THE VALIDITY OF THE MODIFIED STAR EXCURSION BALANCE TEST AS A PREDICTOR OF KNEE EXTENSOR AND HIP ABDUCTOR STRENGTH

Lee Shimwell *1, Francis Fatoye 2, James Selfe 3.


2 Professor of Health Economics and Outcomes, Department of Health Professions, Faculty of Health, Psychology and Social Care, Manchester Metropolitan University (United kingdom).

3 Professor of Physiotherapy, Department of Health Professions, Manchester Metropolitan University (United kingdom).

ABSTRACT

Objectives: To test the construct validity of the modified Star Excursion Balance Test (SEBT) in predicting the strength of knee extensors and hip abductors and to compare performance between dominant and non-dominant legs.

Strength, as measured using hand held dynamometry, has been used as one of a variety of measures in order to subgroup patients with patellofemoral pain. This method is not well received in clinical practice due to time constraints and cost. In contrast, the SEBT is a quick and simple functional measure of dynamic stability. This study examined the validity of the SEBT as an alternate strength measure.

Design: Within subject correlation. Each participant recorded maximum isometric contractions for hip abduction and knee extension using a HHD and performed the modified SEBT, on both legs, within one test period. Paired t tests were used to compare dominant and non-dominant legs and Pearson’s correlation analyses were used to explore for associations.

Setting: Non-clinical environment,

Participants: Eighteen healthy male amateur runner’s between 18 and 39 years old (mean age 36.1 years).

Main outcome measures: Knee and hip moments normalised for leg length (Nm/m) for the HHD and percent of leg length reach score for each SEBT reach as well as a composite of all three.

Results: There was no significant difference between dominant and non-dominant legs for all the tests (knee extension p = 0.72, hip abduction p = 0.90, SEBT composite p = 0.86) therefore data was combined into one set. There was no significant relationship between either hip abduction strength (r = 0.28, p = 0.11), or knee extension strength (r= 0.16, p = 0.17), with any combination of the SEBT. A moderate relationship (r = 0.52, p = 0.38) between hip abduction and SEBT posterolateral reach was seen, however, this was not statistically significant.

Conclusion: There was no statistically significant association between either knee extension or hip abduction strength with the modified SEBT. This suggests that strength is unlikely to be a primary construct of the test and the SEBT is not a useful replacement for the HHD when testing strength.

KEY WORDS: SEBT, Patellofemoral, Subgroups, Hand held dynamometer, Runners.

Address for correspondence: Mr. Lee Shimwell, Extended Scope Physiotherapist, Integrated Musculoskeletal, Pain and Rheumatology Service, East Lancashire Hospitals NHS Trust (United kingdom). E-Mail: lee_shimwell@hotmail.com
INTRODUCTION

In a consensus statement, Witvrouw et al [1] suggested that identification of subgroups was a key priority for Patellofemoral Pain (PFP) research, with the intention that this may help focus future treatment strategies. Subsequently, three subgroups have been identified by Selfe et al [2]. Maximal isometric knee extension (quadriceps) and hip abduction (gluteal) strength, tested using a hand held dynamometer (HHD), was an important factor in defining the subgroups.

It has been demonstrated that the HHD is a reliable test of strength, with a moderate-high correlation to a Cybex isokinetic dynamometer in healthy male and female participants (knee extension r = 0.71, \( r^2 = 0.504 \) [3], p = 0.001; hip abduction r = 0.49-0.65, p < 0.05, knee extension r = 0.48, p < 0.01 [4]; knee extension r = 0.91 [5]). However, anecdotal clinician feedback from Selfe et al [2] suggested that the use of a HHD muscle strength test, although acceptable for research purposes, is regarded as expensive and impractical in routine clinical practice.

It has been suggested that clinician friendly Physical Performance Tests (PPT) are functional and could act as an alternative method of determining strength [6] with the advantage that they incorporate multi-joint neuromuscular control factors. In addition, Hegedus et al [7] suggested that these tests are easy to administer, time efficient, can be used in multiple settings and are inexpensive. In a systematic review they identified the 13 most commonly used in sport lower limb PPTs. The most studied were the vertical jump test, the single leg hop for distance test and the modified Star Excursion Balance Test (SEBT).

