Note. This article will be published in a forthcoming issue of the *Journal of Applied Biomechanics*. The article appears here in its accepted, peer-reviewed form, as it was provided by the submitting author. It has not been copyedited, proofread, or formatted by the publisher.

Section: Original Research

Article Title: Tests to Measure Core Stability in Laboratory and Field Settings: Reliability and Correlation Analyses

Authors: Francisco J. Vera-Garcia, Diego López-Plaza, Casto Juan-Recio, and David Barbado

Affiliations: Department of Sport Science, Sport Research Centre, Miguel Hernández University of Elche, Elche (Alicante), Spain.

Journal: *Journal of Applied Biomechanics*

Acceptance Date: January 31, 2019

DOI: https://doi.org/10.1123/jab.2018-0407

“Tests to Measure Core Stability in Laboratory and Field Settings: Reliability and Correlation Analyses”
by Vera-Garcia FJ, López-Plaza D, Juan-Recio C, Barbado D
Journal of Applied Biomechanics

Tests to measure core stability in laboratory and field settings: reliability and correlation analyses

Vera-Garcia, Francisco J.

López-Plaza, Diego

Juan-Recio, Casto

Barbado, David

Department of Sport Science, Sport Research Centre, Miguel Hernández University of Elche, Elche (Alicante), Spain.

Conflict of interest: None

Corresponding Address:
David Barbado, PhD.
Sport Research Centre.
Miguel Hernández University of Elche.
Avda. de la Universidad s/n. 03202 - Elche (Alicante). SPAIN.
Telephone: 966.65.25.67
E-mail: dbarbado@goumh.umh.es

Running title: Tests to measure core stability

"Tests to Measure Core Stability in Laboratory and Field Settings: Reliability and Correlation Analyses"
by Vera-Garcia FJ, López-Plaza D, Juan-Recio C, Barbado D
Journal of Applied Biomechanics

Abstract

Although core stability (CS) has been assessed through many different tests, the relationships among them are currently unknown. The main objective was to analyse the relationship between five representative tests used to assess CS in: i) laboratory settings: *Sudden Loading Test* (*SLT*) and *Stable and Unstable Sitting Test* (*SUST*); ii) field settings: *Biering-Sorensen Test* (*BST*), *Three-Plane Core Strength Test* (*TPCT*) and *Double-leg Lowering Test* (*DLLT*). The reliability of these tests was also examined. Thirty-three recreationally active males performed the tests twice. The relationship between all variables was examined using Pearson-correlation coefficient in those variables with a good reliability. Only stiffness and angular displacement in the SLT, dynamic unstable tasks in the SUST and the holding-time in the BST showed good reliability (ICC: 0.63-0.91; typical error: 9.8%-21.0%). Few and low correlations were observed between the SLT, SUST and BST. Despite finding several significant correlations among the dynamic unstable tasks of the SUST ($r \geq 0.807$; $p<0.01$), no correlations were found between the loading directions of the SLT. The absence of correlations between these tests suggests that CS measurements are not generalisable, as they probably assess different dimensions of CS, or in the case of the *BST*, a different capacity (i.e. trunk extensor endurance).

Key words: biomechanical test, field test, trunk stability, consistency

Word Count: 4342 words

Introduction

Core stability is a popular concept attracting the interest of coaches, athletes, clinicians and researchers in the last 20 years because of its potential benefits for injury prevention and athletic performance [1-4]. Consequently, many different tests have been used to assess core stability in laboratory and field settings. As there is no single accepted definition of this term [1-5], the characteristics of these tests and the parameters measured are very different, e.g. trunk/spine stiffness [6-8], participant's center of pressure (CoP) fluctuations [9-12], lumbopelvic displacement [13,14], visual/qualitative scores [2] or endurance time [15-17]. In addition, although some of these parameters could be related, the relationships between all these tests are unknown, hindering the possible generalisation of their results.

Following an operational biomechanical concept of core stability (i.e. "the capacity of the body to maintain or resume a relative position (static) or trajectory (dynamic) of the trunk following internal or external forces") [4], two laboratory methodologies have generally been used to assess core stability: i) *Sudden loading/unloading tests*, which measure the trunk's ability to respond to quick and controlled external unidirectional perturbations [6,7,9,10]; and ii) *Stable and unstable sitting tests*, which quantify the trunk's ability to keep a desired position or trajectory while coping with internal fluctuations [9-12,18]. Although both kinds of methodologies are based on the same biomechanical concept of core stability, they seem to measure different stability features or dimensions. In this sense, Barbado et al., [9] compared the performance of competitive judokas and kayakers and recreational athletes in representative tests of both methodologies founding that specific-sport training induces specific core stability adaptations, which were only revealed through specific tests. These findings suggest core stability outcomes are not generalizable, but are highly dependent on the conditions in which they are measured (i.e. dynamic/static, one-/two-dimensional, etc.) and on the characteristics of the target population (i.e. physical fitness level, sport discipline, age, health status, etc.).

Regarding field settings, many different tests have been used to assess core stability. Considering their main characteristics, they can be grouped into three different methodologies [19]: i) *Lumbopelvic postural control tests*, based on clinical concepts of spine stability/instability (e.g. "the ability to control motion of the lumbar spine and pelvis relative to an arbitrarily defined neutral position") [20] and measuring the ability to maintain a given lumbopelvic position in lying supine, such as the *Double-leg Lowering Test* [13,15,21,22] and the *Sahrmann Core Stability Test* [14,20]; ii) *Whole-body stability tests*, which follow the definition provided by Kibler et al., [2] (i.e. the ability to control the trunk motion and position to allow optimum energy transfer through the core to the limbs) and are normally performed in single leg stance, as for example, the *Three Plane Core Strength Test* and the *Star Excursion Balance Test* [2,17,21,23]; and iii) *Trunk muscle fitness tests*, which generally measure isometric trunk endurance, such as the *Biering-Sorensen Test* [17], and follow a core stability conception very close to the concept of core strength/strengthening [15,16,21]. Therefore, the main field methodologies used to assess core stability seem to measure different dimensions/components of this concept (like the biomechanical laboratory tests presented above) or even other related capabilities (e.g. endurance, strength, etc.) that would not fall within the mechanical stability definitions, thus hindering the comparison between core stability studies that use different testing methodologies. Furthermore, despite their low cost and easy application, some of these field tests have shown several methodological limitations (i.e. low reliability, poor sensitivity, etc.) [23,24], making the data interpretation even more difficult. Future studies should analyse the advantages and limitations of these tests, including their validity as stability measures [19,23].

