-Pad showed to be more difficult in a group of healthy athletes and recreationally active healthy young adults [13]. However, the relative responsiveness of the SEBT performed on a firm underground and the SEBT performed on a compliant surface is unknown. Therefore we set out to evaluate the relative responsiveness of the two versions of the SEBT in a group of healthy young adults who performed a five week (nine sessions) sensorimotor and strength exercise programme.

Methods

Experimental Approach to the Problem
To test the main question of our study, i.e. the comparison of the responsiveness (standardized mean response) between the SEBT on firm underground versus SEBT on a soft underground (Airex Balance Pad), we choose a test-retest design with five weeks of sensorimotor and resistance training (nine sessions) between the two measurements. The five week of resistance training was based on published exercise programmes that are known to improve balance (see e.g. [14, 15]). Because reliability of a test should be known, we calculated the Intra Class Correlation Coefficients for the three repetitions.

Subjects
A convenience sample of 30 healthy persons (20 females, 10 males, mean age 24 SD 4.9) was recruited. Two participants dropped out because of illnesses unrelated to the exercise programme. The mean age of the remaining 28 (20 females, 8 males) participants was 24 years with a SD of 4.9.

All participants signed an informed consent form. Ethical approval was provided by the Cantonal Ethical Committee of the Valais.

Procedures

Interventions
The participants performed sensorimotor exercises on solid surface, on a wobble board, on foam, on a small trampoline, on a kettle board and on a slack-line. In the first week the exercises were performed without additional tasks, in the following week additional tasks such as catching and throwing a ball or small jumps were introduced in a progressive way. In addition, the participants did resistance exercises with 10 to 12 repetitions with high movement velocity and submaximal load and three to four sets per exercise. From the third week on, an isometric pre-fatigue-exercise of 30 seconds was introduced before each set. In the last three weeks plyometric exercises and the contrast-method (e.g. after 12 slow squats the participants performed 12 jumps, i.e. different resistance training modalities were combined) were introduced. The rest-duration between the sets was twice as long as the set duration.

The exercises were chosen in order to maximize the effects on lower leg strength (ankle, knee and hip extensors, knee flexors, hip adductors and abductors, as well as trunk muscles) and to improve static as well as dynamic balance.

Measures
Dynamic balance was measured with the Star Excursion Balance Test (SEBT) [16–18]. Measurements were performed before the five week intervention period (t0, baseline) and after the intervention period (t1) in the same procedure and order, and by the same assessors. The participants had to stand bawrefoot on one leg and reach as far as possible with their other leg in different predefined directions, marked with tape. The reached distance was measured in centimetres and standardized on the leg length by dividing each distance by the leg length and then multiplying by 100 [18]. The normalized values are percentages of the excursion distance in relation to the leg length. We measured the leg lengths from the anterior superior iliac spine to the distal tip of the medial malleolus with a tape measure. Participants lay supine.

The SEBT has high intra- and inter-tester reliability (ICCs between 0.81 and 0.96) [19] and is able to detect balance deficits, especially in patients with chronic ankle instability. We hypothesized that a modified SEBT, i.e. performed on a soft and instable surface, the Airex Balance-Pad (Alcan Airex, Sins, Switzerland) would be better able to detect change; therefore, the participants performed the SEBT three times per direction on a hard surface and three times on the Airex Balance-Pad. The best value of the three trials was taken for the analysis. We used a protocol where only the antero-medial, medial and postero-medial directions were tested [17], see Figure 1. The order of the tests was

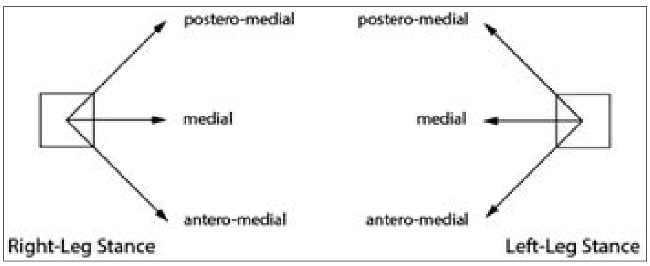

Figure 1: Directions used for the Star Excursion Balance Test. Only one leg was tested, i.e. the leg on which the participant felt more comfortable.

always the same, i.e. we did not randomize the order but kept it the same for all participants. We tested only one leg. To determine the leg, the participants stood for 20 seconds on each leg and decided afterwards on which leg they wanted to test. If they couldn't decide, we ask them on which leg they stand if they kick a ball. This was the leg on which they had to stand during the SEBT. After at least three practice trials for each direction, three trials were measured. The number of practice trials was determined by an observer who evaluated visually with a ruler whether there was still an improvement between two trials. If there was no improvement between the second and the third trial, the measurement trials started. If there was still an improvement, then further practice trials were performed. One examiner marked the most distal location of the reach foot as it contacted the floor. The measurement was discarded if the participant was unable to stay on one leg, when the heel lifted off or when the foot of the reach leg was used to provide support (i.e. did not only touch the ground lightly).

