IJSRM
INTERNATIONAL JOURNAL OF SCIENCE AND RESEARCH METHODOLOGY
An Official Publication of Human Journals

Human Journals
Research Article
March 2017 Vol.:6, Issue: 1

Cross-Cultural Adaptation and Validation of the Bestest Version in Portuguese Language and Evaluation of Reliability

IJSRM
INTERNATIONAL JOURNAL OF SCIENCE AND RESEARCH METHODOLOGY
An Official Publication of Human Journals

Ricardo Pedro*, Patrícia Alexandra Pousinho Moita, Marta Antunes, Mariana Sousa Faria, Diogo Campos

Escola Superior de Saúde de Cruz Vermelha Portuguesa

Submission: 3 March 2017
Accepted: 7 March 2017
Published: 25 March 2017

Keywords: Scale; Balance; Evaluation; BESTest; Adaptation; Validation; Reliability

ABSTRACT

Background: Balance deficits are one of the most common problems in neurorehabilitation, but the most commonly used balance assessment tool in physical therapy, the BESTest (Balance Evaluation Systems Test) has not yet been translated for use in Portugal. Thus, the cultural and language adaptation and validation of the BESTest will be valuable to physiotherapy. **Objective:** Translate and validate the BESTest to the Portuguese language and to evaluate the reproducibility so it can be used in Portugal. **Methods:** five phases of the translation and adaptation process were used: forward translation, synthesis, back translation, expert committee review and pre-test. The pre-test sample included 19 individuals (13 females and 6 male). **Results:** Most of the modification proposals by the expert panel were of linguistic nature (172 modifications corresponding to 90 items), followed by content modifications (41 modifications corresponding to 37 items) and modifications on the scale subjectivity (8 modifications corresponding to 8 items). The pre-test results consisted of four dimensions: sample characterization data, scale application, questionnaire filled in by each sample participant and by the rate assigned to each participant. The results of the pre-test questionnaire showed that the BESTest was clear and perceivable. The BESTest Portuguese version obtained an inter-rater reliability of α=0,964. However, with the distribution by sections, this value (α) ranged between 0,788 and 0,948. Concerning intra-rater reliability, *Cronbach's* alpha was 0,964, ranging between 0,782 and 0,949 in the five scale sections. Finally, *Cronbach's* alpha value concerning the internal consistency within sections ranged between 0,80 and 0,95. **Conclusion:** The cross-cultural adaptation of BESTest was successful, furthermore, the results are positive concerning to content validity of BESTest and with promising measurement properties.

INTRODUCTION

Cross-cultural adaptations should be considered for several different settings[1]. Balance is a condition of extreme importance within functional mobility and self-sufficiency throughout life, allowing normal performance of activities[2-5].

Balance control allows adjustment of the gravity vector to the base of support during any voluntary motor activity[3-8]. At the biomechanical level, it concerns the capacity of maintaining the center of mass within the stability limits determined by the base of support. In the context of a balance disorder, it can be classified according to the general body stance and/or neuromuscular transmission[2,5, 9-13]. This concept has developed in a way to encompass static, dynamic and reactive components[2,5].

In terms of mobility and total functional independence, balance deficit is a frequent problem with which the physical therapist deals daily, representing in most cases neuromuscular and geriatric conditions. Thus stressing out the evaluation, through assessment tools able to differentiate the systems' contribution to balance problems and fall risk[7,11,14-19]. This evaluation is sometimes complex and consists of several tests in order to reach some specificity, therefore a complete assessment tool is needed in order to simplify it[2, 15].

This study aims to achieve the cultural adaptation and validation of the BESTest version in Portuguese language, in order to apply it in studies or assessments of patients with balance disorders, as well as to evaluate its reliability and internal consistency.

Balance Evaluation Systems Test

The Balance Evaluation Systems Test (BESTest) is a sensitive and specific balance assessment tool developed to identify the postural control systems responsible for functional balance deficit, thus allowing the specific evaluation of systems in failure[3, 4, 6, 16, 17, 20].

Therefore, in order to focus on the intervention and to obtain an accurate measurement of results, it is necessary to evaluate primarily the systems interacting and enabling normal balance, identify the affected ones and perceive them as a whole[3, 9, 15, 21, 22]. This assessment system was developed in 2008, aiming to evaluate certain components affecting balance. The BESTest includes 36 items grouped into six systems: biomechanical constraints, stability limits, anticipatory postural adjustments, reactive postural responses, sensory orientation and

walking stability, comprising a maximum score of 108 points, which represents the normal balance[3, 4, 9, 17, 20, 22-24].

BESTest versions already validated are available in the following languages: French (BESTest), Spanish (BESTest), Swedish (BESTest), Greek (Mini BESTest), Japanese (Mini BESTest and BESTest), Persian (Mini BESTest and BESTest) and Brazilian Portuguese (Mini BESTest and BESTest)[25].

METHODOLOGY

Cultural and Linguistic Adaptation

The cultural and linguistic adaptation procedure followed the sequential method based on the recommendations of Beaton *et al*[1]. The BESTest original version was translated by two independent translators having Portuguese as mother language, one with and the other one without knowledge in the health domain[26-28]. The Portuguese consensus version has considered the two translations and the respective translators' reports, as well as the modifications suggested by an expert panel consisting of seven Physical therapists[27,29].

The back-translation was performed by two independent translators, having English as mother tongue and without knowledge in the health domain, including a report on the translation specificities. These documents led to the consensus version, back-translated, which was validated by the author of the original version[26-28].

The pre-test was applied by three raters to a sample of 19 healthy young individuals (13 females and 6 male), in order to evaluate the understanding of the questions and the adequacy to the Portuguese context. After the application of the assessment tool, the individuals (participants and raters) answered a questionnaire concerning their perception and difficulty in performing the assessment.

Inter- and intra-rater reliability and internal consistency

The final Portuguese version was applied to a sample of 35 individuals (19 females and 16 male) selected according to accessibility convenience and complying with the following inclusion criteria: Portuguese nationality, aged 18 or older, with pathology or clinical condition including disturbed balance and followed by physical therapists in their clinical practice place. The exclusion criteria were: to be confined to bed and revealing hemodynamic

instability. Three individuals were excluded: two of them being discharged from the hospital and the third one due to aggravation of his disease. The assessment tool was applied in two moments with an interval of eight days.

The internal consistency of the results and the intra-rater reliability of the BESTest were evaluated by verifying the intra-class correlation coefficient through *Cronbach's* alpha. The inter-rater reliability was verified through the intra-class correlation coefficient (ICC). Data obtained were processed using IBM SPSS Statistics 21 program.

All participants signed consent forms allowing their voluntary participation in the research, within both phases of the tests. All doubts concerning research goals, as well as procedure options, were clarified.

RESULTS

Cultural and Linguistic Adaptation

The expert panel proposed modifications which were scored as "I agree with modifications", corresponding to 110 items, and scored as "I do not agree", corresponding to 12 items. Most of the modification proposals were of linguistic nature (172 modifications corresponding to 90 items), followed by content modifications (41 modifications corresponding to 37 items) and modifications on the scale subjectivity (8 modifications corresponding to 8 items). There were no proposals at the scientific level (Table 1).

Pre-test results

The pre-test results consisted of four dimensions: sample characterization data, scale application, questionnaire filled in by each sample participant and questionnaire filled in by the rate assigned to each participant. The pre-test sample included 19 individuals (13 females and 6 male) with a mean age of 22±3 ([20-25]). All sample individuals were higher education students, without relevant clinical antecedents (Table 2). The mean pre-test time was of 17,68 minutes, (15 minutes - 25 minutes). The mean score obtained was 101 points (85 points - 107 points) (Table 3). The results of the pre-test questionnaire proved that the BESTest was clear and perceivable.

Sample results

The sample selected for application of the BESTest Portuguese version included 32 individuals (16 females and 16 male) with a mean age of 53, 97±14,98. It encompassed different diagnosis conditions, including low back pain, tarsal fracture, spinal cord injury, ankle sprain, chondromalacia, anterior cruciate ligament injury, polyneuropathy, myelitis, Parkinson´s disease, meniscectomy, multiple sclerosis and cerebrovascular accident (Table 4).

Discriminative results on the application of the BESTest Portuguese version

The analysis of the two application moments of the BESTest Portuguese version shows a mean result of 75, 56 (±18, 65), regarding the total scale, for the first moment. All section values were calculated and presented (Table 5).

Concerning the second moment, performed eight days after, the BESTest presented a mean total score of 76, 28 (±18, 38). All section values were calculated and presented (Table 5).

Results of the reliability and internal consistency evaluation

The BESTest version translated into Portuguese obtained an inter-rater reliability of α=0,964. However, with the distribution by sections, this value (α) ranged between 0,788 and 0,948 (Table 6). Concerning intra-rater reliability, *Cronbach's* alpha was 0,964, ranging between 0,782 and 0,949 in the five scale sections (Table 7). Finally, *Cronbach's* alpha value concerning the internal consistency within sections ranged between 0, 80 and 0, 95. For the total results of the scale, the value obtained was α=0,964. Item 13 (section III) and item 19 - eyes open, firm surface (section V) of the scale were excluded from the internal consistency evaluation since they presented a variation equal to zero (Table 8).

DISCUSSION

The cultural and linguistic adaptation of the BESTest to the Portuguese language complied with the parameters regarding not only translation but also cultural adaptation within the Portuguese population. Therefore, its content validity and the good quality of translation were kept up, reason why it can be used with this population. The fact that it has been tested on the Portuguese population allows its validity, given that any assessment tool becomes valid for a surveyed population[32].The expert panel's consensus was supported by the verification of the coherence, linguistic, spelling, sentence, grammatical, punctuation, content or subjective

errors [1].

In spite of the high number of modification proposals by the expert panel, only 45 were accepted. This decision was taken in order to keep the cross-cultural equivalence of the object[1,12,27,29,38]; otherwise, it would have been necessary to adapt or to replace structural aspects of the original scale (e.g., to add details to some items regarding the task or the position of the item itself). Based on the same reason, from the 221 suggestions presented, 115 of linguistic nature were accepted. Otherwise, their modification could interfere with the semantic equivalence of the items[1,12,26-29,33-36].

That said, the linguistic modifications that would not change the original content were accepted. Modifications were based only on linguistic coherence improvement, better sentence construction or concept changes with synonyms, in order to be more understandable and present a better structure from a linguistic point of view.

Taking into account that the experts' modification proposals covered 15,57% of the tool items, there was no need for a new evaluation, since literature recommends a new appreciation by the expert panel[29] only when doubts exceed 20% of the items. In short, the modifications were made at the linguistic, spelling, sentence, grammatical and punctuation levels[1].

The pre-test sample revealed homogeneity, because predominantly feminine, with a mean age of 22 years and a very little comprehensive range of ages, including young higher education students, without relevant clinical antecedents. These factors do not allow drawing conclusions of the Portuguese population representativeness since it is a restricted and convenience based sampling.

The goal of the pre-test was to verify the perceptibility of the items during the BESTest application, from the perspective of both the rater and the sample, in order to identify needed modifications of some items; thus, making application easier and perceivable and, in addition, to verify the mean application time. This last issue is essential so that the BESTest application in users (patients) with pathologies will be useful at the clinical practice level. On the other hand, the idea of the mean time spent with the task allows informing the user and the rater on the time needed.

The minimum time obtained with the pre-test application was half of that indicated in the

BESTest original study[3] and the maximum time was five minutes lower. However, we must bear in mind that the pre-test sample included healthy individuals only and the original study included not only healthy individuals but also individuals with different pathologies.

Taking into account the sample characteristics (healthy young people), the score results obtained with the BESTest application were as expected, i.e., values near the scale total score.

The results of the questionnaire applied to the pre-test participants did not indicate any difficulty in perceiving the assessment tool items, as well as the physical therapist instructions. In addition, these results revealed feedback from the participants on the interpretation and fulfillment of the items, not being necessary the reformulation of any item/instruction to the user during the scale application.

The questionnaire filled in by the raters showed more variable results in comparison with the questionnaire filled in by the pre-test participants. Difficulty in providing the user with the appropriate instructions on fulfillment of tasks (difficulty at the specificity level) related with the rater's expression, was referred to items 12, 13, 18 and 27.

The sample selected for application of the BESTest Portuguese version revealed heterogeneity, because of its diversity (age, gender and diseases). It should also be noted that most of the balance assessment tools were validated within specific conditions, such as the case of the Berg Balance Scale[37], Tinetti's Falls Efficacy Scale[44] and Barthel Index[13]. The BESTest validation is an added-value for the population since it encompasses balance evaluation in different conditions.

The results obtained for the first application moment showed a total mean score of 75,56±18,646, while regarding the second moment (interval of eight days) the total mean score increased to 76,28±18,380. Scores increased in all sections, except for section V (sensory orientation) that revealed the same value. Section VI presented the greatest increase, followed by sections IV and I.

According to several authors, in order to achieve an acceptable variation of the reliability coefficient, its value must be between 0,70 and 0,90. Any value higher than this one is considered excellent[12, 39, 42, 43]. Thus, it is possible to say that the BESTest Portuguese version reveals an excellent reliability, since inter, intra-rater reliability and internal consistency as a whole, present a value of ICC = 0,964. The fact that the sample includes only 32 individuals

does not ensure the population representativeness, since the higher the sample is, the higher is the expected variation[39]. However, with this number of participants, it is already possible to observe the tendency of the results obtained since the normal distribution in the Gauss curve occurred. It is, therefore, possible to apply the statistical inference and the generalization of the results obtained[8].

The inter-rater reliability evaluation showed an expected value. The excellent value (α=1) is unlikely to obtain. In fact, the repeated use of an assessment tool in an equal sized sample leads to different results in each repetition, in general, due to several uncontrollable causes: assessment tool reading, ambient temperature, rater's fatigue and personality, as well as other atmospheric conditions[8].

In comparison with the original version, the Portuguese version showed a better inter-rater correlation coefficient, since the original scale obtained an ICC value of 0,91[3].This superiority of the inter-rater correlation was also observed in the Brazilian version, which indicated an ICC value of 0,93[24]. However, both values are considered excellent for the whole of the test. The ICC value is still higher in relation with the scales normally used**:** *Tinetti Mobility Assessment*: ICC=0,75-1,0; *Functional Reach Test*: ICC = 0,97 (according to Bennie *et al*) and ICC = 0,75 (according to Giorgetti *et al*); *Balance Screening Tool* (BST): rs=0.89 [3,41].

Regarding ICC, we obtained minimum values higher than those of the original scale and maximum values lower than those of the same scale (ICC = 0, 79-0, 96). These values allow us to say that the Portuguese version has a lower variation of the inter-rater correlation in the shared sections, thus being higher than the original version.

Comparatively, the Portuguese version presents only two sections with α values lower than the original version (sections III and V), while comparing with the Brazilian version, only sections I and III show lower values[3, 24]. However, in sections I, II, IV and VI the α value is higher than in the original version, thus demonstrating a higher inter-rater reliability for these sections.

The inter-rater correlation coefficient was of ICC = 0,964. This value proves that the results are consistent since two assessments were performed. That said, the inter-rater correlation coefficient of the BESTest is higher, in comparison with the Tinetti Mobility Assessment scale with a value of ICC = 0,95[44]. It was also observed that section III presented a lower

agreement level (ICC = 0,782) and section IV presented a higher agreement level (ICC = 0,949). These were expected values. On the one hand, section III evaluates mainly subjective items; then the agreement between the two assessment moments is unlikely. On the other hand, section IV showed a higher agreement level, since the items evaluated are mainly objective; then the probability of an agreement between the two assessment moments is higher.

The second moment of the BESTest application met the literature requirements (between two and fourteen days) since it was performed eight days after the first moment[12, 42].

The internal consistency evaluation revealed a *Cronbach's* alpha value of 0,96, therefore demonstrating that it is a reliable assessment tool, with an α value higher than that of the original scale: 0,88[4, 6, 26-29, 31, 33-36, 39, 46-49].

Considering that**:** the *Berg Balance Scale* (BBS), original and Norwegian versions, present values of 0,96 and 0,87, respectively; the Tinetti's balance scale (*Tinetti Balance Test*) presents: 0,91; the *Falls Efficacy Scale International (FES-I)* presents: 0,92; the *Unified Balance Scale (UBS)* presents: 0,98; the *Gait and Balance Scale (GABS* – Brazilian version) presents: 0,93 and the *Trunk Impairment Scale (TIS)* presents: 0,86; we can verify that regarding other balance assessment tools already used worldwide, only one demonstrates higher reliability values than the BESTest Portuguese version[47, 51-54].

CONCLUSION

Based on the test sample, we can conclude that the final BESTest Portuguese version has content validity and obtained excellent reliability (intra-and inter-rater) and internal consistency, both as a whole and within sections and items. Regarding the sample studied, we can also conclude that the BESTest presents a higher reliability and internal consistency in comparison with other balance assessment tools. Therefore, it represents not only an added-value but also an excellent evaluation option in people with balance disorders.

It is still important to refer that with the cultural adaption to the Portuguese language this tool should be promoted and widely applied in the clinical practice, in order to improve quality in intervention. That said, physical therapists will be provided with a unique and complete evaluation tool allowing them to identify balance deficit levels.

The process of cultural adaptation to the Portuguese language, validation, as well as reliability and internal consistency evaluation of the BESTest faced different limitations. The first one, concerning the external validity, is the reduced sample size, selected in a convenience basis; in addition, data collection was carried out in a restricted zone (Lisbon) only. Therefore, it was not possible to obtain the Portuguese population representativeness.

The goal was achieved considering the contribution of a consistent and reliable assessment tool, with reference psychometric characteristics, for using in the clinical practice.

The comparison with other similar assessment tools is suggested, in order to use a larger number of tools for testing psychometric characteristics, such as responsiveness, precision and acceptability. Further development of studies in the future with the application of the BESTest to specific and larger population samples is recommended.

The application of this assessment tool in other country districts is also suggested, in order to obtain a comparison of results in different regions of Portugal. A larger time period between the first and the second application moments is recommended, thus allowing the comparison of the values obtained in the first and the second moments, with the follow-up of the patients after physical therapy treatments.

The application of the BESTest as a balance assessment tool in specific conditions is another suggestion for future studies, in order to understand if this scale is reliable for a particular condition. Conducting studies with groups presenting several pathologies is also recommended, with the aim of perceiving in which condition the BESTest reveals capable measurement properties.

REFERENCES

1. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000;25(24):3186-91.
2. Yim-Chiplis PK, Talbot LA. Defining and measuring balance in adults. Biological research for nursing. 2000;1(4):321-31.
3. Horak FB, Wrisley DM, Frank J. The Balance Evaluation Systems Test (BESTest) to differentiate balance deficits. Physical therapy. 2009;89(5):484-98.
4. Leddy AL, Crowner BE, Earhart GM. Utility of the Mini-BESTest, BESTest, and BESTest sections for balance assessments in individuals with Parkinson disease. J Neurol Phys Ther. Jun;35(2):90-7.
5. Horak FB. Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls? Age and ageing. 2006;35 Suppl 2:ii7-ii11.
6. Jacobs JV, Kasser SL. Balance impairment in people with multiple sclerosis: preliminary evidence for the Balance Evaluation Systems Test. Gait & posture. 2012;36(3):414-8.

7. Alexander NB. Postural control in older adults. Journal of the American Geriatrics Society. 1994;42(1):93-108.
8. Viali L. Algumas Considerações Sobre A Denominada Curva Normal. VIDYA. 2014 34(1):99-116.
9. Huxham FE, Goldie PA, Patla AE. Theoretical considerations in balance assessment. The Australian journal of physiotherapy. 2001;47(2):89-100.
10. Lawton M, Teresi J. Annual Review of Gerontology and Geriatrics: Springer Publishing Company; 1994.
11. Maki BE, McIlroy WE. The role of limb movements in maintaining upright stance: the "change-in-support" strategy. Physical therapy. 1997;77(5):488-507.
12. Fortin M. O processo de investigação da concepção à realização. 1ª ed. Loures: Lusociência; 1996.
13. Araújo F, Ribeiro J, Oliveira A, Pinto C. Validação do Índice de Barthel numa amostra de idosos não institucionalizados. Revista Portuguesa de Saúde Pública. 2007;25(2).
14. Pollock A, Durward B, Rowe P. What is balance? Clinical Rehabilitation. 2000;14:402-6.
15. Mancini M, Horak FB. The relevance of clinical balance assessment tools to differentiate balance deficits. European journal of physical and rehabilitation medicine. 2010;46(2):239-48.
16. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation System's Test: the mini-BESTest. Journal of rehabilitation medicine: official journal of the UEMS European Board of Physical and Rehabilitation Medicine. 2010;42(4):323-31.
17. Maia A. Tradução e adaptação para o português - brasil do "balance evaluation systems test" e do minibestest e análise de suas propriedades psicométricas em idosos e indivíduos com doença de parkinson. Biblioteca Digital - UFMG: Escola de Educação Física, Fisioterapia e Terapia Ocupacional da Universidade Federal de Minas Gerais; 2012.
18. Thurman DJ, Stevens JA, Rao JK, Quality Standards Subcommittee of the American Academy of N. Practice parameter: Assessing patients in a neurology practice for risk of falls (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology. 2008;70(6):473-9.
19. Hummel N, Hufner K, Stephan T, Linn J, Kremmyda O, Brandt T, *et al*. Vestibular loss and balance training cause similar changes in human cerebral white matter fractional anisotropy. PloS one. 2014;9(4):e95666.
20. Chinsongkram B, Chaikeeree N, Saengsirisuwan V, Viriyatharakij N, Horak FB, Boonsinsukh R. Reliability and validity of the Balance Evaluation Systems Test (BESTest) in people with subacute stroke. Physical therapy. 2014;94(11):1632-43.
21. Berg WP, Alessio HM, Mills EM, Tong C. Circumstances and consequences of falls in independent community-dwelling older adults. Age and ageing. 1997;26(4):261-8.
22. Bambirra C, Magalhães L, Rodrigues-de-Paula F. Confiabilidade e validade do BESTest e do MiniBESTest em hemiparéticos crônicos. Rev Neurocienc. 2015;23(1):30-40.
23. Jones k, Horak FB, Winters K, Morea J, Bennett R. Fibromyalgia is Associated with Impaired Balance and Falls. J Clin Rheumatol. 2009;15(1):16-21.
24. Rodrigues LC, Marques AP, Barros PB, Michaelsen SM. Reliability of the Balance Evaluation Systems Test (BESTest) and BESTest sections for adults with hemiparesis. Brazilian journal of physical therapy. 2014;18(3):276-81.
25. Horak F. Translations Test Copies2009 [cited 2016 02-01-2016]. Available from: http://bestest.us/test_copies/.
26. Martín I, Paul C, Roncon J. Estudo de Adaptação e Validação da Escala de Avaliação de Cuidado Informal. Psicologia, Saúde & Doenças. 2000; 1(1):3-9.
27. Ramada-Rodilla J, Serra-Pujadas C, Delclós-Clanchet G. Adaptación cultural y validación de cuestionarios de salud: revisión y recomendaciones metodológicas. Salud publica de Mexico. 2013;55(1):57-66.
28. Pitta F, Probst V, Kovelis D, Segretti N, Leoni A, Garrod R, *et al*. Validação da versão em português da escala London Chest Activity of Daily Living (LCADL) em doentes com doença pulmonar obstrutiva crónica. Revista Portuguesa de Pneumologia. 2008; 14(1):27-47.
29. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. Journal of evaluation in clinical practice. 2011;17(2):268-74.

30. Horak FB. Clinical measurement of postural control in adults. Physical therapy. 1987;67(12):1881-5.
31. Vilelas J. Investigação: O processo de construção do conhecimento. 1ª ed. Lisboa: Edições Sílabo; 2009.
32. Guccione AA. Physical therapy diagnosis and the relationship between impairments and function. Physical therapy. 1991;71(7):499-503; discussion -4.
33. Touma Z, Ghandour L, Sibai A, Puzantian H, Hamdan A, Hamdan O, *et al*. Cross-cultural adaptation and validation of Behcet's disease quality of life questionnaire. BMC medical research methodology. 2011;11:52.
34. Domingues L, Cruz E. Adaptação Cultural e Contributo para a Validação da Escala Patient Global Impression of Change. ifisionline. 2011;2(1):31-7.
35. Gjersing L, Caplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. BMC medical research methodology. 2010;10:13.
36. Gonçalves R, Cavalheiroa L, Gil J, Rodrigues A, Coutinho A, Alves Henriques G, *et al*. Adaptação cultural e validação da versão portuguesa do Living with Asthma Questionnaire. Rev Port Pneumol. 2013;19(4):157-62.
37. Capucho A. Contibuto para a adaptação de uma escala de avaliação do equilíbrio em utentes idosos: escala de Berg. M16 FT A98: ESSA; 1998.
38. Streiner D. Being inconsistent about consistency: when coefficient alpha does and doesn´t matter. Journal of Personality Assessment. 2003;80:217-22.
39. Bland J, Altman D. Cronbach´s alpha. Br Med J (Clin Res Ed). 1997;314:314-572.
40. Langley F, Mackintosh S. Functional Balance Assessment of Older Community Dwelling Adults: A Systematic Review of the Literature. The Internet Journal of Allied Health Sciences and Practice. 2007;5(4).
41. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health technology assessment. 1998;2(14):i-iv, 1-74.
42. Polgar S, Thomas S. Introduction to research in the health scienses. 5ª ed: Elsevier; 2008.
43. Fortin M. Fundamentos e etapas do processo de investigação. 1ª ed. Loures: Lusodidacta; 2009.
44. Melo CA. Adaptação Cultural e Validação da Escala "Falls Efficacy Scale" de Tinetti. Ifisionline. 2011;1(2):33-43.
45. Oliveira Junior B, Jardim J, Nascimento O, Souza G, Baker T, Santoro I. Tradução, adaptação cultural e reprodutibilidade da Wisconsin Smoking Withdrawal Scale para o português do Brasil. Jornal Brasileiro de Pneumologia 2012;28(6).
46. Berg K, Wood-Dauphine S, Williams J, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. Physiotherapy Canada Journal. 1989;41(6):304-11.
47. Padgett P, Jacobs J, Kasser S. Is the BESTest at Its Best? A Suggested Brief Version Based on Interrater Reliability, Validity, Internal Consistency, and Theoretical Construct. Physical therapy. 2012;92:1197-207.
48. Miyamoto ST, Lombardi Junior I, Berg KO, Ramos LR, Natour J. Brazilian version of the Berg balance scale. Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas / Sociedade Brasileira de Biofisica [*et al*]. 2004;37(9):1411-21.
49. Halsaa KE, Brovold T, Graver V, Sandvik L, Bergland A. Assessments of interrater reliability and internal consistency of the Norwegian version of the Berg Balance Scale. Archives of physical medicine and rehabilitation. 2007;88(1):94-8.
50. Tinetti ME, Richman D, Powell L. Falls efficacy as a measure of fear of falling. Journal of gerontology. 1990;45(6):P239-43.
51. Kempen GI, Yardley L, van Haastregt JC, Zijlstra GA, Beyer N, Hauer K, *et al*. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. Age and ageing. 2008;37(1):45-50.
52. La Porta F, Franceschini M, Caselli S, Cavallini P, Susassi S, Tennant A. Unified Balance Scale: an activity-based, bed to community, and aetiology-independent measure of balance calibrated with Rasch analysis. Journal of rehabilitation medicine: official journal of the UEMS European Board of Physical and Rehabilitation Medicine. 2011;43(5):435-44.
53. Baggio J, Rodrigues G, Tumas V. Validação da versão brasileira da escala de marcha e equilíbrio e comparação com a escala de equilíbrio de Berg. Arq Neuro-Psiquiatr. 2013;71(9A):621-6.

54. Castellassia C, Ribeiro E, Fonseca V, Beinotti F, Oberg T, Lima N. Confiabilidade da versão brasileira da escala de deficiências de tronco em hemiparéticos. Fisioterapia em Movimento. 2009;22(2):189-99.

LIST OF TABLES

Table 1 - Statistical values results of the expert panel.

		Number of answers	Number of items	Mean	Standard deviation	Percentage (% Suggestions / items)
Total number of submitted items			194			
Experts panel answers	**Total**	1355				
	I agree unchanged	1134	72	5,85	1,41	83,69 / 37,11
	I agree with modifications	209	110	1,08	0,71	15,42 / 56,70
	I do not agree	12	12	0,06	0	0,86 / 6,19
Total number of suggestions		221	122*	5,84	1,41	15,57 / 62,89
Types of suggested changes	Linguistic nature	172	90	0,90	0	77,82 / 73,77
	Content modifications	41	37	0,22	0,71	18,55 / 30,33
	Scientific change	0	0	0	0	0 / 0
	Changes related to the subjectivity of the item	8	8	0,05	0	3,62 / 6,56
Changed suggestions		115	45 (100% linguistic amendment)			52,04 (66,86 das linguistic

					amendment) / 36,89
Proposals for modification unchanged	106	77			47,96 / 63,11
Number of suggested changes depending on the number of items	**Number of items**	**Percentage of changes made**	**Percentage of total items proposed to change**		
1 Panel member	12	26,67	9,84		
2 Panel members	16	35,56	13,11		
3 Panel members	10	22,22	5,15		
4 Panel members	3	6,67	2,46		
5 Panel members	3	6,67	2,46		
6 Panel members	1	2,22	0,82		

* There are 13 items of mixed suggestions or contain both language suggestions as subjectivity and content. This means that the total number of tips 122 is not 135.

Table 2 - Sample characteristics of the pre-test and scale application.

	Characteristics		**Individuals number**	**Percentage**
Sample pretest	**Gender**	Female	13	68,42%
		Male	6	31,58%
	Professional occupation	Education students	19	100 %
	Relevant Clinical antecedents	Displays	0	0 %
		Does not display	19	100 %
		Mean (Standard deviation)	**Minimum**	**Maximum**
	Age	22,05 (±1,31)	20	25

Table 3 - Pretest application results.

	Score (Percentage of the total score)
Mean	100,95 (93,47%)
Standard deviation	4, 55 (4, 21%)
Minimum	85 (78, 70%)
Maximum	107 (99, 07%)

Table 4 - Characteristics of data collection of the sample.

	Characteristics		**Individuals number**	**Percentage**
Sample for the application of the Portuguese version of the bestest scale	**Gender**	Female	19	54,29 %
		Male	16	45,71 %
	Professional occupation	Retireds	11	18,66 %
		Domestics	5	8,48 %
		Workers	16	27,14 %
	Diagnosis	Polyneuropathy	1	3,13 %
		Anterior cruciate ligament injury	1	3,13 %
		Chondromalacia	1	3,13 %
		Ankle sprain	1	3,13 %
		Spinal Cord Injury	1	3,13 %
		Tarsal fracture	1	3,13 %
		Low back pain	1	3,13 %
		Myelitis	2	6,25 %
		Parkinson's disease	2	6,25 %
		Meniscus arthroscopy rehabilitation	3	9,38 %
		Multiple sclerosis	3	9,38 %
		Cerebrovascular accident	15	46,88 %

		Right Hemisphere	7	46,67 %
		Left Hemisphere	8	56, 33 %
	Relevant Clinical antecedents	Displays	1	3,13 %
		Does not display	31	96,88 %
		Mean (Standard deviation)	**Minimum**	**Maximum**
	Age	53,97 (±14,98)	25	75

Table 5 - Application discriminative results Bestest the Portuguese version in the first and second time.

Bestest Sections	Mean obtained for each section by application of Bestest scale at the first time (standard deviation)	Mean obtained for each section by application of Bestest scale at the second time (standard deviation)
Section I	9,22 (±3,50; 2-15)	9,38 (±3,45; 2-15)
Section II	18,00 (±2,27; 9-21)	18,03 (±2,27; 9-21)
Section III	13,25 (±3,13; 7-18)	13,28 (±3,08; 7-18)
Section IV	9,88 (±5,02; 0-18)	10,06 (±5,09; 0-18)
Section V	12,56 (±1,85; 10-15)	12,56 (±1,88; 10-15)
Section VI	12,66 (±5,67; 2-21)	12,97 (±5,40; 2-21)
Total	75,56 (±18,65; s^2=347,67; x =71,50; X=67; 42-107)	76,28 (±18,38; s^2=337,82; x =71,50; X=65; 44-107)

Table 6 - Statistics of interobserver reproducibility for the sections and for the total score of the application of the Portuguese version of the bestest.

Bestest Sections	Intraclass coefficient (95% confidence interval)
Section I	0,826
Section II	0,846
Section III	0,788
Section IV	0,930
Section V	0,790
Section VI	0,948
Total	0,964

Table 7 - Statistics of intra-observer reproducibility for the sections and for the total score of the application of the Portuguese version of the bestest.

Bestest Sections	Cronbach's α (95% confidence interval)
Section I	0,824
Section II	0,847
Section III	0,782
Section IV	0,933
Section V	0,797
Section VI	0,949
Total	0,964

Table 8 - Statistics of internal consistency for the sections and for the total score

Bestest Sections	Cronbach's α (95% confidence interval)
Section I	0,83
Section II	0,85
Section III	0,80
Section IV	0,93
Section V	0,80
Section VI	0,95
Total	0,96