Considering the ambiguity of the core stability term in the literature and the wide variety of tests used to measure it in laboratory and field settings [9], it's difficult to think that core stability scores obtained in specific conditions (for example, supporting the trunk in horizontal position in the *Biering-Sorensen test*) are generalisable or transferable to other conditions (for example, controlling the trunk posture or trajectory in *Stable and unstable sitting tests*). However, in scientific and practical settings all these variables are normally used under the same term (i.e. core stability), which can be confusing for coaches, physical trainers, clinicians and researchers and make the selection of core stability tests a complex and challenging decision. Therefore, further research is needed to understand the characteristics of these tests better and to explore the possible relationships between their scores.

In order to facilitate the decision-making process when selecting core stability tests, the main objective of this study was to analyse the relationship among five representative tests of some of the most common methodologies used to assess core stability in biomechanics laboratories (*Sudden Loading Test* and *Stable and Unstable Sitting Test*) and field settings (*Biering-Sorensen Test*, *Three Plane Core Strength Test* and *Double-leg Lowering Test*). In addition, in order to avoid the potential bias caused by the low consistency of the variables on the correlational analysis [25,26], the relative and absolute reliability of these tests were analysed. This reliability analysis also allowed us to discuss the advantages and disadvantages of using the biomechanical and field tests in different contexts. Based on testing specificity when measuring and training core stability [9,27], it was hypothesized that there would be no correlations among the parameters obtained in the different core stability tests, highlighting the importance of a proper selection of the most suitable tests for each individual and situation.

Methods

Participants

Thirty-three healthy recreationally active males (1-3 hours of moderate physical activity; 1-3 days per week) voluntarily took part in this study (age: 24.06±2.89 years; mass: 75.02±9.30 kg; height: 176.58±5.51 cm). Participants completed a questionnaire about their health status and physical activity habits during the year before testing. Exclusion criteria for this study were: known medical problems, especially neurological or musculoskeletal disorders and/or episodes of low back pain in the last 1 year, and participating in a trunk muscle conditioning program at the time of the study. All subjects signed an informed consent based on the 2013 Helsinki Declaration. All procedures were approved by the University Office for Research Ethics.

<u>Experimental procedure</u>

Participants participated in four testing sessions each spaced a month apart, in which they performed the five tests twice. In the first and second session they carried out the laboratory tests in this order: *Sudden Loading Test* and *Stable and Unstable Sitting Test*. In the third and fourth session they performed the three field tests in the following order: *Three Plane Core Strength Test*, *Double-leg Lowering Test* and *Biering-Sorensen Test*. The two repetitions of each test (test and retest) were separated by a month with the intention of reducing the learning effect [28]. None of the participants had previously carried out these tests, so the initial test familiarisation was the same for all participants. Before each testing session, participants performed a warm-up consisting of 5 min of cycling and 4 sets of standardised trunk exercises: 2 sets of 15 curl-ups and 2 sets of 15 back extensions in a roman chair (recovery time between sets and exercises was 30 s).

<u>Testing protocol description</u>

Although the laboratory tests analysed in this study have been comprehensively reported elsewhere [9], they are briefly described here. In the *Sudden Loading Test* participants were placed in a semi-sitting position (Figure 1A-B-C) on a stable and rigid wooden chair which limits leg motion and promotes a neutral spine position and an elastic equilibrium for the core structures. Sudden and unexpected loads were applied to the participants' upper-body centre of mass through a pulling mechanism formed by a pneumatic piston (pressure: 4.2 bar;

speed: 0.5 m/s), a steel cable tensioner and an adjustable trunk harness (Figure 1A-B-C). The pneumatic piston was placed in front, behind and at the right side of the participants to apply five sudden loads in anterior, posterior and right-lateral direction, respectively. Each perturbation took place within a 15 s window, in which participants were instructed to maintain a neutral spine position and not to react voluntarily to the perturbation. In order to keep participant's forces constant (25-27.5 N) before sudden loading, biofeedback of load-cell forces was provided in real time.

In the *Stable and Unstable Sitting Test* participants performed 10 balance tasks while sitting on a stable or an unstable seat placed on a force-plate (Kistler, Switzerland, Model 9286AA) (Figure 1D-E). The CoP displacement was measured during static and dynamic conditions (sampling frequency: 1000 Hz) to quantify trunk postural control. The stable seat was a wooden structure with leg and foot supports, while a polyester resin hemisphere (diameter: 35 cm, height: 12 cm), stuck on the bottom of the stable seat with Velcro® tape, was used for the unstable sitting tasks (Figure 1E). Participants performed two static and three dynamic 70-second tasks on each seat (with 1 min rest between tasks) [9,10]. 2D visual feedback of CoP displacement was provided to participants in real time during the dynamic and one of the static tasks (Figure 1F). In addition, during these tasks, a target point was shown to the participants to assess their ability to adjust their CoP position to the target location. As a result, the following conditions were analysed: stable sitting with and without feedback, stable sitting while performing medial-lateral, anterior-posterior and circular displacements with feedback, unstable sitting with and without feedback, and unstable sitting while performing medial-lateral, anterior-posterior and circular displacements with feedback. All conditions were counterbalanced to reduce a possible learning or fatigue effect.

During the *Three Plane Core Strength Test* [2], participants' postural control was examined in single leg stance (dominant leg) while their trunk slowly moved in the frontal

(Figure 2A), sagittal (Figure 2B) and transverse (Figures 2C and 2D) planes to lightly touch a wall located 8 cm away from the participants' shoulder/s and then returned to the starting position. After a familiarisation period (i.e. verbal instructions, visual demonstration and six practice repetitions for each plane), participants performed two trials of six repetitions for each testing direction. They were instructed to keep their head and pelvis in neutral position during the movement in the three planes. The same examiner scored the tests for all participants from 1-poor to 4-excellent following the criteria established by Chmielewski et al. [29].

As described by Krause et al., [13] to perform the *Double-leg Lowering Test* the participants were placed in a laying supine position on a semi-rigid mat with their arms on their chest and an examiner on each of their sides (Figure 2E). Examiner 1 helped participants to place their legs as close to the vertical position as possible with their knees extended and placed their fingers between their low back and the mat in order to monitor the position of the low back during the test. Participants were asked to keep their pelvis posteriorly rotated, and their lumbar spine held firm against the mat, while slowly lowering both legs from the vertical position to the horizontal position. The time execution for lowering the legs was limited to 10 s and counted aloud using a metronome. Examiner 1 verbally indicated examiner 2 when the participants' back began to lift from the monitoring fingers, which represented the end of the test. Examiner 2 recorded the participant's performance with a goniometer, which had a 40 cm–long wooden dowel at the top for placement along the axis of the femur. The goniometer remained parallel to the participants' left femur during leg lowering (Figure 2E). Two trials were performed with a 1 min rest between them, using the lower value for subsequent analyses.