A SEBT composite score was calculated as the sum of the three directions, i.e. by simply adding the values of each of the three directions.

Statistical Analyses

The test-retest reliability of the three tests for each direction in the SEBT, performed on one day before the intervention started was assessed with Intraclass Correlation Coefficients with a 2-way random model and the absolute agreement method implemented in SPSS [20]. To compare the responsiveness across both versions of the SEBT, standardized response means (SRMs) were calculated for each in participant by dividing the change score (end of exercise period minus baseline values) by the standard deviations of the change scores of the whole sample. We compared SRMs across both versions with a paired t-test, following the method used by Puhan et al. [21]. SRMs were calculated in STATA, version 12.0 (StataCorp LP, USA). In addition, we calculated effect sizes, i.e. mean score change divided by the standard deviation of time point one, as well as paired t-tests to evaluate the statistical significance of the difference before-and-after the interventions in the SEBT test on (a) firm surface an (b) soft surface.

We calculated the standard error of the measurement (SEM) with the formula: SD(pooled)*(√(1-ICC).

The SEM can be used to determine the "minimum detectable change" (MDC95%) for the measurement, i.e., the amount of change required in an individual's balance score, in order to be sure with a given level of confidence, that the "change is real", over and above measurement error. The MDC95% was calculated with the formula: 1.96*(√2)*SEM, which is equivalent to 2.77*SEM.

Results

Reliability of the modified SEBT, i.e. performed on the Airex Balance-Pad was good and comparable with the classic SEBT. The ICC values were (first value for SEBT on the Balance-Pad, second values for classic SEBT): 0.87 vs. 0.78 (antero-medial), 0.86 vs. 0.89 (medial), 0.85 vs. 0.87 (postero-medial).

Table 1 shows values at baseline, at the end of the exercise period and the change. All changes were statistically significant and the effect sizes were moderate to large. However, the changes were smaller than the minimal detectable change at a 95% confidence level (see table 1).

The responsiveness for the sum of all three directions was better for the modified SEBT compared to the classic SEBT: SRM of 1.27 (95% CI 0.88 to 1.66) for the classic SEBT and a SRM of 1.42 (95% CI 1.04 to 1.81) for the modified SEBT. The difference in the SRM between both variants was -0.15 (95% -0.51 to 0.21). There is only a small advantage for the responsiveness of the modified SEBT and this difference is still compatible with the Null-Hypothesis that there is no difference in the responsiveness. Table 1 shows the SRMs for all directions. There was only one difference in the SRMs statistically significant: In the postero-medial direction, the SRM was significantly higher for the modified SEBT. The difference was -0.53 with a 95% CI from -0.96 to -0.10.

Direction	Mean t0 firm (sd)	Mean t1 firm (sd)	Change firm (95% CI)	Mean t0 Airex (sd)	Mean t1 Airex (sd)	Change Airex (95% CI)	SRM firm ground SRM (95% CI)	SRM Airex SRM (95% CI)	Difference SRM (95% CI)
Antero-Medial	90.9 (5.2)	93.6 (6.0)	2.6 (1.0 to 4.2)	86.0 (6.3)	90.1 (8.1)	4.1 (1.9 to 6.3)	0.64 (0.25 to 1.03)	0.73 (0.35 to 1.12)	-0.09 (-0.60 to 0.41)
			ES: 0.51			ES: 0.50			
			MDC95: 5.2			MDC95: 8.2			
Medial	94.1 (7.1)	99.1 (7.0)	5.0 (3.4 to 6.7)	87.6 (8.0)	94.1 (8.4)	6.5 (4.2 to 8.9)	1.19 (0.80 to 1.57)	1.07 (0.68 to 1.46)	0.11 (-9.29 to 0.52)
			ES: 0.44			ES: 0.78			
			MDC95: 7.4			MDC95: 7.4			
Postero-Medial	97.7 (10.0)	103.7 (8.7)	6.0 (3.9 to 8.2)	90.5 (10.7)	97.7 (8.9)	7.2 (5.5 to 8.9)	1.09 (0.71 to 1.48)	1.62 (1.24 to 2.01)	-0.53 (-0.96 to -0.10)
			ES: 0.69			ES: 0.80			
			MDC95: 10.8			MDC95: 10.7			
Sum over three directions	282.7 (20.2)	296.4 (23.7)	13.7 (9.5 to 17.9)	264.1 (23.7)	281.9 (24.0)	17.8 (13.0 to 22.7)	1.27 (0.88 to 1.66)	1.42 (1.04 to 1.81)	-0.15 (-0.51 to 0.21)
			ES: 0.67			ES: 0.74			

Table 1: Standardized Response Means for the different SEBT versions.