Whereas there is strong evidence that the vertical jump test is responsive to sport specific training [8] there is no evidence of validity [7] for other constructs, such as strength.

For the single leg hop for distance test, Weng et al [3] tested the correlation between it and both the HHD and Cybex but found no significant relationship with either dynamometry method for knee extension strength. A possible reason for this they suggested that the participants were not allowed to swing their arms in a natural way which could have impacted on the results. However, in agreement with these findings, Hegedus et al [7] summarise that despite being considered by clinicians as the best of a range of hop tests, and as the best candidate to measure quadriceps strength at the knee, the single leg hop test is likely to assess a different construct than strength. The SEBT is primarily a balance test that combines both hip and knee control elements. It has been described as a test of dynamic postural control because it challenges the participant to maintain a single leg stance whilst reaching with the other leg consecutively along the three arms of a ‘Y’, which requires the components that make up neuromuscular control such as strength, proprioception and flexibility [9].

McMullen et al [10] demonstrated reduced performance (p <0.05) on the SEBT after fatiguing the gluteus medius muscle by 15%. Gribble et al [11] also found that knee fatigue, achieved through repeated lunging, produced reductions in dynamic stability for all reaches with men showing a 4% larger decline than women (p = 0.2). These findings could indirectly imply that strength is one factor in performance on the SEBT. Clagg et al’s [12] primary aim was to compare SEBT performance between patients with an ACL reconstruction (ACLR) and healthy controls at the point of return to sport. However, they also looked at a secondary comparison of isokinetic dynamometry strength with SEBT performance finding a statistically significant association on the ACLR involved limb between posterolateral reach and knee extension strength (r = 0.28, p = 0.03) and all three reaches for hip abduction strength (r = 0.28-0.41, p = 0.01-0.03). However, although statistically significant, r values were low indicating a weak association. As the focus of the study was on ACLR patients, the healthy control group scores were reported in relation to these and not in sufficient detail to consider the question of strength and SEBT performance in healthy participants.

Thorpe and Ebersole [13] also found low correlations (r = -0.58-0.20) and with no statistical
significance (p > 0.05) in a sample of healthy soccer athletes. They chose to use three reaches (anterior, posterior and medial) of the original eight, instead of the more recently recommended modified SEBT [14] that uses anterior (ANT), posteromedial (PM) and posterolateral (PL). Importantly they did not measure hip abduction, which was actually the strongest correlation shown by Clagg et al [12] in the ACLR group. Therefore, it remains unknown if the SEBT is a valid test of both hip and knee strength in a healthy population.

These previous studies into the SEBT’s association with strength have focused on either patients with pathological conditions [12] or used a variation of the original SEBT reaches [13] and not the subsequently standardised and recommended modified version. The primary aim of this study was to validate the SEBT as a measure of hip abduction and knee extension strength in comparison to a HHD and the primary hypothesis is that there will be a positive association. This may allow the SEBT to replace the HHD in future studies that subgroup PFP patients. In addition, Thorpe and Ebersole [13] demonstrated equal performance between limbs, and that limb preference did not result in differences in SEBT performance supporting similar findings in healthy individuals [7, 9]. Thus, the secondary hypothesis is that there will be no statistically significant difference in performance between limb dominance. If this is accepted there is a methodological advantage that both legs of each participant can be tested to potentially increase the power of the study.

MATERIALS AND METHODS

Participants: 20 Participants were recruited from two UK amateur running clubs. Each had time to consider their involvement, ask questions and provided written informed consent prior to participation. Ethical approval was obtained from Manchester Metropolitan University Ethics Committee in accordance with the principles documented in the Declaration of Helsinki (2013). Participants were made aware that they could withdraw at any time without providing a reason.