Finally, in the *Biering-Sorensen Test* [30] participants were positioned in a prone position with their lower body fastened to a test bench by Velcro® inextensible tape and with their upper body off the bench (extended horizontally and unsupported), matching the anterior-superior iliac spines with the bench edge (Figure 2F). Participants were instructed to maintain their trunk cantilevered in the horizontal position for as long as possible while their arms were crossed over their chest. A digital stopwatch (Casio HS-30W-N1V) was used to record the time execution.

Data analysis and reduction

For the *Sudden Loading Test*, the stiffness, maximum angular displacement and damping of the trunk were calculated in each direction according to Cholewicki et al. [6] In order to obtain the highest reliability for these parameters, the calculations were performed for the 110 ms after the perturbation [9,10], which means that they mainly represent the combination of the passive and reflex response of the trunk following perturbation [6]. The mean of the three best trials for each direction was used for the reliability and the correlation analysis.

The CoP signal of the *Stable and Unstable Sitting Test* was filtered through a low-pass, 4th-order, zero-phase-lag Butterworth filter with a cut-off frequency of 5 Hz [31] and then downsampled at 20 Hz. To avoid the non-stationary CoP behaviour related to the beginning of the test, the signal from the first 10 s of each 70 s trial was removed from further analyses [12]. In order to quantify trunk postural control while sitting, the mean radial error was calculated as described by Barbado et al. [9] The best trial of each condition (i.e. lower mean radial error) and an unstable dynamic composite index (the averaged mean radial error of the three unstable dynamic tasks) were used for the reliability and correlational analysis.

Statistical analyses

After confirming the normality of the variables and eliminating the outliers (±3 standard deviation values), the relative reliability was analysed by the intraclass correlation coefficient for each direction ($ICC_{2,1}$) and for the composite indexes ($ICC_{2,k}$), calculating their confident limits at 90% [26]. ICC values were interpreted according to the following criteria: <0.1, trivial; 0.1-0.29, small; 0.3-0.49, moderate; 0.5-0.69, large; 0.7-0.89, very large; 0.9-1, nearly perfect

[32]. In order to detect the agreement of chance for ordinal variables used in the *Three Plane Core Strength Test*, weighted Kappa index (*k*) was estimated, with confidence limits calculated at 90%. This index was interpreted according to the following scale: 0.00 (poor); 0.01-0.20 (slight); 0.21-0.40 (fair); 0.41-0.60 (moderate); 0.61-0.80 (substantial); 0.81-1.00 (almost perfect) [33].

Absolute reliability was assessed through the typical error (TE), minimum detectable change and change in the mean. The TE was calculated dividing the difference between consecutive pairs of trials by √2 (intra-subject variability); then, the minimum detectable change was calculated as 1.5 times the TE [26]. A one-way ANOVA was performed for each test score to explore the existence of statistically significant differences in the mean between sessions.

After analysing the reliability of the test scores, the data obtained in the second testing session was used to perform a Pearson correlation analysis (*r*) between those variables which obtained an acceptable level of relative reliability [25], i.e. ICC values higher than 0.60. All analyses were performed with SPSS version 22.0 (SPSS Inc., Chicago, IL, USA).

Results

In relation to the reliability analysis (Table 1), the *Sudden Loading Test*, showed moderate to high reliability for the trunk angular displacement and trunk stiffness in most directions (0.63≤ICC≤0.91; 9.80%≤TE≤20.97%); however, the reliability of the trunk damping was low (0.25≤ICC≤0.71; 34.83%≤TE≤46.67%). No statistical differences were found for the sudden loading parameters between sessions. Concerning the *Stable and Unstable Sitting Test,* the dynamic unstable sitting tasks and their composite index showed higher reliability (0.70≤ICC≤0.81; 13.42%≤TE≤15.55%) than the static unstable and the static and dynamic stable sitting tasks (0.08≤ICC≤0.58; 12.71%≤TE≤39.15%). On the other hand, all the dynamic conditions (dynamic stable, dynamic unstable and composite index) showed significant increases in test performance between sessions. Regarding the field tests, the *Three Plane Core Strength Test* showed fair Kappa indexes ($0.26 \leq k \leq 0.29$) and TE values higher than 23.6% and the *Double-leg Lowering Test* showed a moderate ICC value (0.55), while the *Biering-Sorensen Test* showed a large ICC value (0.81) and a TE value of 12.3%, with a significant increase of the mean endurance time between both testing sessions. It should be noted that during the *Double-leg Lowering Test*, more than 75% of the sample were able to completely lower both legs until they touched the mat without pelvic anterior rotation (score=0°).

Regarding the correlation analysis (Table 2), only four significant correlations were found between the sudden loading variables of the *Sudden Loading Test*, i.e. significant negative correlations among stiffness and angular displacement in frontal (r=-0.694; p<0.01) and posterior (r=-0.857; p<0.01) loading directions, and significant positive correlations between posterior stiffness and lateral angular displacement (r=0.561; p<0.01) and between frontal and lateral angular displacement (r=0.626; p<0.01). On the contrary, for the dynamic unstable conditions of the *Stable and Unstable Sitting Test*, all the analysed correlations were significant, finding high significant positive correlations among the three dynamic unstable sitting tasks ($r \geq 0.807$; p<0.01) and between these tasks and their composite index ($r \geq 0.927$; p<0.01). Regarding the correlations between different tests, only two low significant correlations were obtained between the angular displacement after frontal and posterior loading and dynamic unstable sitting tasks. In addition, no significant correlations were found between the biomechanical parameters and the *Biering-Sorensen Test* scores.

Discussion

Despite the potential benefits of core stability for injury prevention and sport performance [1-4], there is no consensus in research and field settings about which tests should be used to measure it, maybe because it is a multidimensional and complex concept which has received many different definitions [5,9]. Therefore, in order to assist with the selection and application of core stability measures, this study analysed the reliability and the relationships among five representative tests of the most popular methodologies used to assess core stability in biomechanics laboratories and field settings. The major finding was the lack of relationships between those tests which obtained an acceptable level of reliability (i.e. ICC>0.60): *Sudden Loading Test* (trunk angular displacement and stiffness), *Stable and Unstable Sitting Test* (dynamic unstable conditions) and *Biering-Sorensen Test*. Consequently, it seems that the results obtained in one test are not generalisable to other measures of core stability, highlighting the importance of choosing the most appropriate tests for each situation based on their reliability, specificity, cost and availability. For example, the *Stable and Unstable Sitting Test* could be used in high performance settings in sports with great trunk stability demands while sitting, like kayaking or canoeing [9], and the *Biering-Sorensen Test* could be used in health and clinical settings and/or sports with great isometric back endurance demands, as hockey, gymnastics or ski descent [28].

Reliability of core stability measures

Regarding the laboratory tests, our data showed moderate to high reliability for trunk angular displacement and trunk stiffness after sudden loading (ICC≥0.63; TE≤20.97%) and for the dynamic unstable conditions of the *Stable and Unstable Sitting Test* (ICC≥0.70; TE≤15.55%). However, these 1-month reliability results were lower than those previously obtained by Barbado et al., [9] (ICC≥0.90; SEM≤14.8%) in the same tests, perhaps because they examined intra-session reliability. On the other hand, low reliability was found for the damping scores after sudden loading (ICC≥0.25; TE≤46.7%) and for the static unstable conditions and the static and dynamic stable conditions of the *Stable and Unstable Sitting Test* (ICC≥0.08;

TE≤39.1%). Interestingly, the relative reliability in the *Stable and Unstable Sitting Test* was higher for the most difficult conditions (dynamic unstable tasks and composite index), maybe because the other conditions did not represent a challenge to our participants (being recreationally active males) and therefore they were not able to discriminate between them [18]. In relation to the comparison between testing sessions, no learning effect was found for the trunk response to sudden loading in the 110 ms after perturbation, as this response depends mainly on spinal reflexes and passive trunk structures [6,34]. However, all the dynamic conditions of the *Stable and Unstable Sitting Test* showed significant increases in test performance between sessions, indicating the need of a long familiarisation period to avoid learning effect in this protocol [9,35].

Concerning the field tests, the *Three Plane Core Strength Test* showed a fair intra-rater agreement (0.26≤k≤0.29). These results are consistent with those obtained by Weir et al., [23] which found a poor inter and intra-tester reliability (0.31≤ICC≤0.55). Possibly, the narrow 4-point scale used to score the participants in this test caused most participants to score in the central values (homogenising the sample), which could affect the correlation index used to assess the relative reliability as it is sensitive to sample homogeneity [25,26]. In addition, it should be noted that these test scores may reflect whole-body stability rather than core stability as they are clearly influenced by the stance leg performance. In this sense, although core stability seems to play an important role in whole-body stability [36], this test measures postural control in single-leg stance and therefore it may be strongly affected by lower limb characteristics and capabilities (e.g. muscle strength, joint stability, leg length, etc.).

Regarding the *Double-leg Lowering Test*, the low sensitivity showed for this protocol (>75% of participants obtained a 0º score) affected its reliability and correlation analysis in this study. Although there are conflicting results in the literature [13,15,22,24], it seems that the *Double-*

leg Lowering Test sensitivity is affected by the age and physical condition of the participants. In this sense, most of our young and recreationally active male participants were able to lower both legs completely without loss of pelvic control, supporting a previous study which showed lack of sensitivity of this protocol to discriminate between individuals with high physical condition [15]. Nevertheless, the *Double-leg Lowering Test* scores obtained in the current study could also be influenced by the relatively little experience of the examiner to monitor the position of the low back with his fingers. Possibly, if we had used a sphygmomanometer to monitor the pelvic anterior rotation [37], the test would have been challenging for our participants and therefore more sensitive to discriminate between them.

In comparison to the *Three Plane Core Strength Test* and the *Double-leg Lowering Test*, the *Biering-Sorensen Test* showed a high reliability (ICC=0.80; TE=12.3%), which in general is in agreement with earlier studies [28,38,39]. It is important to notice that the mean endurance time of our participants increased with test repetition, showing the need of a familiarisation period to avoid learning effect [28,38].

Correlations between reliable core stability measures

The very few correlations found between the *Sudden Loading Test* and the *Stable and Unstable Sitting Test*, and especially between the different loading directions in the *Sudden Loading Test*, show the complexity of measuring core stability, as different biomechanical parameters seem to assess different features of this multidimensional capability [9,35]. As previously stated by Reeves et al., [5] core stability is context dependent, so the result of its assessment depends on the measurement characteristics (e.g. applied forces magnitude, direction and duration), which may involve different motor control mechanisms. In this sense, trunk performance during the dynamic tasks on the unstable seat are influenced by the feedback mechanisms of the cerebellar-cortical system [11,40]; however, trunk responses (i.e. angular displacement and stiffness) in the 110 ms after unexpected sudden perturbations mainly depend on passive trunk structures and spinal reflex responses [6,34].

Moreover, the lack of relationship between the posterior, right-lateral and anterior loading directions in the *Sudden Loading Test* suggests that trunk kinetic and kinematic responses to quick perturbations are specific to the plane being evaluated. These results are in line with a biomechanical-epidemiologic study [41] which showed that the trunk displacement after sudden force release was a significant predictor of ligament injury in female athletes when the perturbations were applied in lateral direction, but not when applied in anterior or posterior direction. In addition, Barbado et al., [9] showed that competitive judokas displayed higher stiffness after lateral loading than competitive kayakers and recreational athletes but they did not show statistical differences in anterior or posterior directions. Thus, specific sport demands may induce specific core stability adaptations in a plane of motion which are not easily transferable to others.

Regarding the *Stable and Unstable Sitting Test*, high significant correlations were found among all the dynamic unstable conditions (r>0.807; p<0.01), as they have very similar characteristics and may measure the same core stability dimension. Additionally, as a methodological consideration, these high correlations suggest this protocol could be reduced to a single task (for example the most demanding, i.e. the unstable sitting while performing circular displacements) to increase its efficiency.

The *Biering-Sorensen Test* was the only field test that obtained an acceptable level of relative reliability (ICC=0.81), and consequently the relationship between this test and the laboratory tests were analysed, finding no significant correlations between them (r≤-0.267; p>0.05). During the *Biering-Sorensen Test*, participants must maintain the trunk cantilevered in the horizontal position against gravity, which probably requires a certain level of stability. However, participants' performance in this test is quantified as the longest time they are able to maintain the horizontal position [42], which seems an endurance index rather than a stability parameter [19,28]. Overall, taking into account the limitations of the three field tests analysed in this study, it seems necessary to develop new tests to assess core stability in field settings.

As in any research, it is important to note some limitations which could bias the data presented above and their analysis. Although a sample of 25 participants has been considered a sufficient sample size in reliability studies [43], a much larger sample would be desirable for minimizing the random change for the measurements [26]. Secondly, the sample characteristics could have affected the reliability of the analysed tests because some of them were designed to evaluate patients with low back pain or instability. Therefore, our results should only be applied to a healthy, young and recreationally active male population. Future studies would need to explore the reliability and the relationships of these and other tests in other populations. Finally, although we analysed three representative tests of some of the most popular types of field methodologies used to measure core stability, there are many different field tests in the literature, so we could have chosen other tests, and therefore, have obtained different results.

Overall, the results of this study show the complexity of core stability assessment and provide information about some popular tests used to assess core stability, which may help coaches, clinicians and researchers to choose the most appropriate tests for each situation and to interpret their results.

Acknowledgments

This research was made possible by the financial support of Ministerio de Ciencia e Innovación (DEP2010-16493) and Generalitat Valenciana (ACOMP/2011/130), Spain. Casto Juan-Recio was supported by a pre-doctoral grant (Val i+d) given by Generalitat of Valencia. The authors wish to thank the participation of the University students who offered their time to take part in this research.

References

1. Borghuis J, Hof AL, Lemmink KA. The importance of sensory-motor control in providing core stability: implications for measurement and training. *Sports Med.* 2008;38(11):893-916.

2. Kibler WB, Press J, Sciascia A. The role of core stability in athletic function. *Sports Med.* 2006;36(3):189-198.

3. Vera-García FJ, Barbado D, Moreno-Pérez V, Hernández-Sánchez S, Juan-Recio C, Elvira JLL. Core stability. Concepto y aportaciones al entrenamiento y la prevención de lesiones. *Revista Andaluza de Medicina del Deporte.* 2015;8(2):79-85.

4. Zazulak B, Cholewicki J, Reeves NP. Neuromuscular control of trunk stability: clinical implications for sports injury prevention. *J Am Acad Orthop Surg.* 2008;16(9):497-505.

5. Reeves NP, Narendra KS, Cholewicki J. Spine stability: the six blind men and the elephant. *Clin Biomech.* 2007;22(3):266-274.

6. Cholewicki J, Simons AP, Radebold A. Effects of external trunk loads on lumbar spine stability. *J Biomech.* 2000;33(11):1377-1385.

7. Vera-Garcia FJ, Brown SH, Gray JR, McGill SM. Effects of different levels of torso coactivation on trunk muscular and kinematic responses to posteriorly applied sudden loads. *Clin Biomech.* 2006;21(5):443-455.

8. Vera-Garcia FJ, Elvira JL, Brown SH, McGill SM. Effects of abdominal stabilization maneuvers on the control of spine motion and stability against sudden trunk perturbations. *J Electromyogr Kinesiol.* 2007;17(5):556-567.

9. Barbado D, Barbado LC, Elvira JLL, Dieën JHv, Vera-Garcia FJ. Sports-related testing protocols are required to reveal trunk stability adaptations in high-level athletes. *Gait Posture.* 2016;49:90-96.

10. Barbado D, Lopez-Valenciano A, Juan-Recio C, Montero-Carretero C, van Dieen JH, Vera-Garcia FJ. Trunk Stability, Trunk Strength and Sport Performance Level in Judo. *PLoS One.* 2016;11(5):e0156267.

11. Cholewicki J, Polzhofer GK, Radebold A. Postural control of trunk during unstable sitting. *J Biomech.* 2000;33(12):1733-1737.

12. van Dieen JH, Koppes LL, Twisk JW. Postural sway parameters in seated balancing; their reliability and relationship with balancing performance. *Gait Posture.* 2010;31(1):42-46.

13. Krause DA, Youdas JW, Hollman JH, Smith J. Abdominal muscle performance as measured by the double leg-lowering test. *Arch Phys Med Rehabil.* 2005;86(7):1345-1348.

14. Stanton R, Reaburn PR, Humphries B. The effect of short-term Swiss ball training on core stability and running economy. *J Strength Cond Res.* 2004;18(3):522-528.

15. Leetun DT, Ireland ML, Willson JD, Ballantyne BT, Davis IM. Core stability measures as risk factors for lower extremity injury in athletes. *Med Sci Sports Exerc.* 2004;36(6):926-934.

16. Nesser TW, Huxel KC, Tincher JL, Okada T. The relationship between core stability and performance in division I football players. *J Strength Cond Res.* 2008;22(6):1750-1754.

17. Chuter VH, de Jonge XAKJ, Thompson BM, Callister R. The efficacy of a supervised and a home-based core strengthening programme in adults with poor core stability: a three-arm randomised controlled trial. *Br J Sports Med.* 2015;49(6):395-399.

18. Lee H, Granata KP. Process stationarity and reliability of trunk postural stability. *Clin Biomech.* 2008;23(6):735-742.

19. Vera-García F, Barbado D, Moreno-Pérez V, Hernández-Sánchez S, Juan-Recio C, Elvira J. Core stability: evaluación y criterios para su entrenamiento. *Revista Andaluza de Medicina del Deporte.* 2015;8(3):130-137.

20. Mills JD, Taunton JE, Mills WA. The effect of a 10-week training regimen on lumbo-pelvic stability and athletic performance in female athletes: A randomized-controlled trial. *Phys Ther Sport.* 2005;6(2):60-66.

21. Radwan A, Francis J, Green A, et al. Is there a relation between shoulder dysfunction and core instability? *Int J Sports Phys Ther.* 2014;9(1):8.

22. Sharrock C, Cropper J, Mostad J, Johnson M, Malone T. A pilot study of core stability and athletic performance: is there a relationship? *Int J Sports Phys Ther.* 2011;6(2):63-74.

23. Weir A, Darby J, Inklaar H, Koes B, Bakker E, Tol JL. Core stability: inter- and intraobserver reliability of 6 clinical tests. *Clin J Sport Med.* 2010;20(1):34-38.

24. Walker ML, Rothstein JM, Finucane SD, Lamb RL. Relationships between lumbar lordosis, pelvic tilt, and abdominal muscle performance. *Phys Ther.* 1987;67(4):512-516.

25. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* 1998;26(4):217-238.

26. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med.* 2000;30(1):1-15.

27. Aggarwal A, Kumar S, Madan R, Kumar R. Relationship among different tests of evaluating low back core stability. *J Musculoskelet Res.* 2011;14(02):1250004.

28. Juan-Recio C, López-Plaza D, Barbado Murillo D, García-Vaquero MP, Vera-García FJ. Reliability assessment and correlation analysis of 3 protocols to measure trunk muscle strength and endurance. *J Sports Sci.* 2017:1-8.

29. Chmielewski TL, Hodges MJ, Horodyski M, Bishop MD, Conrad BP, Tillman SM. Investigation of clinician agreement in evaluating movement quality during unilateral lower extremity functional tasks: a comparison of 2 rating methods. *J Orthop Sports Phys Ther.* 2007;37(3):122-129.

30. Biering-Sørensen F. Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine.* 1984; 9:106–119.

31. Lin D, Seol H, Nussbaum MA, Madigan ML. Reliability of COP-based postural sway measures and age-related differences. *Gait Posture.* 2008;28(2):337-342.

32. Hopkins W, Marshall S, Batterham A, Hanin J. Progressive statistics for studies in Sports Med and exercise science. *Med Sci Sports Exerc.* 2009;41(1):3.

33. Landis JR, Koch GG. The measurement of observer agreement for categorical data. B*iometrics.* 1977:159-174.

34. Cholewicki J, McGill KC, Shah KR, Lee AS. The effects of a three-week use of lumbosacral orthoses on trunk muscle activity and on the muscular response to trunk perturbations. *BMC Musculoskelet Disord.* 2010;11(1):154.

35. Barbado D, Moreside J, Vera-Garcia FJ. Reliability and repetition effect of the center of pressure and kinematics parameters that characterize trunk postural control during unstable sitting test. *PMR.* 2017;9(3):219-230.

36. Kang K-Y. Effects of core muscle stability training on the weight distribution and stability of the elderly. *J Phys Ther Sci.* 2015;27(10):3163-3165.

37. Ladeira CE, Hess LW, Galin BM, Fradera S, Harkness MA. Validation of an abdominal muscle strength test with dynamometry. *J Strength Cond Res.* 2005;19(4):925.

38. Juan-Recio C, Barbado D, López-Valenciano A, Vera-García FJ. Test de campo para valorar la resistencia de los músculos del tronco *Apunts. Educación física y deportes.* 2014;117.

39. McGill SM, Childs A, Liebenson C. Endurance times for low back stabilization exercises: clinical targets for testing and training from a normal database. *Arch Phys Med Rehabil.* 1999;80(8):941-944.

40. Jacobs J, Horak F. Cortical control of postural responses. *J Neural Transm.* 2007;114(10):1339-1348.

41. Zazulak BT, Hewett TE, Reeves NP, Goldberg B, Cholewicki J. Deficits in neuromuscular control of the trunk predict knee injury risk: a prospective biomechanical-epidemiologic study. *Am J Sports Med.* 2007;35(7):1123-1130.

42. Demoulin C, Vanderthommen M, Duysens C, Crielaard JM. Spinal muscle evaluation using the Sorensen test: a critical appraisal of the literature. *Joint Bone Spine.* 2006;73(1):43-50.

43. Springate SD. The effect of sample size and bias on the reliability of estimates of error: a comparative study of Dahlberg's formula. *Eur J Orthod.* Apr 2012;34(2):158-163.

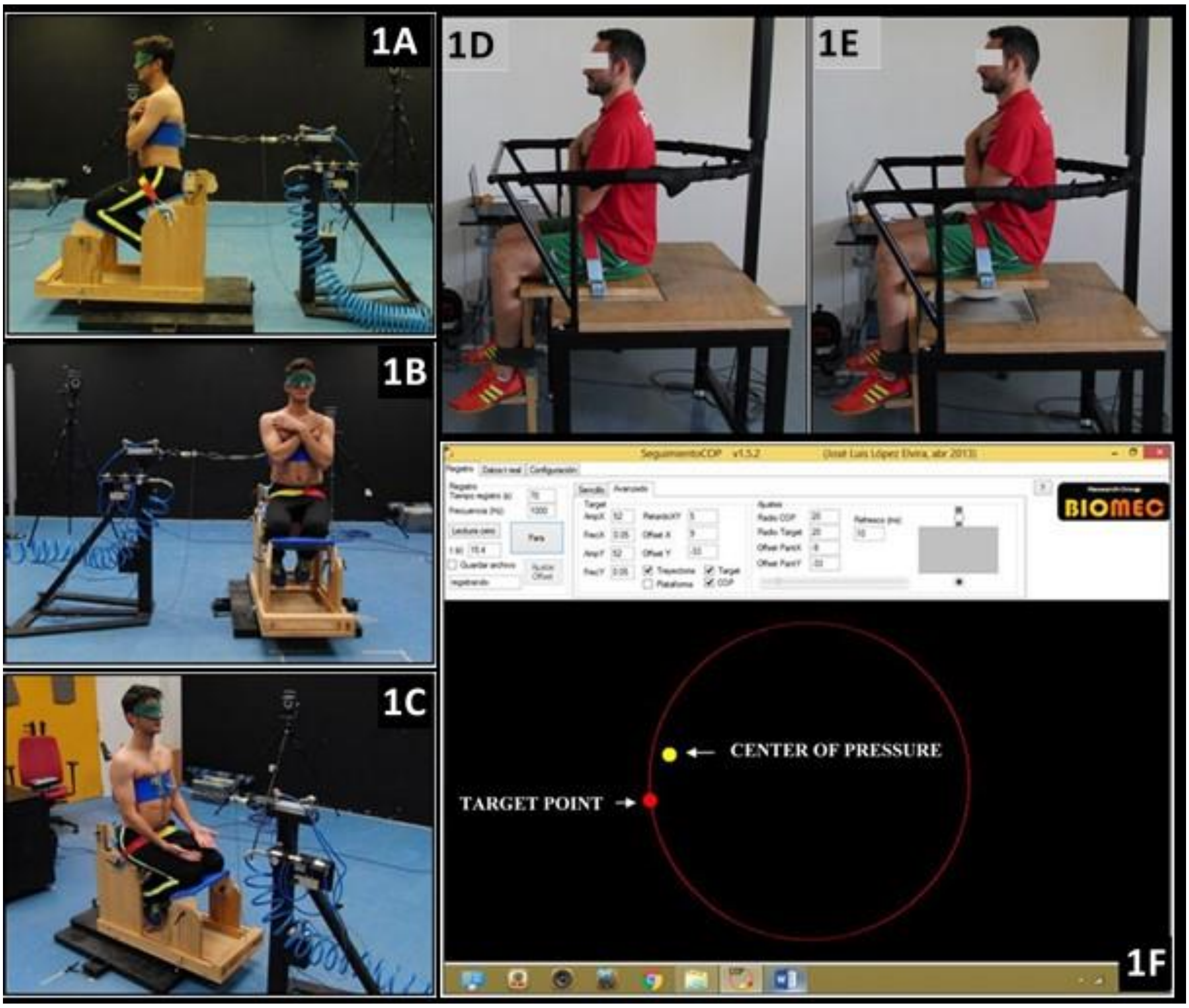

Figure 1. Pictures showing: the set-up for applying sudden loads in the posterior (1A), lateral (1B) and anterior (1C) direction; lateral view of a participant performing the *Stable and Unstable Sitting Test* on a stable seat (1D) and on an unstable seat (1E); and the software (1F) used to provide feedback to the participants in real time (unstable sitting circular task).