T0 = baseline values, T1 = values after the intervention period. ES: Effect Size (mean difference divided by baseline standard deviation). MDC95: minimal detectable change at the 95% confidence level. SRM: Standardized Response Mean (dividing the change scores of each participant (follow-up minus baseline score) by the standard deviation of the change scores of the whole sample. CI: Confidence Interval.

Discussion

This study with 28 healthy young adults showed that a modified version of the Star Excursion Balance Test (SEBT) performed on a soft and unstable underground (Airex Balance-Pad) was only slightly more responsive than the classic version of the SEBT on firm underground. Both versions showed good reliability.

There are some limitations of this study: (a) the small sample size led to wide confidence intervals in the differences of the SRM and therefore we cannot exclude that the true difference in the responsiveness between the two versions of the SEBT could be as large as 0.51, which might be considered as a large difference. (b) we only included healthy participants which provides information for a prevention but not for a rehabilitation setting. (c) we did not analyse all possible directions of the SEBT but only three. (d) Although the effect sizes for the changes in the SEBT values before the interventions and after the five weeks intervention are moderate to large, the changes are smaller than the minimal detectable change at the 95% confidence level. Therefore, if an individual person shows a change in the SEBT similar to our mean change, we still couldn't exclude that the observed change is only due to a measurement error. Munro & Herrington [22] reported similar values for the minimal detectable change (they called it smallest detectable difference, but the formula were the same as in our study). If we adhere strictly to the MDC95, we have to state that the SEBT might only detect very large changes. Some experts state that the 95% confidence level for the minimal detectable change is to stringent and that lower confidence values, e.g. at the 90% confidence level should be reported [23, 24]. However, even with these lower MDC confidence levels, our changes are lower than the MDC values. For the group level, this does not pose any problem, but for the assessment of an individual person we should be aware of this limitation. It is important to consider a further limitation that our study and the study of Munro & Herrington have in common [22]: in our and in their study the MDC (they called it SDD) was calculated with standard deviations and standard error of measurements assessed from healthy young participants. Strictly spoken, we cannot extrapolate these values to patients and therefore MDCs for patients are still to be evaluated. Furthermore, it could be that more practice trials could have improved the reliability of our measurements. Only one study reported on the use of the SEBT on a soft foam underground [13]. The SEBT-scores were 5% to 10% lower for the version on soft underground. Our results are consistent with this.

Several studies used the SEBT to evaluate changes before and after an intervention [25–27]. Compared to the study of Filipa et al [26], we had a shorter training period (five weeks versus eight weeks and less training sessions (9 versus 16). The exercises used were similar. The changes were in a similar order compared to our study. Filipa et al and Hale et al. had a control group that showed no improvement in the SEBT test. These findings support the statement that our observed changes are not only due to a simple learning effect.

Further research should evaluate the SEBT on a soft compliant underground in patients with chronic ankle instability of patients after knee injuries. Furthermore, it is unclear whether temperature of the foam has an influence on the compliance, i.e. on the amount of the destabilizing effect of the cushion (pad). The influence of temperature or other parameters (age of the tissue, time in use, etc.) on the compliance of the pad should be evaluated.

Implication for the practice: Because of this increased difficulty and the improved responsiveness, although not in all directions statistically significant, we recommend for healthy persons to perform the Star Excursion Balance Test on a soft underground.

Conclusion

The reliability of the Star Excursion Balance Test performed on firm underground and the version performed on a compliant, soft underground is good. The responsiveness is slightly better with the version on the soft, compliant and thus unstable underground, but this difference was only significant for the postero-medial direction. Therefore, we suggest that the SEBT version on the soft underground is used in a prevention setting or in rehabilitation of sportive patients.

Acknowledgement

The authors thank all participants.

Ethical Approval: Ethical approval was provided by the Cantonal Ethical Committee of the Valais (Nr: CCVEM 046/08).
Funding: No funding.
Conflict of Interest: All authors declare that there are no conflicts of interest.