A power calculation performed on SEBT using effect size (0.81 and SD 7.13), from a comparable population of healthy and active participants [15], revealed that 17 participants were required for the study. To allow for dropouts a total of 20 were recruited into the study. Using a purposive sample of healthy male runners in the 18-39 age range ensured there was good ecological validity with the strong and physically active PFP subgroup as identified by Selfe et al [2]. All participants met this inclusion criteria with running at least once per week set as an acceptable frequency for physical activity. In addition they were required to have had no reduction in running frequency due to injury or pain within the last three months, no exercise on the day of testing and no history of lower limb surgery as these have all been shown to negatively influence SEBT performance [14].

Procedures: Participants were asked to attend only one testing session, lasting approximately 30 minutes. They were first asked to determine their dominant leg by kicking a soccer ball [16] which has been shown to have high test-retest reliability (Kappa = 0.77) and high internal consistency (Cronbachs alpha p = 0.83) with three other bilateral lower limb skill dominance tests [17]. A warm up consisting of mid-range free standing squats was performed, in two sets of five repetitions each and with a 30 second break between. Using a within subject correlation design the following three tests were then performed on both legs of each participant and in a randomized order, determined by the toss of a coin: 1) Maximum isometric knee extension Quadriceps strength using a Lafayette (USA) manual muscle testing system HHD (model 01163, range 0-136kg), 2) Maximum isometric hip abduction Gluteal strength using the same Lafayette HHD and 3) Modified SEBT, using the Y-balance test system (functionalmovement.com, Danville, VA). The lead researcher performed all measurements as Intra-tester reliability of both the HHD (ICC = 0.70-0.94[18]) and the Y-balance system (ICC = 0.85-0.89[19]) are reported to be high.

For the isometric strength tests the HHD was mounted against a stabilisation strap as described in detail in previous studies [2,3,20,21]. For knee extension the participant was seated on a plinth (Figure 1) and for hip abduction they were in a side lying position
difference between limb dominance in any of the measures the data were combined to further increase the power of the data set (n = 36). Pearson’s correlation coefficients (r = 0.1-0.3 low, 0.3-0.5 med, 0.5-1.0 strong) were used to look for associations in the data between knee extension strength and hip abduction strength against all four components of the SEBT. A Bonferroni calculation was used with significance as p = 0.01.

Fig. 1: Hand Held Dynamometer Knee Extension.

Fig. 2: Hand Held Dynamometer Hip Abduction.

Fig. 3: Performance on the modified Star Excursion Balance Test (SEBT) on the left leg using the Y balance test system on the:

A) Posterolateral reach  B) Posteromedial reach  C) Anterior reach

Statistical Analysis: The Statistical Package for the Social Sciences (SPSS, IBM Version 22, USA) was used to conduct all statistical tests. Descriptive statistics for knee extension and hip abduction strength are displayed in normalised torques (Nm/m) by multiplying the maximal force (Newton’s) with the moment arm (in metres) and for the SEBT are displayed as a percentage of leg length (metres) to normalise for height [9,20].

RESULT

Data were recorded from 18 participants (age 36.1± 2.43 years, leg length 0.94± 0.05 m, Tibia
moment arm 0.33± 0.02 m, Femur moment arm 0.36± 0.02 m) for both dominant and non-dominant legs. One participant had a recent history of pain in one leg and was therefore excluded and one participant failed to attend.

Table 1: Mean scores for all tests and difference for leg dominance using paired samples t test.