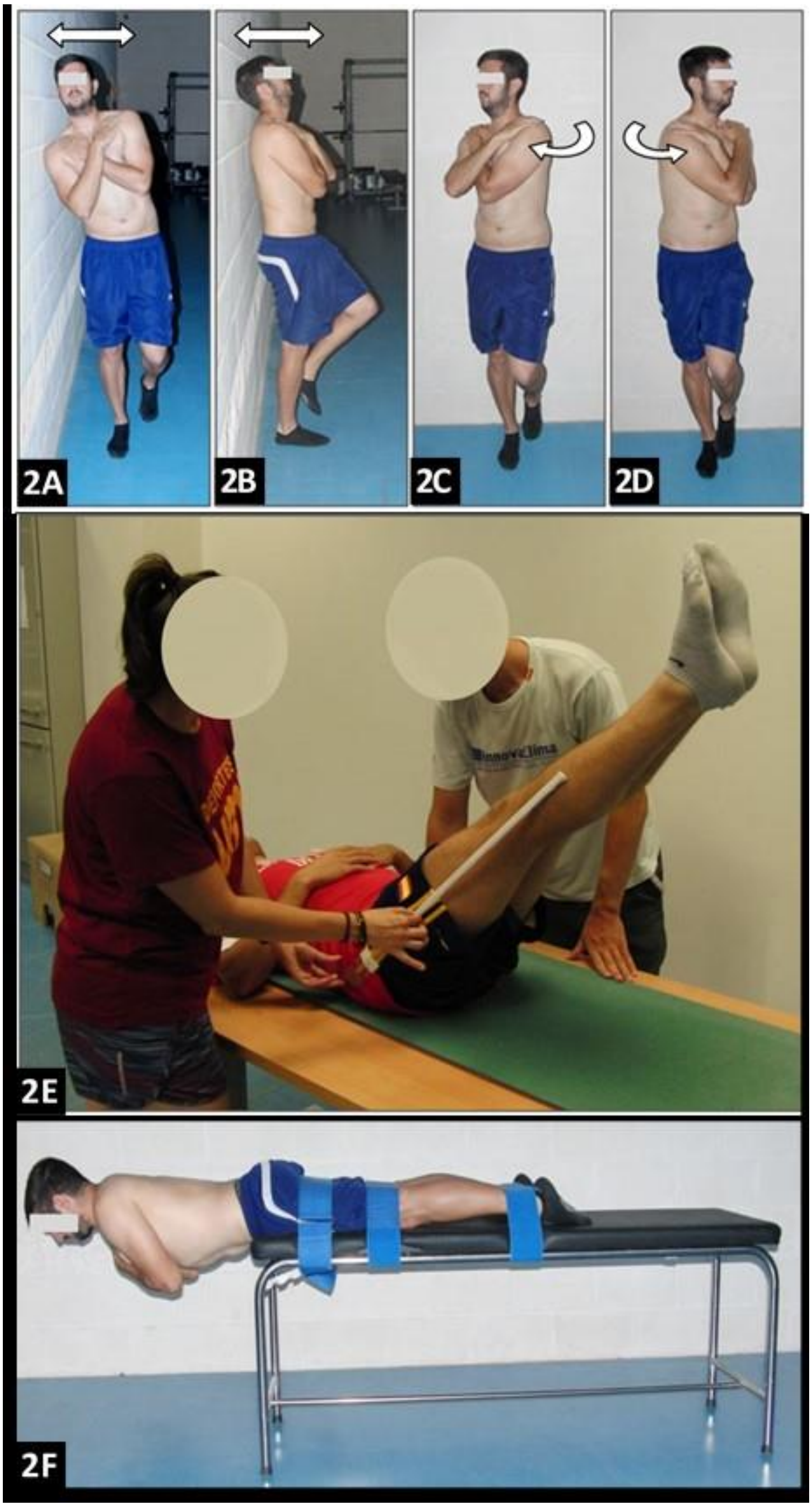

Figure 2. Pictures showing: a participant performing the *Three Plane Core Strength Test* in the frontal (2A), sagittal (2B) and transversal (2C-D) plane; a participant performing the *Double-leg Lowering Test* (2E); and a participant performing the *Biering-Sorensen Test* (2F).

"Tests to Measure Core Stability in Laboratory and Field Settings: Reliability and Correlation Analyses"
by Vera-Garcia FJ, López-Plaza D, Juan-Recio C, Barbado D
Journal of Applied Biomechanics

Table 1. Descriptive statistics and relative and absolute reliability for the variables obtained in the different tests.

Tests and variables			Session 1	Session 2	CM (mean - 90% CL)	Typical error (%) (mean - 90% LC)	MDC75 (%)	ICC(2,1) (mean - 90% LC)
Sudden Loading Test	**Θ** (rad)	**Anterior**	0.087 ± 0.022	0.088 ± 0.025	0.001 (-0.005 - 0.006)	11.46 (9.23 - 15.30)	17.18 (13.84 - 22.95)	0.87 (0.76 - 0.94)†
		Lateral	0.075 ± 0.019	0.072 ± 0.018	-0.003 (-0.010 - 0.003)	18.52 (14.92 - 24.73)	27.78 (22.37 - 37.10)	0.63 (0.28 - 0.81)†
		Posterior	0.207 ± 0.028	0.196 ± 0.025	-0.006 (-0.015 - 0.003)	9.80 (8.01 - 12.74)	14.70 (12.02 - 19.12)	0.72 (0.47 - 0.86)†
	K (N*m/rad)	**Anterior**	1499.30 ± 589.74	1538.72 ± 772.74	39.42 (-102.73 - 181.57)	19.52 (15.65 - 26.28)	29.28 (23.48 - 39.41)	0.91 (0.82 - 0.95)†
		Lateral	855.00 ± 291.34	938.97 ± 268.70	83.98 (-3.36 - 171.31)	20.97 (16.96 - 27.79)	31.45 (25.43 - 41.69)	0.71 (0.43 - 0.85)†
		Posterior	530.30 ± 135.60	571.45 ± 158.80	41.16 (4.23 - 78.09)	14.66 (11.99 - 19.07)	21.99 (17.98 - 28.60)	0.81 (0.64 - 0.90)†
	β (N*m*s/rad)	**Anterior**	362.98 ± 177.40	360.6 ± 188.11	-2.36 (-85.31 - 80.58)	46.67 (37.58 - 62.32)	70.01 (56.38 - 93.48)	0.25 (-0.48 - 0.62)†
		Lateral	712.98 ± 222.14	703.01 ± 349.80	-9.96 (-124.62 -104.70)	34.83 (28.17 - 46.17)	52.25 (42.25 - 69.26)	0.50 (0.03 - 0.74)†
		Posterior	73.69 ± 37.37	77.85 ± 32.51	8.34 (-4.46 - 21.13)	35.93 (29.38 - 46.72)	53.90 (44.07 - 70.08)	0.71 (0.44 - 0.85)†
Stable and Unstable Sitting Test	**MRE** (mm)	**SNF**	1.01 ± 0.45	1.12 ± 0.51	0.11 (-0.08 - 0.31)	39.15 (31.53 - 52.28)	58.72 (47.29 - 78.41)	0.08 (-0.27 - 0.41)
		SWF	0.76 ± 0.44	0.62 ± 0.28	-0.13 (-0.29 - 0.02)	33.47 (27.07 - 44.37)	50.21 (40.60 - 66.56)	0.42 (0.09 - 0.66)
		SML	2.18 ± 0.47	1.89 ± 0.51	-0.29 (-0.49 - -0.09)*	21.01 (17.06 - 27.66)	31.51 (25.58 - 41.49)	0.24 (-0.09 - 0.53)
		SAP	2.07 ± 0.44	1.77 ± 0.32	-0.40 (-0.62 - -0.19)*	12.71 (10.28 - 16.85)	19.07 (15.42 - 25.28)	0.57 (0.30 - 0.76)
		SCD	3.09 ± 0.72	2.51 ± 0.60	-0.58 (-0.79 - -0.36)*	16.27 (13.21 - 21.42)	24.40 (19.81 - 32.13)	0.52 (0.23 - 0.72)
		UNF	5.57 ± 1.71	4.85 ± 1.12	-0.72 (-1.28 - -0.16)	22.93 (18.61 - 30.18)	34.39 (27.92 - 45.28)	0.35 (0.02 - 0.61)
		UWF	4.98 ± 1.12	4.36 ± 1.60	-0.63 (-1.11 - -0.15)	21.87 (17.75 - 28.79)	32.80 (26.63 - 43.18)	0.58 (0.31 - 0.76)
		UML	6.57 ± 1.87	5.88 ± 1.55	-0.54 (-1.04 - -0.03)*	15.55 (12.47 - 20.93)	23.33 (18.70 - 31.40)	0.70 (0.46 - 0.84)
		UAP	7.07 ± 2.01	5.92 ± 1.66	-1.16 (-1.62 - -0.04)*	13.42 (10.85 - 17.78)	20.12 (16.27 - 26.67)	0.72 (0.51 - 0.85)
		UCD	8.55 ± 2.87	7.06 ± 2.05	-1.50 (-2.09 - -0.91)*	14.08 (11.38 - 18.66)	21.11 (17.07 - 27.99)	0.81 (0.64 - 0.90)
		Composite index	7.62 ± 2.49	6.36 ± 1.74	-1.26 (-1.74 - -0.78)*	14.08 (11.76 - 19.06)	21.72 (17.63 - 28.59)	0.79 (0.62 - 0.89)
TPCST		**Frontal**	2.04 ± 0.76	2.33 ± 0.88	0.30 (-0.07 - 0.66)	35.69 (29.18 - 46.41)	53.54 (43.78 - 6961.)	0.26 (0.02 - 0.50)¥
		Sagittal	2.56 ± 0.89	2.78 ± 0.70	0.22 (-0.07 - 0.51)	23.64 (19.33 - 30.74)	35.46 (29.00 - 46.11)	0.29 (0.04 - 0.54)¥
		Transverse	2.00 ± 0.83	2.00 ± 0.73	0.00 (-0.30 - 0.30)	32.52 (26.59 - 42.29)	48.78 (39.89 - 63.43)	0.29 (0.04 - 0.55)¥
DLLT (º)			2.19 ±7.14	3.37 ± 8.35	1.19 (-1.27 - 3.64)	**	**	0.55 (0.28 - 0.74)
BST (s)			138.56 ± 41.90	155.30 ± 49.50	16.74 (5.10 - 28.38)*	12.28 (10.00 - 16.06)	18.41 (15.00 - 24.09)	0.81 (0.66 - 0.90)

θ (rad): trunk angular displacement; *K* (N*m/rad): trunk stiffness coefficient; *β* (N*m*s/rad): trunk damping coefficient; MRE (mm): mean radial error; SNF: stable sitting without feedback; SWF: stable sitting with feedback; SML: stable sitting while performing medial-lateral displacements with feedback; SAP: stable sitting while performing anterior-posterior displacements with feedback; SCD: stable sitting while performing circular displacements with feedback; UNF: unstable sitting without feedback; UWF: unstable sitting with feedback; UML: unstable sitting while performing medial-lateral displacements with feedback; UAP: unstable sitting while performing anterior-posterior displacements with feedback; UCD: unstable sitting while performing circular displacements with feedback; TPCST: *Three Plane Core Strength Test*; DLLT (º): *Double-leg Lowering Test*; BST (s): *Biering-Sorensen Test*; CM: Change in mean; ICC: intraclass correlation coefficient; MDC: minimum detectable change; * signification p<.05; †: $ICC_{(2,k)}$; ¥: weighted kappa index; **: Values were not included due to their arbitrary reference system.

“Tests to Measure Core Stability in Laboratory and Field Settings: Reliability and Correlation Analyses”
by Vera-Garcia FJ, López-Plaza D, Juan-Recio C, Barbado D
Journal of Applied Biomechanics

Table 2. Relationship between *Sudden Loading Test*, *Stable and Unstable Sitting Test* and *Biering-Sorensen Test*.

Tests and variables			*Sudden Loading Test*						*Stable and Unstable Sitting Test*			
			Frontal		Lateral		Posterior		MRE			
			K	***θ***	***K***	***θ***	***K***	***θ***	**UML**	**UAP**	**UCD**	**Composite index**
Biering-Sorensen test			-.161	.122	-.093	.123	-.192	-.267	-.071	-.213	-.143	-143
Sudden Loading Test	**Frontal**	***K***	-	**-.694****	.163	-.282	-.093	.291	.236	.191	.176	.116
		θ		-	-.419	**.626****	.223	-.181	-.227	-.371	**-.439***	-.357
	Lateral	***K***			-	-.189	-.012	-.188	.178	.125	.180	.183
		θ				-	**.561****	-.282	.404	-.053	-.136	-.041
	Posterior	***K***					-	**-.857****	.373	.306	.200	.268
		θ						-	**-.447***	-.322	-.225	-.308
Stable and Unstable Sitting Test	**MRE**	**UML**							-	**.807****	**.850****	**.927****
		UAP								-	**.941****	**.965****
		UCD									-	**.980****
		Composite index										-

θ: trunk angular displacement; *K*: trunk stiffness coefficient; MRE: mean radial error; UML: unstable sitting while performing medial-lateral displacements with feedback; UAP: unstable sitting while performing anterior-posterior displacements with feedback; UCD: unstable sitting while performing circular displacements with feedback; Composite index: averaged MRE of the three unstable dynamic tasks; *Signification p<.05; **Signification p<.001.