Corresponding author:

Roger Hilfiker
roger.hilfiker@hevs.ch
Rathausstrasse 8, 3954 Leukerbad, Switzerland

References

1 Plisky PJ, Rauh MJ, Kaminski TW, Underwood FB. Star Excursion Balance Test as a predictor of lower extremity injury in high school basketball players. The Journal of orthopaedic and sports physical therapy. 2006;36(12):911-9.
2 Olmsted LC, Carcia CR, Hertel J, Shultz SJ. Efficacy of the Star Excursion Balance Tests in Detecting Reach Deficits in Subjects With Chronic Ankle Instability. Journal of Athletic Training. 2002;37(4):501.
3 McGuine TA, Greene JJ, Best T, Leverson G. Balance as a predictor of ankle injuries in high school basketball players. Clin J Sport Med. 2000;10(4):239-44.
4 Waterman BR, Owens BD, Davey S, Zacchilli MA, Belmont PJ, Jr. The epidemiology of ankle sprains in the United States. J Bone Joint Surg Am. 2010;92(13):2279-84.
5 Gurney B, Milani J, Pedersen M. Role of fatigue on proprioception of the ankle. J Exerc Phys online. 2000;3(1).
6 DiStefano LJ, Clark MA, Padua DA. Evidence supporting balance training in healthy individuals: a systemic review. J Strength Cond Res. 2009;23(9):2718-31.
7 Zech A, Hubscher M, Vogt L, Banzer W, Hansel F, Pfeifer K. Balance training for neuromuscular control and performance enhancement: a systematic review. Journal of athletic training. 2010;45(4):392-403.

8 McKeon PO, Hertel J. Systematic review of postural control and lateral ankle instability, part II: is balance training clinically effective? J Athl Train. 2008;43(3):305-15.
9 Hubscher M, Zech A, Pfeifer K, Hansel F, Vogt L, Banzer W. Neuromuscular training for sports injury prevention: a systematic review. Medicine and science in sports and exercise. 2010;42(3):413-21.
10 Aaltonen S, Karjalainen H, Heinonen A, Parkkari J, Kujala UM. Prevention of sports injuries: systematic review of randomized controlled trials. Arch Intern Med. 2007;167(15):1585-92.
11 Abernethy L, Bleakley C. Strategies to prevent injury in adolescent sport: a systematic review. Br J Sports Med. 2007;41(10):627-38.
12 Gribble PA, Hertel J, Plisky P. Using the Star Excursion Balance Test to assess dynamic postural-control deficits and outcomes in lower extremity injury: a literature and systematic review. J Athl Train. 2012;47(3):339-57.
13 Sabin MJ, Ebersole KT, Martindale AR, Price JW, Broglio SP. Balance performance in male and female collegiate basketball athletes: influence of testing surface. J Strength Cond Res. 2010;24(8):2073-8.
14 Howe T, Rochester L, Jackson A, Banks P, Blair V. Exercise for improving balance in older people (Review). 2008.
15 Granacher U, Muehlbauer T, Zahner L, Gollhofer A, Kressig RW. Comparison of traditional and recent approaches in the promotion of balance and strength in older adults. Sports Med. 2011;41(5):377-400.
16 Kinzey SJ, Armstrong CW. The reliability of the star-excursion test in assessing dynamic balance. J Orthop Sports Phys Ther. 1998;27(5):356-60.
17 Hertel J, Braham RA, Hale SA, Olmsted-Kramer LC. Simplifying the star excursion balance test: analyses of subjects with and without chronic ankle instability. J Orthop Sports Phys Ther. 2006;36(3):131-7.
18 Gribble PA, Hertel J. Considerations for Normalizing Measures of the Star Excursion Balance Test. Measurement in Physical Education and Exercise Science. 2003;7(2):89-100.
19 Hertel J, Miller SJ, Denegar CA. Intratester and Intertester Reliability During the Star Excursion Balance Tests. JOURNAL OF SPORT REHABILITATION. 2000;9(2):104-16.
20 Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19(1):231-40.
21 Puhan MA, Guyatt GH, Goldstein R, Mador J, McKim D, Stahl E, et al. Relative responsiveness of the Chronic Respiratory Questionnaire, St. Georges Respiratory Questionnaire and four other health-related quality of life instruments for patients with chronic lung disease. Respir Med. 2007;101(2):308-16.
22 Munro AG, Herrington LC. Between-session reliability of the star excursion balance test. Phys Ther Sport. 2010;11(4):128-32.
23 Hopkins WG. Measures of reliability in sports medicine and science. Sports medicine. 2000;30(1):1-15.
24 Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Physical Therapy. 2006;86(5):735-43.
25 Dannelly BD, Otey SC, Croy T, Harrison B, Rynders CA, Hertel JN, et al. The effectiveness of traditional and sling exercise strength training in women. Journal of strength and conditioning research / National Strength & Conditioning Association. 2011;25(2):464-71.
26 Filipa A, Byrnes R, Paterno MV, Myer GD, Hewett TE. Neuromuscular training improves performance on the star excursion balance test in young female athletes. The Journal of orthopaedic and sports physical therapy. 2010;40(9):551-8.
27 Plisky PJ, Gorman PP, Butler RJ, Kiesel KB, Underwood FB, Elkins B. The reliability of an instrumented device for measuring components of the star excursion balance test. N Am J Sports Phys Ther. 2009;4(2):92-9.