<table>
<thead>
<tr>
<th></th>
<th>All (n = 36)</th>
<th>Dominant (n = 18)</th>
<th>Non-Dominant (n = 18)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>*HHD knee extension (Nm/m)</td>
<td>163.48 (59.46)</td>
<td>162.23 (61.15)</td>
<td>164.73 (59.47)</td>
<td>-0.35</td>
<td>0.72</td>
</tr>
<tr>
<td>*HHD hip abduction (Nm/m)</td>
<td>127.16 (33.74)</td>
<td>127.36 (32.54)</td>
<td>126.92 (35.86)</td>
<td>0.12</td>
<td>0.9</td>
</tr>
<tr>
<td>†SEBT comp (% of leg length)‡</td>
<td>92.16 (5.56)</td>
<td>92.24 (5.76)</td>
<td>92.08 (5.53)</td>
<td>0.16</td>
<td>0.86</td>
</tr>
<tr>
<td>†SEBT anterior (%LL)</td>
<td>64.41 (5.39)</td>
<td>63.34 (5.30)</td>
<td>64.47 (4.42)</td>
<td>-0.15</td>
<td>0.82</td>
</tr>
<tr>
<td>†SEBT posteromedial (%LL)</td>
<td>107.14 (8.84)</td>
<td>107.39 (8.53)</td>
<td>106.89 (9.38)</td>
<td>0.36</td>
<td>0.72</td>
</tr>
<tr>
<td>†SEBT posterolateral (%LL)</td>
<td>104.94 (7.43)</td>
<td>105.02 (7.96)</td>
<td>104.87 (7.38)</td>
<td>0.12</td>
<td>0.9</td>
</tr>
</tbody>
</table>

p > 0.05 shows no significant difference in means

*HHD- Hand Held Dynamometer
†SEBT- Star Excursion Balance Test
‡ Calculated by dividing the combined total by three.

Table 2: Pearson’s correlation coefficient between tasks.

<table>
<thead>
<tr>
<th></th>
<th>(n = 36)</th>
<th>Dominant and Non-Dominant limbs</th>
<th>r</th>
<th>r²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHD knee ext- SEBT ant</td>
<td>0.27</td>
<td>-0.07</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHD knee ext- SEBT post.med</td>
<td>0.2</td>
<td>0.04</td>
<td>0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHD knee ext- SEBT post.lat</td>
<td>0.52</td>
<td>0.27</td>
<td>0.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHD knee ext- SEBT comp</td>
<td>0.28</td>
<td>0.08</td>
<td>0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHD hip abd- SEBT ant</td>
<td>0.12</td>
<td>0.01</td>
<td>0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHD hip abd- SEBT post.med</td>
<td>0.26</td>
<td>0.07</td>
<td>0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHD hip abd- SEBT post.lat</td>
<td>-0.26</td>
<td>0.07</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHD hip abd- SEBT comp</td>
<td>0.16</td>
<td>0.02</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistically Significant at p = <0.01. r = 0.1-0.3 low, 0.3-0.5 med, 0.5-1.0 strong.

Table 3: A comparison of Hand Held dynamometer (HHD) scores with other studies.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>Healthy male amateur runners</td>
<td>Male PFP patients Strong sub-group age &gt;30</td>
<td>Healthy male</td>
</tr>
<tr>
<td>Nm/m Hip</td>
<td>127.1</td>
<td>163</td>
<td>-</td>
</tr>
<tr>
<td>Nm/m Knee</td>
<td>163.5</td>
<td>226</td>
<td>-</td>
</tr>
<tr>
<td>Newtons Hip</td>
<td>305</td>
<td>-</td>
<td>103.8</td>
</tr>
<tr>
<td>Newtons Knee</td>
<td>492</td>
<td>-</td>
<td>215</td>
</tr>
</tbody>
</table>

There were no statistically significant differences between the dominant and non-dominant legs for any of the tests.

The results of correlation analysis can be seen in Table 2. From the table it can be seen that there was a moderate, but statistically non-significant, positive correlation between HHD knee extension strength and SEBT posterolateral reach (r = 0.52, r² 0.27, p = 0.38). All other combinations between knee extension or hip abduction and the three SEBT tasks, plus the composite score, showed only low associations (r = <0.3) that did not achieve statistical significance.

Visual analysis of the scatterplots (Figure 4) showed two distinct groups of data for hip abduction Vs SEBT posteromedial, posterolateral and composite scores, but not for any other of the comparisons.

