

Test-Retest Reliability and Concurrent Validity of The Brief Balance Evaluation Systems Test (Brief-BESTest) in Typical Children Aged Five to Seven Years Old

Zin May Kyaw, Duangporn Suriyaamarit, Soontharee Taweetanalarp*

Department of Physical Therapy, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand, 10330

Received March 22, 2024
Accepted July 24, 2024
Published August 31, 2024

***Corresponding author:**

Soontharee Taweetanalarp
Department of Physical Therapy, Faculty
of Allied Health Sciences, Chulalongkorn
University, Bangkok, Thailand, 10330

E-mail:
soontharee.t@chula.ac.th

©2024, Kyaw, ZM et al.

This is an Open Access article published
under a Creative Commons (CC BY-NC-
ND 0.4) license.



ABSTRACT

Introduction: The brief balance evaluation systems test (Brief-BESTest) is a balance assessment tool consisting of eight items that assess the six components of the postural control development system. It is a valid and reliable test for assessing balance for various physical conditions in different populations; however, it has not been investigated in typical school-aged children. The study aimed to establish the reliability and validity of the Brief-BESTest in children aged five to seven years old.

Methods: A total of 26 children (13 boys, 13 girls, age = 76.65 ± 9.63 months) were included, and their balances were assessed by the Brief-BESTest and pediatric balance scales (PBS). The same rater completed the test-retest of the Brief-BEST test one week later. The intraclass correlation coefficient (ICC) models 3,1 were used to investigate test-retest reliability. The Spearman correlation coefficient was used to determine the correlation between the Brief-BESTest and PBS. The standard error of measurement (SEM) and minimal detectable change (MDC) were also investigated. Differences were considered statistically significant when the p-value was less than 0.05.

Results: The test-retest reliability of the Brief-BESTest results ICC (3,1) values of 0.94 (95%CI = 0.86 - 0.97) with the p-value of < 0.001 . The correlation between the Brief-BEST test scores and the PBS scores was the Spearman's rho values of 0.509 with significance p values of 0.008. The SEM and MDC were 0.72 and 1.9,9, respectively.

Conclusion: Brief-BESTest has excellent test-retest reliability and moderate concurrent validity when used in typical children aged five to seven years old.

Keywords: Postural control; Brief-BESTest; Reliability; Validity; Typical children

Introduction

The ability to maintain the body's position in space while maintaining stability and orientation during a task is known as postural control [1]. Postural control develops throughout life and is related to motor milestones such as head control or lifting, sitting, creeping, crawling, standing, walking, and running [1, 2]. Infants coordinate sensory and motor input between the head and trunk, extending the sensory-motor principles of head postural control to the trunk muscles [2].

They learn to sit by stimulating the postural responses with the somatosensory and vestibular systems and use their postural muscles between the ages

of one and six months [3-5]. Until age three, the trunk serves as the major postural reference, consistent with egocentric orientation control [2, 6]. Some proprioceptive systems mature between the ages of three and four years. Visual information is the first sensory system to emerge and fully mature just by age five [7]. At the age of seven, visual and somatosensory inputs are fully developed, and the structures responsible for motor control are fully developed and have generally reached mature postural control [7, 8]. The entire postural control system is fully developed at the age of 12 because the vestibular system is only fully matured at that age [7]. Therefore, it can be said that the postural control system is not completely developed between the

ages of five and seven, and thus, it is necessary for early detection.

The postural control development of the children can be assessed with various balance assessment tools, which are commonly used in clinical settings. These tools include the Timed Up and Go Test (TUG), which measures the time taken to stand up, walk, turn, and sit down, providing a quick assessment of basic mobility [9, 10], Pediatric Balance Scale Test (PBS), which evaluates static and dynamic balance through 14 tasks but can be time-consuming [11-13], Pediatric Reach Test (PRT), which assesses forward reach while maintaining balance, focusing on dynamic aspects of balance [14, 15], Mini Balance Evaluation Systems Test (Mini-BESTest), which includes 14 tasks derived from the original BESTest, providing a comprehensive assessment of dynamic balance across four components of the balance systems but requiring a longer administration time [16, 17] and brief balance evaluation systems test (Brief-BESTest) [18-20].

