Brazilian Journal of Physical Therapy

https://www.journals.elsevier.com/brazilian-journal-of-physical-therapy

ORIGINAL RESEARCH

Modified Fresno test to assess Physical therapists' use of evidence-based practice: measurement properties of the Brazilian-Portuguese short version

Modified Fresno test Brazilian-Portuguese short version

Anderson Martins da Silva[a,b], Dafne Karen da Silva Vieira Rosa[a], Rosimeire Simprini Padula[a,c,*]

[a] *Masters and Doctoral Programs in Physical Therapy, Universidade Cidade de São Paulo, São Paulo, SP, Brazil*
[b] *Course of Physical Therapy, Centro Universitário do Vale do Ribeira (UNIVR/ UNISEPE), Registro, SP, Brazil*
[c] *Department of Physical Therapy, Universidade Cidade de São Paulo (UNICID), São Paulo, Brazil*

Abstract
Background: The modified Fresno test is a questionnaire developed to evaluate Evidence-Based Practice among physical therapists. A previous study has indicated that a shorter version with 9 items would be more appropriate for Brazilian physical therapists.
Objectives: To test the measurement properties of the modified Fresno test - Brazilian-Portuguese short version.
Methods: To analyze the reliability of the modified Fresno test - Brazilian-Portuguese short version by two raters (intra and inter-rater) in a sample of 133 physical therapists. The intraclass correlation coefficient ($ICC_{2,1}$) and standard error of measurement (SEM) was used to assess the test-retest reliability. Internal consistency was tested using the Cronbach's alpha coefficient. An expert committee analyzed content validity using the content validity index (CVI). Confirmatory factor analysis (CFA) using the principal components method was used to assess construct validity. Responsiveness was estimated using effect size, and ceiling and floor effects were also investigated.
Results: Inter-rater and intra-rater reliability were, respectively: ICC= 0.93 (95 % CI 0.91, 0.95); rater 1 - ICC= 0.95 (95 % CI 0.94, 0.97); rater 2 - ICC= 0.98 (95 % CI 0.98, 0.99). The agreement was very good (values $\leq$ 5 %). Internal consistency was good for most instrument items ($\geq$ 0.80). The CVI showed agreement among the expert committee members (0.96). The Cronbach's alpha coefficients calculated for the corrected item total showed values greater than 0.40. In the CFA,

* Corresponding authors at: Masters and Doctoral Programs in Physical Therapy, Universidade Cidade de São Paulo, Rua Cesário Galeno 475, 03071-000, São Paulo, SP, Brazil.
E-mails: amsfisio@gmail.com (A.M. da Silva), rosipadula@gmail.com (R.S. Padula).

the "model 2" showed acceptable indices (≥ 0.90). Responsiveness was classified as very small. No ceiling and floor effects.

Conclusion: The Fresno Modified Test - Brazilian-Portuguese short version has good to excellent reliability. CFA showed that the fit indices were adequate to be used in the population of interest.

Introduction

Evidence-based practice (EBP) can be defined as the conscious and judicious use of the best scientific evidence to guide therapeutic decisions.[1-4] The decision-making process must carefully consider three requirements:[5,6] high-quality clinical research, professional knowledge, and patient preferences. The EBP movement has steadily gained ground among physical therapists worldwide over the past decade.[3] It is directly related to the rise of global health problems[7] and advances in physical therapy studies and practice.[8] For some areas of physical therapy, the degree of scientific evidence available for clinical decision-making is very advanced,[9] although, for others, it is still very scarce.[10] However, EBP is of paramount importance for patients, physical therapists, and health services as it empowers physical therapists to accurately diagnose diseases, predict their course, and provide effective interventions. Consequently, we expect EBP to reduce healthcare costs while enhancing investments from public and private agencies.[11,12]

Several measurement tools have been developed to evaluate the use of EBP in clinical practice.[13-15] According to Shaneyfelt et al.,[16] the Fresno test is the only instrument that evaluates all stages of the EBP implementation process using more realistic clinical scenarios, enabling the assessment of EBP knowledge and skills in medical professionals.[16-19] The Fresno Test has adequate measurement properties and has been adapted to be used in other languages[8,16-17] and by different health professionals.[2,20-21]

The modified Fresno test for physical therapists is an adapted version of the Fresno test.[2,17] Regarding its measurement properties, there are strengths and weaknesses. The modified Fresno test has shown good internal consistency, excellent reliability, and little variability in intra- and inter-rater agreement. which may support its reliability, including the Brazilian-Portuguese version.[1,2,22] For exploratory and confirmatory factor analyses, the Brazilian-Portuguese version of the modified Fresno test showed good factor validity and internal consistency suitable for its use by physical therapist and students.[1,22] However, these analyses also revealed low reliability and low factor loading for items 2, 9, 10, and 11 of the instruments. Based on these results, a committee of experts suggested the modified Fresno test − Brazilian-Portuguese short version consisting of 9 items. This short version with 9 items showed better reliability than the modified full version with 13 items.[22] Hence, the modified Fresno test - short version may be more suitable to objectively assess all steps in the implementation of EBP among Brazilian physical therapists. Therefore, the objective of this study was to test the reliability (i.e., intra-rater end inter-rater reliability, measurement error, and internal consistency), validity (i.e., content validity and construct validity), responsiveness, and ceiling and floor effects of the modified Fresno test − Brazilian-Portuguese short version (i.e., 9 items) among Brazilian physical therapists.

Methods

Study design

This is a measurement property study of the modified Fresno test − Brazilian-Portuguese short version.[22] We followed the recommendations from the COnsensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) initiative.[23-25]

Participants and recruitment

A convenience sample of physical therapists was invited to participate in the study. The sample size was estimated as proposed by the COSMIN initiative for measurement properties tests.[23-25] Inclusion criteria were: (1) physical therapist professors from public and private institutions involved in clinical practice, and (2) physical therapist specialists in any clinical practice areas. Participants were excluded from the study if they refused to answer the test and if they started to answer but did not finish the online survey. The Universidade Cidade de São Paulo local research ethics committee (number #3.636.011/2019) approved the study. This measurement property study was conducted through an online survey using the SurveyMonkey platform. The study was advertised through direct e-mail and text messages to potential participants and posts on social networks (Instagram, Facebook, and Linked In).

Procedure

The following sociodemographic variables were collected for each participant: age, sex, institution where they graduated (private or public funded), time (years) since graduation, and qualification. The SurveyMonkey platform recorded the time spent by participants to complete the test.

Modified Fresno test − Brazilian-Portuguese short version

Data were collected using the modified Fresno Test - Brazilian-Portuguese short version with 9-items, proposed by Silva et al.[22] The test features an entry-level text with completion instructions and two clinical scenarios. The participant

must choose one of the clinical scenarios proposed by the test and have a maximum of 1 hour to answer all 9 items.[2,17] The total score of the modified Fresno test — Brazilian-Portuguese short version is calculated by summing up the answers from each item. For example, item 2 asks the participant to name the research databases to answer the chosen clinical scenario, which has the following scoring criteria: (a) variety of sources, (b) convenience, (c) clinical relevance, and (d) validity.[2,17] The answer to each item is scored on a 5-point Likert scale: (1) not evident, (2) limited, (3) minimal, (4) strong, and (5) excellent. The sum of the scores for each criterion results in a score per question that varies between 0 and 24 points: questions Q1 to Q6 allow a maximum score of 24 points for each question; for question Q7, the maximum score is 16 points, and for questions Q8 to Q9, a maximum score of 4 points per question. The total score is the sum of points for all questions and ranges from 0 to 168 points, with a higher score meaning a better understanding of EBP.[22]

Each participant answered the questionnaire twice (i.e., test-retest design), at the initial assessment (first assessment) and 7 days later (second assessment). At the first assessment, participants received instructions on how to answer the questionnaire and they completed the participant informed consent form. After agreeing to participate in the study and choosing one of the clinical scenarios presented, the participants had up to 1 hour to complete the questionnaire. After this first assessment, all participants were asked to answer the test again seven days later according to the instructions described in the questionnaire (second assessment). The instrument's total score was calculated by two independent assessors.[22]

Data analysis

Two assessors with experience in EBP analyzed the answers and calculated the total score of the questionnaire. They received a 3-hour training session about the scoring of the modified Fresno test — Brazilian-Portuguese short version. The first hour of the training was dedicated to explaining the questionnaire's scoring criteria. In the second hour of the training, a pilot test was conducted with each rater scoring one questionnaire. Finally, the third hour was dedicated to analysing and discussing the results of the pilot test. After the training, the assessors randomly scored each questionnaire blinded to the participants' details. Each assessor received the questionnaires after the first and second assessments. Responsiveness of the instrument was analyzed including participants who answered the questionnaire again four months after the first assessment. Our hypothesis is that changes in effect size will be detected after 4 months following the administration of the questionnaire. The effect is related to the fact that when using the modified Fresno test — Portuguese-Brazilian short version, the participant will benefit from the training and access to the information contained therein, providing learning and the search for new knowledge for the use of EBP.[22]

Statistical analysis

Descriptive statistics, including measures of central tendency (simple and relative frequencies, mean and median) and measures of dispersion (standard deviation), were used to describe the sociodemographic data, as well as the total scores and the score for each of the dimension (i.e., confirmatory factor analysis) of the modified Fresno test — Brazilian-Portuguese short version.[22]

For reliability and agreement we compared the score for each item and the total score of the test obtained in the first and second assessments for each assessor (intra-rater reliability) and between assessors for the score obtained in the first assessment (inter-rater reliability) with a 7-day interval. The reliability analyses followed the same method used in the analysis of the original instrument using intraclass correlation coefficient (ICC) of type 2,1 for continuous variables.[2] The ICC was classified according to *Fleiss* (1986)[26]: poor (<0.4), good (0.4−0.75), and excellent reliability (>0.75). The Standard Error of Measurement (SEM) was used to assess agreement.[13] The percentage of the SEM related to the total score of a questionnaire was interpreted according to Ostelo et al.[27] as ≤ 5 %, very good; > 5 % and ≤ 10 %, good; > 10 % and ≤ 20 %, doubtful; and > 20 %, negative.[27] Internal consistency analysis was performed by calculating the *Cronbach's alpha* coefficient for all items of the questionnaire and the corrected item-total correlations on data from assessor 1 collected at the first and second assessment.[28] Alpha value of ≥ 0.7 was defined as "acceptable" and ≥ 0.80 as "good." A value greater than 0.2 was considered for the corrected item-total correlations.[29]

Content validity was analyzed through the evaluation of a committee of 16 experts (physical therapists, statisticians, and epidemiologists), considering the content validity index. The content validity index measures the percentage of agreement among the members of the expert committee on the aspects addressed in the test items on items of the modified Fresno Test — Brazilian-Portuguese short version.[22] The acceptable agreement index among the expert committee members was 0.80. Confirmatory factor analysis was performed using the principal components method for structural validity.[30] The total score of the 266 responses in the test and retest was evaluated, totaling 1064 observations. The confirmatory factor analysis was used to investigate the adequacy of the modified Fresno test — a Brazilian-Portuguese short version, with 9 items, represented as "model 2", with the model proposed by Silva et al.[22] of 13 items, defined as "model 1". The first model evaluated the questionnaire's structure composed of 3 (three) factors. The second model tested the hypothesis that the questionnaire is composed of 2 (two) factors. The maximum likelihood method was used for the analysis. The following indices were used for model fitting: chi-square index, where values with significance levels $p < 0.05$ are estimated. The Goodness of Fit Index indicates how much the proposed models explain variance and covariance. The comparative fit index, the normed fit index, and the non-normed fit index, representing the proportional improvement in model fit, were also analyzed. Values greater than 0.90 for these indices indicate an adequate model. The root mean square error of approximation considers the error of approximation in the population in a covariance matrix. Values of 0.08 or less represent a reasonable error. The expected cross-validation index, which indicates the best fit of the models and is appropriate for comparing ungrouped models, was also analyzed. There are no benchmark values to rank model fit, and it should be as low as possible.

The responsiveness of the modified Fresno test - Brazilian-Portuguese short version was analyzed by calculating the effect size.[22] The effect size is the difference between the mean total score at the 4-month minus the mean total score at baseline (i.e., first assessment) divided by the standard deviation of the baseline scores. The effect size was interpreted as 0.01 as very small, 0.2 as small, 0.5 as medium, 0.8 as large, 1.2 as very large, and 2.0 as huge.[31]

To analyze the ceiling and floor effect, the assessor's percentage of minimum and maximum possible scores in the first and second assessments were calculated. Ceiling and floor effects were considered present when >15% of respondents achieved the lowest (i.e., 0 points) or highest (i.e., 168 points) possible total score. The presence of ceiling effects may indicate that the items are not challenging enough or too challenging, respectively.[23] The Statistical Package for the Social Sciences (SPSS IBM® - version 22.0) software was used for the statistical treatment of the data.

Results

Descriptive and reliability analysis

Eligible physical therapists were recruited for this study between January 2020 and July 2021. A total of 133 physical therapists responded to the questionnaire at the first and second assessments and were included in the reliability analysis. Of these, 82 (62%) participants answered the questionnaire again four months after the initial assessment and were included in the responsiveness analysis. Participants who were missing at four months did not complete the questionnaire due to lack of knowledge or disinterest in the study.

Table 1 presents the sociodemographic characteristics of the study participants. The sample consisted mainly of female participants (66.9%), predominantly between 30 and 39 years old (46.6%). Most of them have a maximum qualification of a master's degree (36.1%), having graduated from a private educational institution (79.7%) and completed the degree between 6 and 10 years ago (27.1%).

The intra-rater reliability for assessor 1 showed ICCs for each item ranging from 0.63 to 1.0, and SEM ranged from 0.11 to 2.76. Intra-rater reliability was rated as excellent for 8 out of the 9 items. The only item that showed good intra-rater reliability was item 2. The ICCs for most of the items were higher for assessor 2 compared to the ones found for assessor 1. For assessor 2, ICCs for each item ranged from 0.88 to 0.99, and SEM ranged from 0.34 to 2.12 points. These results should be interpreted as excellent intra-rater reliability for each item rated by assessor 2.

For the total score of the modified Fresno test − short version, the intra-rater reliability was rated as excellent for both assessors. Assessor 1 and assessor 2 showed ICCs values of 0.95 (95% CI: 0.94, 0.97) and 0.98 (95% CI: 0.98, 0.99), respectively. Inter-rater reliability was rated as excellent, with ICC 0.93 (95% CI: 0.91, 0.95). The SEMs were 10.38 points for the inter-rater reliability, 8.72 points for the intra-rater reliability of assessor 1, and 4.52 points for the intra-rater reliability of assessor 2. The SEMs were classified as very good (values ≤ 5% of the total score) (Table 2).

Table 1 Sociodemographic characteristics of the sample (*N* = 133).

Characteristics		N (%)
Participants		133 (100)
Sex		
	Male	44 (33.1)
	Female	89 (66.9)
Qualification		
	Undergraduate	6 (4.5)
	Specialization	32 (24.1)
	Masters	48 (36.1)
	Doctorate	36 (27.1)
	Postdoctoral	11 (8.3)
Graduation Institution		
	Public	27 (20.3)
	Private	106 (79.7)
Time since Graduation (years)		
	1−5	14 (10.5)
	6−10	36 (27.1)
	11−15	30 (22.6)
	16−21	33 (24.8)
	22−27	9 (6.8)
	≥ 28	11 (8.3)
Age (years)		
	21 to 29	21 (15.8)
	30 to 39	62 (46.6)
	40 to 49	41 (30.8)
	50 to 59	7 (5.3)
	Between 60 and 63	2 (1.5)

The internal consistency obtained with Cronbach's alpha indicated a value of 0.93 for the nine items of the instrument. The α coefficients calculated for the corrected item-total correlations showed values greater than 0.40.

Content validity

Content validity was assessed using the content validity index, which measures the proportion or percentage of agreement among the expert committee on the questionnaire items. A content validity index of 0.96 was found.

Construct validity

The confirmatory factor analysis showed the "Model 2"; with nine items and two factors provided a better fit compared to "Model 1" with 13 items and three factors (Fig. 1).

The *goodness of fit index* and *comparative fit index* reached an acceptable level (≥ 0.90). "Model 2", which tested a two-factor structure, provided a better fit to the data (Table 3). However, the fit indices of the second model also did not reach acceptable levels. Therefore, the second model, although better than the first model, did not adequately fit the data.

The responsiveness analysis of the modified Fresno test - Brazilian-Portuguese short version is presented in Table 4, with the results of the effect sizes of the mean differences of the total scores in the first and second stages. The table analyzed the effect size appropriate for the paired *t*-test.

Table 2 Intra- and inter-rater reliability of the modified Fresno test — Brazilian-Portuguese short version, per item, and the sum total of the items.

	Inter-rater		Intra-rater 1		Intra-rater 2	
	(N = 133)	(N = 133)	(N = 133)	(N = 133)	(N = 133)	(N = 133)
Questions	ICC $_{2,1}$ (95 % CI)	Agreement— (SEM)	ICC $_{2,1}$ (95 % CI)	Agreement— (SEM)	ICC $_{2,1}$ (95 % CI)	Agreement— (SEM)
Q1	0.93 (0.90, 0.95)	2.61	0.94 (0.92, 0.96)	2.46	0.99 (0.99, 0.99)	0.80
Q2	0.70 (0.52, 0.81)	3.44	0.63 (0.33, 0.78)	2.42	0.92 (0.88, 0.94)	1.69
Q3	0.91 (0.87, 0.93)	2.59	0.88 (0.84, 0.92)	2.76	0.95 (0.93, 0.97)	1.85
Q4	0.76 (0.67, 0.83)	3.82	0.92 (0.88, 0.94)	2.19	0.92 (0.87, 0.95)	2.12
Q5	0.91 (0.88, 0.94)	2.42	0.96 (0.94, 0.97)	1.68	0.95 (0.92, 0.96)	1.80
Q6	0.90 (0.86, 0.93)	1.79	0.78 (0.44, 0.89)	2.89	0.93 (0.90, 0.95)	1.64
Q7	0.64 (0.50, 0.75)	3.07	0.99 (0.99, 0.99)	0.44	0.88 (0.84, 0.92)	1.95
Q8	0.70 (0.51, 0.81)	1.20	1.00 (1.00, 1.00)	0.11	0.97 (0.96, 0.98)	0.42
Q9	0.81 (0.72, 0.86)	1.03	1.00 (1.00,1.00)	0.11	0.98 (0.97, 0.98)	0.34
Total	0.93 (0.91, 0.95)	10.38	0.95 (0.94, 0.97)	8.72	0.98 (0.98, 0.99)	4.52

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; SEM, standard error of measurement.

For the total score, the effect size was 0.13 points and rated as "very small," i.e., 4-month follow-up minus the baseline.

No ceiling and floor effects were found among respondents.

Discussion

Our results showed excellent intra- and inter-rater reliability, very good intra- and inter-rater agreement, good internal consistency, and no ceiling or floor effects for the modified Fresno test — Brazilian-Portuguese short version. Content validity showed agreement among the expert committee members, and construct validity was adequate as a suitable alternative to assess EBP in physical therapists and students. The responsiveness analysis showed a very small effect size.

The scores on the individual questions and the total score of the instrument showed excellent intra- and inter-rater reliability results. These values were also equivalent to the original English version, which showed moderate to excellent intra-rater reliability for both raters for all questions and moderate to excellent inter-rater reliability for all questions. Internal consistency by Cronbach's alpha indicated satisfactory reliability for all instrument items. These values agree with the results obtained in studies developed by Silva et al.,[1] Tilson,[2] Ramos et al.[17] and Argimon-Pallas et al.[32] These studies state that the test is reliable in different languages and can be used to evaluate other professionals.

Confirmatory factor analysis (CFA) tested the proposed 2-factor model (model 2) against the short version proposed by Silva et al.[22] Two of the nine fit indices (GFI and CFI) used by the maximum likelihood method reached acceptable values ($\geq$ 0.90) in Model 2. The *Normed Fit Index (NFI)* and *Non-Normed Fit Index (NNFI)* indexes also showed values close to acceptable and represented the proportional improvement in the fit of model 2.
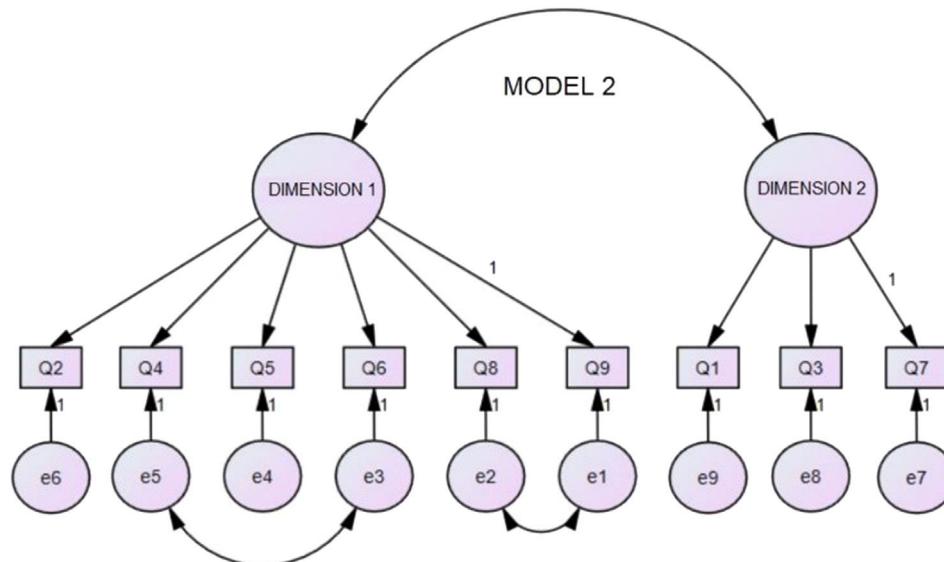

Fig. 1 Confirmatory factor analysis diagram (CFA).

Table 3 Fit indices for the two-factor models tested according to confirmatory factor analysis.

Models	X^2	df	GFI	CFI	RMSEA	SRMR	AIC	ECVI	NFI	NNFI
Model 1	254.58	62	0.85	0.82	0.11	0.10	312.58	1.38	0.78	0.73
Model 2	52.03	24	0.91	0.92	0.09	0.08	94.03	0.71	0.87	0.81

Abbreviations: AIC, akaike's information criterion; CFI, comparative fit index; ECVI, expected cross-validation index; GFI, goodness of fit index; NFI, normed fit index; NNFI, non-normed fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual.

Table 4 Effect sizes of the mean differences of the total scores of the modified Fresno test — Brazilian-Portuguese short version, according to the evaluation times (first and second stage).

	Evaluation	N	ES (Glass's Delta)	95 % CI		Interpretation[1] (Effect size)
				Inf	Sup	
Total Sum of Answers	1st and 2nd STAGES	133	0.13	−0.11	0.37	Very Small

Abbreviations: CI, confidence interval; DP, standard deviation; ES, effect size.
[1] Classification based in Sawilowsky (2009).[21]

The responsiveness analysis evaluated the effect size for the total sum of responses, using Glass Delta values, presenting values of 0.13 points, classified as "very small" and" showing a small difference between the groups analyzed. This highlights that the modified Fresno test — Brazilian-Portuguese short version, was responsive in detecting group changes over time. Unresponsive instruments are unable to detect changes in a test over time.

The presence of ceiling and floor effects can influence the sensitivity and responsiveness of a measuring instrument. However, no ceiling and floor effects were observed for the respondents in the instrument's total score. This also highlights the sensitivity and ability of the instrument to detect changes in relation to knowledge about EBP over time. The ceiling and floor effects were not analyzed in the 13-item Modified Fresno Test.

The lack of knowledge about the adoption of EBP was the justification of 51 (38 %) of the participants who did not complete the instrument on the second occasion. These difficulties may be related to the biggest obstacle presented by students, physical therapists, and other health professionals in adopting EBP, which is the difficulty in searching, interpreting, and transposing evidence to clinical practice. This obstacle is directly related to the competencies and skills of each professional. Other barriers to the adoption of EBP are also pointed out in other studies, such as limited access to databases and full texts, language issues, and the time available for updating.

The sample used in this study may characterize a limitation. As in other studies of measurement properties, a convenience sample was used instead of determining a sample size to achieve statistically significant results. The 7-day time interval between test and retest applications may also be a study limitation and may bias the correlation coefficient calculation. More research should be done to confirm the instrument's measurement properties in other samples with different characteristics. It is suggested that studies be conducted comparing the test version presented here with other versions. Even so, it would be important to attest to the ability of the instrument to discriminate between theoretically different groups.

The strength of this study was to demonstrate that the modified Fresno test — Brazilian-Portuguese short version, maintained the quality of its measurement properties. This makes the instrument easier for respondents and more reliable for assessing EBP in online assessments.

Conclusion

The modified Fresno test — Brazilian-Portuguese short version showed excellent intra- and inter-rater reproducibility, very good agreement, and good internal consistency. The content validity test was in agreement among the experts and, through Confirmatory Factor Analysis, showed good fit indicesThese results confirm that the short version is a suitable alternative to comprehensively assess EBP in physical therapy professionals and students.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

References

1. Silva AM, Costa LCM, Comper ML, Padula RS. Cross-cultural adaptation and reproducibility of the Brazilian-Portuguese version of the modified FRESNO Test to evaluate the competence in evidence based practice by physical therapists. *Braz J Phys Ther*. 2016;20:26−47.

2. Tilson JK. Validation of the modified Fresno Test: assessing physical therapists' evidence based practice knowledge and skills. *BMC Med Educ*. 2010;10:38.

3. Herbert R, Jamtvedt G, Mead J, Hagen KB. *Practical Evidence-Based Physiotherapy*. 2nd ed. London: Churchill Livingstone; 2011.

4. Elkins MR, Moseley AM, Pinto RZ. Usage evaluation of the physiotherapy evidence database (PEDro) among Brazilian physical therapists. *Braz J Phys Ther*. 2015;19.

5. Silva TM, Costa LC, LO Costa. Evidence-Based Practice: a survey regarding behavior, knowledge, skills, resources, opinions and perceived barriers of Brazilian physical therapists from São Paulo state. *Braz J Phys Ther*. 2015;19(4):294—303.

6. Saragiotto BT, Costa LC, Oliveira RF, Lopes AD, Moseley AM, Costa LO. Description of research design of articles published in four Brazilian physical therapy journals. *Braz J Phys Ther*. 2014;18(1):56—62.

7. Claridge JA, Fabian TC. History and development of evidence-based medicine. *World J Surg*. 2005;29(5):547—553.

8. Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990—2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*. 2016;388(10053):1545—1602.

9. WCfP Therapy. Physical therapy: a global profile. *World Confederation for Physical Therapy*. 2015.. http://www.wcpt.org. [cited 2017 agost].

10. Zadro JR, Ferreira G. Has physical therapists' management of musculoskeletal conditions improved over time? *Braz J Phys Ther*. 2020;24(5):458—462. Epub 2020 May 5. PMID: 32387047; PMCID: PMC7563797.

11. Kaizik MA, Hancock MJ, Herbert RD. DiTA: helping you search for evidence of diagnostic test accuracy in physical therapy. *Braz J Phys Ther*. 2024;28(3):101082. Epub ahead of print. PMID: 38848626; PMCID: PMC11192776.

12. Munneke W, Demoulin C, Roussel N, et al. Comparing physical therapy students' attitudes and beliefs regarding chronic low back pain and knee osteoarthritis: an international multi-institutional comparison between 2013 and 2020 academic years. *Braz J Phys Ther*. 2024;28(1):100592.

13. Boström A-M, Sommerfeld DK, Stenhols AW, Kiessling A. Capability beliefs on, and use of evidence-based practice among four health professional and student groups in geriatric care: a cross sectional study. *PLoS ONE*. 2018;14:13.

14. Schreiber J, Stern P. A review of the literature on evidence-based practice in physical therapy. *The Int J Allied health Sci Practice*. 2005;3:1—10.

15. Hoogeboom TJ, Jette AM. Using evidence hierarchies to find the best evidence: a procrustean bed? *Phys Ther*. 2021;101(11): pzab235.

16. Shaneyfelt T, Baum KD, Bell D, et al. Instruments for evaluating education in evidence-based practice. *JAMA*. 2006;296:1116—1127.

17. Ramos KD. Validation of the Fresno test of competence in evidence-based medicine. *BMJ*. 2003;326:319—321.

18. Kathleen DR, Sean S, SM T. Validation of the Fresno test of competence in evidence based medicine. *BMJ*. 2003;326:319—321.

19. Ilic D. Assessing competency in Evidence Based Practice: strengths and limitations of current tools in practice. *BMC Med Educ*. 2009;9(53).

20. Laibhen-Parkes N, Kimble LP, Melnyk BM, Sudia T, Codone S. An adaptation of the original fresno test to measure evidence-based practice competence in pediatric bedside nurses. *Worldviews Evid Based Nurs*. 2018;15(3):230—240. Epub 2018/05/08.

21. Rothberg B, Feinstein RE, Guiton G. Validation of the colorado psychiatry evidence-based medicine test. *J Grad Med Educ*. 2013;5(3):412—416. Epub 2014/01/10.

22. da Silva AM, Padula RS. Factor structure and short version of the modified Fresno test to assess the use of the evidence-based practice in physiotherapists. *BMC Med Educ*. 2021;21:135.

23. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34—42.

24. Mokkink LB, Prinsen CA, Bouter LM, Vet HC, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther*. 2016 Jan 19;20 (2):105—113.

25. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010 May;19 (4):539—549.

26. Fleiss JL. *The Design and Analysis of Clinical Experiments*. New York: John Wiley & Sons; 1986.

27. Ostelo RWJG, de Vet HCW, Knol DL, van den Brandt PA. 24-item Roland-Morris disability questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. *J Clin Epidemiol*. 2004;57:268—276.

28. Thorndike R.M. Book review : psychometric theory (3rd ed.) by Jum Nunnally and Ira Bernstein New York: McGraw-Hill, 1994, xxiv + 752 pp. Vol. 19. 1995.

29. Kline P. *A Handbook of Test Construction (Psychology Revivals)*. 1nd ed. London: Routledge; 2015.

30. Brusaca LA, Moriguchi CS, Barbieri DF, Stevens ML, Oliveira AB. Brazilian version of need for recovery scale: assessment of structural validity, criterion validity, and internal consistency. *Braz J Phys Ther*. 2022;26(6):100465.

31. Sawilowsky SS. New effect size rules of thumb. *J Mod Appl Stat Methods*. 2009;8(2):467—474.

32. Argimon-Pallas JM, Flores-Mateo G, Jimenez-Villa J, Pujol-Ribera E. Psychometric properties of a test in evidence based practice: the Spanish version of the Fresno test. *BMC Med Educ*. 2010;10:45.. Epub 2010/06/18.