Table 4: A comparison of Star Excursion Balance Test (SEBT) scores with other studies.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior (%LL)</td>
<td>64.4</td>
<td>67.1</td>
<td>79.2</td>
<td>0.89</td>
<td>94.1</td>
</tr>
<tr>
<td>Post.med (%LL)</td>
<td>107.1</td>
<td>98.1</td>
<td>95.6</td>
<td>-</td>
<td>116.1</td>
</tr>
<tr>
<td>Post.lat (%LL)</td>
<td>104.9</td>
<td>79.9</td>
<td>90.4</td>
<td>-</td>
<td>128.7</td>
</tr>
<tr>
<td>Composite (%LL) ‡</td>
<td>92.2</td>
<td>77</td>
<td>88.4</td>
<td>-</td>
<td>105</td>
</tr>
</tbody>
</table>

*NLL- Normalized score for height represented as a percentage of limb length.
‡ Calculated by dividing the combined total by three.

Fig. 4: Scatterplots for hip abduction comparison with SEBT posteromedial, posterolateral and composite correlations showing two distinct groups; One >150Mm and one <150Mm.
DISCUSSION

The aim of the study was to test the construct validity of the modified SEBT in predicting the strength of knee extensors and hip abductors and to compare performance between dominant and non-dominant legs. There was no statistically significant correlation between either hip abduction strength, or knee extension strength, with any of the four components of the SEBT, including the commonly used composite score. Therefore the results reject our a priori primary hypothesis. The secondary hypothesis for the current study was that there would be no difference in performance between dominant and non-dominant legs and this is accepted. This supports previous findings in athletic and healthy populations [7,9] including soccer [13] which, unlike running, has a specific limb dominance. The lack of difference also extended to both of the HHD strength tests.

To the best of our knowledge this is the first study to examine SEBT performance compared to lower extremity maximum isometric strength tests using a HHD and also the first time that SEBT normative scores have been quantified in a population of healthy amateur runners (Table 1). As the SEBT is not only used in a clinical setting, but is also used in sport as a physical performance screening test, it is useful to know that strength may not be one of its dominant constructs and this may help prevent erroneous interpretations of its outcomes.

Similar to findings by Clagg et al [12] on ACLR participants, there was a moderate but statistically non-significant positive correlation (r = 0.52, p = 0.38) between knee extension strength and the posterolateral reach. However, in the healthy control group participants they also reported a correlation between hip abduction strength and the posterolateral reach that we have not found.

Our findings are also in agreement with those of Thorpe and Ebersole [13], in healthy and athletic (soccer) participants. However the results are not directly comparable mainly because they used a different variation of reaches and not the modified recommended version of the SEBT, or report a composite score as we have here, but also because female participants were used.

Both of these studies used isokinetic strength testing and as there is no consensus as to which method most effectively measures strength [20] it is more difficult to directly compare findings. Both Weng et al [3] and Katoh et al [4] found higher scores in the isokinetic method and speculated that this may because of the increased stabilisation that an isokinetic chair provides but also concluded that the HHD isometric method showed moderate to high correlations and could be considered a valid measure of strength. This is supported in a literature review of hip and knee muscle function testing by Maffiuletti [20] who concluded that it may even be preferable because of its relative simplicity and reliability. Direct comparison of the HHD results from this study with others can be seen in Table 3. Both Weng et al [3] (Nm/Kg) and Katoh et al [4] (N) reported isometric strength differently to us (Nm/m). Whilst both used healthy subjects Katoh et al [15] did not report the participants as athletic and this could explain the higher scores in this study. However, in contrast, Selfe et al [2] reported higher scores in a PFP patient population for the strong male subgroup (age >30; comparable to this study) which could support the theory that this subgroup of PFP patients have overload of the knee as a possible cause of pain.

When interpreting the SEBT findings it should be noted that negative performance on the posteromedial and posterolateral reach directions have been linked to hip and knee joint range of movement (ROM) [22] and the anterior reach direction to ankle dorsiflexion ROM [24]. None of these were controlled for in this study and may have influenced the results. To do so would have reduced the SEBT’s practical use as a potential screening tool of strength in the wider population.