Among these tests, Brief-BESTest was developed in 2012 by Padgett et al. based on the original balance evaluation systems test (BESTest). It comprises eight items with the highest item-section correlation, representing the six postural control components [20, 21]. Previous research indicates that the Brief-BESTest was utilized to evaluate balance impairment in different populations with various conditions, including multiple sclerosis, Parkinson's disease [20, 21], stroke, community-dwelling older individuals [22], children with cerebral palsy and youth with visual impairments [18, 19]. The reliability of the Brief-BESTest was excellent, with the intraclass correlation coefficient (ICC) with 95% confident interval (CI) values of 0.94 (0.98 - 0.99), which matched the interrater reliability levels of the original BESTest [18-22]. The Brief-BESTest showed a strong correlation (Spearman's $\rho = 0.93$) with the PBS [18], a 14-item balance assessment tool, which is a gold standard commonly used in clinical settings [11, 12]. When comparing the Brief-BESTest to the Mini-BESTest, which derives from the original BESTest, the Brief-BESTest is more practical and particularly advantageous in clinical settings where time constraints are a concern. The Brief-BESTest covers six components of the postural control systems, while the Mini-BESTest covers only four. This makes the Brief-BESTest less time-consuming while still maintaining comprehensive assessment capabilities. Overall, Brief-BESTest is easy to use and less time-consuming, combines strong validity and reliability, and helps evaluate functional balancing performance [18, 19].

However, there were fewer studies of the Brief-BESTest in the children population [18, 19], and the measurement properties of the test have been established only for atypical children such as those with cerebral palsy [18] and children and adolescents with visual impairment [19]. Furthermore, it is not

appropriate to reference the properties of the Brief-BESTest scores in other populations [23], and the validity and reliability of the Brief-BESTest in typical school-aged children have not been demonstrated. Further research was necessary to determine the psychometric properties of the Brief-BESTest, particularly in typical school-aged children. Additionally, the whole postural control development system is not fully mature in children ages five and seven. Therefore, this study aims to establish the concurrent validity and test-retest reliability of the Brief-BESTest in typical school-aged children between the ages of five and seven. Based on earlier research and clinical expectations, we hypothesized that the Brief-BESTest would exhibit good test-retest reliability and concurrent solid validity when compared to established measures such as PBS. This study aims to provide detailed points for these measurement properties in typical school-aged children.

Methods

Study designs and participants

An observational study design was used to investigate the test-retest reliability and concurrent validity of the Brief-BESTest. A random sample of children aged 5-7 years old was recruited for this study. Participants were included if they were between five and seven and could understand and follow instructions in the Myanmar Language. The study excluded participants who had undergone neurologic or orthopedic surgery within the previous 12 months or whose intellectual or behavioral limitations prevented them from fully participating in the tests. Furthermore, participants who used mobility assistance equipment or had any medicine with a sedative effect within 24 hours of testing and could not finish the test were also excluded. The characteristics of the children are shown in Table 1. Before collecting data, the parents of potential participants were given a consent form that included the purpose of the study and the research protocol, and they consented to participate.

Procedures

Data collection was done by the physiotherapist, who was trained to use the Brief-BESTest and PBS testing procedures before starting the data collection. The training involved standardized instruction sessions and practice assessments to ensure consistency. The participants were asked to perform the items as instructed, with the items taken randomly from the Brief-BESTest assessment form. The items in the Brief-BEST assessment form are (i) hip/trunk lateral strength for biomechanical constraints, (ii) functional reach forward for stability limits, (iii) standing on one leg left and right for transitions-anticipatory postural control, (iv) compensatory stepping lateral, left and suitable for the reactive postural response, (v) stance with eyes closed on foam surface for sensory

orientation, and (vi) timed up and go test for gait stability. The score is rated 0 to 3, with 3 reflecting no balance impairment and 0 representing severe balance impairment, for a maximum score of 24 [20, 21]. After the Brief-BESTest, the PBS assessment was performed and scored with the PBS assessment form with standardized instructions, including 14 items related to static and dynamic balance, which are sitting to standing, standing to sit, transfers, standing unsupported, sitting unsupported, standing with eyes closed, standing with feet together, standing with one foot in front, standing on one foot, turning 360 degrees, turning to look behind, retrieving an object from the floor, placing alternate foot on a stool and reaching forward with an outstretched arm [11]. The scores for the items range from 0 to 4 on a 5-point ordinal scale (0-1-2-3-4) with a maximum score of 56 points; a higher number indicates better performance in terms of functional balance [11]. After 7-day intervals, the participants were evaluated with the Brief-BESTest by the same assessor at the same place for the test-retest reliability.

Statistical analysis

Data analysis was performed using SPSS version 28.0 software. Shapiro-Wilk analyzed the normality of the Brief-BESTest and PBS data and the result showed the data were not normally distributed. Therefore, the concurrent validity was analyzed using the Spearman correlation coefficient to evaluate correlations between the Brief-BESTest and PBS, with the rho value of 0.70 or higher generally considered to show a strong relationship. In contrast, a coefficient between 0.40 and 0.70 indicated a moderate relationship, and below 0.40 was a weak relationship [24]. The p-value < 0.05 was considered statistically significant.

Test-retest reliability of the Brief-BESTest total scores in typically developing children is calculated using the ICC model (3,1) values of 95% CI. ICC values are interpreted whereby below 0.50 is considered poor reliability, between 0.50 and 0.75 is moderate reliability, between 0.75 and 0.90 is good reliability, and above 0.90 is excellent reliability [24]. Additionally, absolute reliability was determined by calculating the standard error of measurement (SEM) and minimal detectable change (MDC) using the equations

$$\text{SEM} = \text{Standard Deviation (SD)}_{\text{pooled}} \times \sqrt{(1-\text{ICC})}$$

$$\text{and MDC} = 1.96 \times \sqrt{2} \times \text{SEM}$$

Ethical approval

This study proposal was approved by the Institutional Review Board of the Research Ethics Review Committee for Research Involving Human Subjects, Chulalongkorn University, with approval number 215/66 on October 20, 2023.

Results

The mean scores of the Brief-BESTest, along with SD values, are shown in Table 2. The test-retest reliability of the Brief-BESTest results in significant-excellent test-retest reliability with the ICC (3,1) values of 0.94 (95% CI = 0.86 - 0.97) with the p-value of < 0.001. This high ICC indicates a very high level of consistency in the test scores over repeated administrations, suggesting that the Brief-BESTest is a reliable tool for assessing balance in typical school-aged children. The SEM and MDC of the Brief-BESTest test-retest reliability were 0.72 and 1.99, respectively (Table 2). The SEM indicates the precision of individual test scores; a lower SEM reflects higher precision. The MDC represents the smallest change that exceeds measurement error, thus indicating a real change in balance performance if the score changes by at least 1.99 points. The score of the PBS is 54.46 ± 1.17 (Mean $\pm$ SD). The correlation between the Brief-BESTest test scores and the PBS scores was the Spearman's rho values of 0.51 with significance p-values of 0.008, which means that there was moderate concurrent validity of the Brief-BESTest, as it aligns reasonably well with another established standardized measure of balance, the PBS.

Table 1 Demographic data of the participants (n=26)

Demographic Data	n (%)	Mean $\pm$ SD
Gender	26	-
Male	13 (50%)	-
Female	13 (50%)	-
Age (months)	-	76.65 $\pm$ 9.63
Weight (kg)	-	21.76 $\pm$ 5.83
Height (cm)	-	117.12 $\pm$ 7.54
BMI (kg/m²)	-	15.66 $\pm$ 2.79
Arm length (cm)	-	49.65 $\pm$ 3.42
Leg length (cm)	-	59.61 $\pm$ 4.32

Note: n (%) = number of the participants (percentage of the participants)

Table 2 Outcome values of the test-retest reliability of the Brief-BESTest in school-aged children

Outcome	Mean Scores (SD)		ICC _(3,1)	95% CI	p-value	SEM	MDC
	Baseline	After 7 Days					
Brief-BESTest	19.96 ± 3.27	19.69 ± 2.77	0.94	0.86 - 0.97	< 0.001	0.72	1.99

Note: ICC_(3,1)= Intraclass correlation coefficient model (3,1), 95% CI= 95% confidence interval, p-value= significant level of ≤ 0.05 , SEM= standard error of measurement, MDC= minimal detectable change

Discussion

The purpose of this study was to determine the Brief-BESTest's concurrent validity and test-retest reliability in children with typical development. Our study found that the Brief-BESTest had high test-retest reliability, with low SEM and MDC values, and moderate concurrent validity when employed in typically developing children.

Our result showed excellent reliability since the current study used the ICC model (3,1) and involved test-retests with the same rater [24], who practiced before starting the data collection, and the Brief-BESTest instructions were standardized [20]. The consistency of the test administration might also be one of the reasons why our results were excellent, as the rater tested at the same place, with the same environment, with the same standardized procedure, and with the same time interval for the retest. Furthermore, this study focused on children between the ages of five and seven who have immature and varied postural control, and the results showed high reliability, which is consistent with the findings from the studies in children with cerebral palsy [18] and neurological patients [20, 25, 26]. In the stroke population, there was likewise excellent reliability (ICC(2,1)= 0.97-0.97, $p < 0.001$) [25]. ICC(3,1) values of 0.98 with 95% CI 0.95-0.99 were observed in the study involving children with cerebral palsy, indicating strong intra-rater reliability [18]. Excellent reliability suggests that the Brief-BESTest can be a reliable tool in clinical settings for evaluating balance in children with typical development. High reliability is crucial as it ensures that the test results are consistent over time, making it a dependable tool for tracking balance performance and guiding therapeutic interventions.

For the concurrent validity, the PBS and the Brief-BESTest showed a moderate association with a Spearman's rho of 0.51, indicating the Brief-BESTest has moderate concurrent validity. There can be various reasons, as both assessments focus on four aspects of postural control: biomechanical restrictions, stability limitations, transitions for anticipatory postural adjustment, and sensory orientation. However, some aspects of postural control, like gait stability and the

time up-and-go test, can be better reflected by the Brief-BESTest than by the PBS [18]. Moreover, the children got full scores in most of the PBS items, but in some of the Brief-BESTest items, they did not get the full scores; it might be because of the structures of the items; for example, in the item for sensory orientation, of the Brief-BESTest, "stand with eyes closed on the foam surfaces" for 30 seconds is the full score, but for the PBS, the item is only to stand with eyes closed and only for 10 seconds for the full score. Additionally, a few of the children who participated in our research had varied body weights—that is, some were underweight and some overweight. Our study differs slightly from earlier findings; the Brief-BESTest and PBS showed a good relationship [18, 27]. Strong correlations with Berge Balance measures (Spearman's rho = 0.87, p -value < 0.001) were found in the stroke patient trial, indicating good concurrent validity [25]. However, in a study of children, adolescents, and youth with visual impairments, the correlation between the Brief-BESTest and activities-specific balance confidence scale (ABC) scores was slightly lower than in other populations [19, 21].

Understanding the error of a measurement tool is essential for determining its application for therapeutic decision-making [24]. Previous studies discovered that the SEM was 1.55 and 1.40 in patients with neurological problems [26], as well as in elderly people living in nursing homes [28], which indicates less precision compared to our findings. The SEM in our study was 0.72, which is satisfactory for the Brief-BESTest's absolute reliability. Furthermore, the test-retest reliability of the Brief-BESTest is precisely excellent when used in the typical children population because of the indication for MDC of 1.99. The lower the MDC, the more excellent the reliability [24]. The MDC values of the Brief-BESTest were 3.19, 4.00, and 4.30 in patients undergoing complete knee replacement [27], elderly residents of nursing homes [28], and individuals suffering from neurological conditions [26], respectively, suggesting less ability to identify a noticeable change in the overall scores of the test [27, 28].

There were some limitations in our study. The first one is age limitation since it only included children who were normally developing between the ages of 5 and 7 years. Secondly, the generalizability of our study cannot be generalized to the overall child population worldwide as the study was conducted in the Southeast Asia child population, and the lifestyle and nurturing style is different. Therefore, the scores for the Brief-BESTest for other regions might be different from those of our study. Lastly, we didn't implement specific controls to ensure participants maintained the same physical status during the 7-day interval between tests for the test-retest. Nonetheless, the clinical use of Brief-BESTest is still possible for evaluating the postural aspects of children's balance development procedures. Future research is needed to ensure that the Brief-BESTest is reliable and valid for application in a wide range of child populations, incorporating stricter controls or monitoring to ensure the consistency in participants' physical activity levels during the study period.

Conclusion

The Brief-BESTest has acceptable concurrent validity and excellent test-retest reliability. Therefore, this increases the confidence in using this test in clinical settings. The high reliability of the test ensures consistent results, providing confidence in its use for evaluating postural control in young children. Additionally, including all six components of the postural control system in the Brief-BESTest makes it a comprehensive tool for identifying balance issues and informing targeted intervention strategies for the impacted components. This makes the Brief-BESTest a valuable resource for clinicians aiming to assess and improve balance in the child population.

Conflicts of interests

The authors affirm that there were no conflicts of interest in this study.

Acknowledgments

The authors would like to thank all the participant guardians and participants who volunteered to participate in this study.

References

- [1] Shumway-Cook A, Woollacott MH, Rachwani J, Santamaria V. Motor control: translating research into clinical practice. 6th ed. Philadelphia: Wolters Kluwer; 2023.
- [2] Assaiante C, Mallau S, Viel S, Jover M, Schemitz C. Development of postural control in healthy children: a functional approach. *Neural Plast.* 2005;12(2-3):109-18.
- [3] de Graaf-Peters VB, Blauw-Hospers CH, Dirks T, Bakker H, Bos AF, Hadders-Algra M. Development of postural control in typically developing children and children with cerebral palsy: possibilities for

intervention? *Neurosci Biobehav Rev.* 2007;31(8):1191-200.

[4] Hedberg Å, Carlberg EB, Forssberg H, Hadders-Algra M. Development of postural adjustments in sitting position during the first half year of life. *Dev Med Child Neurol.* 2005;47 (5):312-20.

[5] Woollacott M, Debû B, Mowatt M. Neuromuscular posture control in the infant and child: is vision dominant? *J Mot Behav.* 1987;19(2):167-86.

[6] Mansion J. Postural control systems in developmental perspective. *Neurosci Biobehav Rev.* 1998;22(4):465-72.

[7] Cardosodesa CS, Boffino CC, Ramos RT. Development of postural control and maturation of sensory systems in children of different ages a cross-sectional study. *Braz Jr Phys Ther.* 2018;22(1):70-6.

[8] Ludwig O, Kelm J, Hammes A, Schmitt E, Frohlich M. Neuromuscular performance of balance and posture control in childhood and adolescence. *Heliyon.* 2020;6(7):e04541.

[9] Hallemans A, Klingels K, Crieckinge TV, Vereeck L, Verbecque E. Reliability and concurrent validity of a modified Timed Up and Go test for healthy preschoolers. *Eur J Pediatr.* 2020;179(10):1579-86.

[10] Verbecque E, Vereeck L, Boudewyns A, deHeyning PV, Hallemans A. A modified version of the Timed Up and Go test for children who are preschoolers. *Pediatr Phys Ther.* 2016;28(4):409-15.

[11] Franjoine MR, Gunther JS, Taylor M. Pediatric balance scale: a modified version of the Berg Balance Scale for the school-age child with mild to moderate motor impairment. *Pediatr Phys Ther.* 2003;15(2):114-28.

[12] Kolehmainen N, Stodden D, Smith AL, Larkin D. Validity and reliability of the Pediatric Balance Scale in typically developing children. *Arch Phys Med Rehabil.* 2018;99(6):1076-82.

[13] Franjoine MR, Darr N, Held SL, Kott K, Young BL. The performance of children developing typically on the Pediatric Balance Scale. *Pediatr Phys Ther.* 2010;22(4):350-9.

[14] Norris RA, Wilder E, Norton J. The functional reach test in 3- to 5-year-old children without disabilities. *Pediatr Phys Ther.* 2008;20(1):47-52.

[15] Deshmukh AA, Ganesan S, Tedla J. Normal values of functional reach and lateral reach tests in Indian school children. *Pediatr Phys Ther* 2011;23(1):23-30.

[16] Dewar R, Claus AP, Tucker K, Ware R, Johnston LM. Reproducibility of the Balance Evaluation Systems Test (BESTest) and the Mini-BESTest in school-aged children. *Gait Posture.* 2017;55:68-74.

- [17] Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: the Mini-BESTest. *J Rehabil Med*. 2010;42(4):323-31.
- [18] Nakkeaw C, Patite A, Sukkeawnareng K, Soonthoanon K. Reliability and validity of the Brief-Balance Evaluation Systems Test (Brief-BESTest) in children with cerebral palsy. *Thai J Phys Ther*. 2020;42(1): 1-11.
- [19] Pennell A, Fisher J, Patey M, Miedema ST, Stodden D, Lieberman L, et al. Measurement properties of Brief-BESTest scores from children, adolescents, and youth with visual impairments. *Disabil Rehabil*. 2023:1-10.
- [20] Padgett PK, Jacobs JV, Kasser SL. Is the BESTest at its best? A suggested brief version based on interrater reliability, validity, internal consistency, and theoretical construct. *Phys Ther Rehabil J*. 2012;92(9): 1197-207.
- [21] Godi M, Giardini M, Arcolin I, Ferrante S, Nardone A, Corna S, et al. Is the Brief-BESTest brief enough? suggested modifications based on structural validity and internal consistency. *Phys Ther Rehabil J*. 2019;99(11):1562-73.
- [22] Marques A, Almeida S, Carvalho J, Cruz J, Oliveira A, Jacome C. Reliability, validity, and ability to identify fall status of the Balance Evaluation Systems Test, Mini-Balance Evaluation Systems Test, and Brief-Balance Evaluation Systems Test in older people living in the community. *Arch Phys Med Rehabil*. 2016;97(12):2166-73.
- [23] American Educational Research Association. Standards for educational and psychological testing. Washington, USA: American Educational Research Association; 2014.
- [24] Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 4th ed. Philadelphia: F.A. Davis Company; 2009.
- [25] Huang M, Pang MYC. Psychometric properties of Brief-Balance Evaluation Systems Test (Brief-BESTest) in evaluating balance performance in individuals with chronic stroke. *Brain Behav*. 2017;7(3):e00649.
- [26] Bravini E, Nardone A, Godi M, Guglielmetti S, Franchignoni F, Giordano A. Does the Brief-BESTest meet classical test theory and Rasch analysis requirements for balance assessment in people with neurological disorders? *Phys Ther*. 2016;96(10):1610-9.
- [27] Chan AC, Pang MYC. Assessing balance function in patients with total knee arthroplasty. *Phys Ther*. 2015;95(10):1397-407.
- [28] Viveiro LA, Gomes GC, Bacha JM, Junior NC, Kallas ME, Reis M, et al. Reliability, validity, and ability to identify fall status of the Berg Balance Scale, Balance Evaluation Systems Test (BESTest), Mini-BESTest, and Brief-BESTest in older adults who live in nursing homes. *Geriatr Phys Ther*. 2019;42(4):E45-54.