Another contributing factor to strength testing is the time of day effect when maximal muscle contractions are influenced by the circadian fluctuations in muscle temperature that occur throughout the day with a 10% positive difference having been observed in an evening when compared to the morning within the same subjects [25]. In contrast, Gribble et al [26] reported better SEBT performance in a morning
compared to afternoon or evening. Again, this was not controlled for in this study. Most tests were performed in the evening, as it was more convenient to participants, and this could have positively benefited the HHD tests over the SEBT. All participants were distance runners and this activity requires muscle stamina more than peak strength. As the tests here are a measure of maximal strength this could also have influenced the results. Sprinters, for example, may have had different outcomes. The SEBT scores from this study and other studies using healthy participants can be seen in Table 4 to allow for a visual comparison. Neither Clagg et al [12] or Thorpe and Ebersole [13] provide adequate SEBT data to be included. The participants in this study (composite reach score = 92.2%, anterior reach = 66.4%) generally performed better than in a group of non-sporting, healthy, males (composite = 77%, anterior = 67.1% [17]) as well as ‘recreational’ healthy male athletes (composite = 88.4%, anterior = 79.2% [9]) although less well on the anterior reach. In contrast, they performed less well on all four components than a group of high school male basketball players (composite = 103%, anterior = 84.1% [27]). Basketball requires more explosive muscle actions than running and this for example provides support to the theory that the type of sporting activity, or even lack of, might bias the SEBT outcome.

These factors would likely affect the overall scores consistently and do not explain the presence of two distinct groups in the data on the hip abduction scatterplots (Figure 4). As the only data recorded was participant age, leg dominance, moment arm length and test outcome scores, it is not possible to retrospectively find a link to these two groups based on any of the above and is a limitation of our study design. However, it was observed that some participants developed a technique during their practice reaches that involved flexing forwards at the trunk without violating the rules of the test. Others maintained a more upright posture. It could be speculated that a measurement of isometric strength does not allow for the trunk rotation required to perform the posterior reaches, and especially the posterolateral (showing the most distinct groupings). The variation in techniques could be a better adaptation to this requirement resulting in the two distinct groups. Eltoukhy et al [28] provides some evidence for this. They used tortuosity (an estimate of twisting during a pattern of motion) to look more closely at these segmental movement strategies finding that each reach distance on the SEBT could be predicted using tortuosity. Further research could look more closely at the effect these strategies have on correlations with strength.

As a PPT, that it does not correlate with lower extremity strength is in keeping with the results of Weng et al [3] who demonstrated, also in healthy participants, that a single leg hop for distance does not have association with quadriceps strength when tested using both isometric HHD and isokinetic techniques. In contrast, Hamilton et al [29] demonstrated that a triple hop for distance test was able to predict the variance in Quadriceps strength of healthy soccer players using isokinetic dynamometry at 60°/s ($R^2 = 0.43, p < 0.01$) and at 180°/s ($R^2 = 0.52, p < 0.01$), although a 70° arc of motion and short rest periods could have influenced peak torque. Although this test did not include hip abduction strength comparisons, it could point to an alternative for future research or it may also simply be, as Hegedus et al [7] state, that PPT’s and strength are simply different constructs.

One of the strengths of this study is that it was adequately powered. The finding that there was no difference between leg dominance, and thus being able to use both legs of each participant further increased this. Another positive is that the participant group shared ecological validity with the ‘strong’ PFP subgroup, and also other research into SEBT performance. However, it is possible that a different result may have been found on females. Using healthy participants, instead of runners with PFP, limits the confounding effects of clinical conditions that have been shown to influence SEBT performance [14] but also limits the ability to extrapolate the results to patient groups. Caution should also be used when applying the findings to other sporting groups and age ranges.

CONCLUSION
There is no statistically significant association
between knee extension and hip abduction strength, when measured using maximum isometric HHD tests, and the modified SEBT in healthy athletic male participants. There is a moderate, but statistically non-significant, positive correlation between knee extension and the posterolateral reach but it is unlikely that strength is a major construct of the SEBT. Further studies are required to confirm these findings in patients with pathological conditions and whether the SEBT may have a use when combined with other measures of strength.

Conflicts of interest: None

REFERENCES


How to cite this article: