

RESEARCH

Open Access



Psychometric and edumetric properties of the Turkish version of the assessment of physiotherapy practice

Halil İbrahim Çelik^{1*}, Demet Öztürk², Mustafa Sari², Numan Bulut³ and Nilgun Bek²

Abstract

Background Clinical education is a crucial process in which students learn how to use theoretical knowledge in clinical settings. However, there is no standardized assessment tool routinely used to evaluate the clinical performance of physiotherapy students in Turkey. This study aimed to examine the psychometric (validity and reliability) and edumetric (acceptability, feasibility, and educational impact) properties of the Turkish version of the Assessment of Physiotherapy Practices (APP-T).

Methods This methodological study included 7 clinical educators and 174 4th-grade physiotherapy students at three universities in Turkey. Students were assessed with the APP-T on completion of 4-week clinical placements. The construct validity was examined using the exploratory factor analysis. The internal consistency was determined using Cronbach's α -coefficient. The test-retest and inter-rater reliability were evaluated with the intra-class correlation coefficient (ICC). For the measurement error of the APP-T, standard error of measurement (SEM) and minimum detectable change (MDC) values were calculated. After the administration of the APP-T was completed, clinical educators were requested to provide feedback on the acceptability, applicability, and educational impact of the APP-T.

Results Exploratory factor analysis demonstrated the 20 items of the APP-T represent a single dominant dimension explaining 76.28% of the variance. Excellent test-retest and inter-rater reliability was detected by ICC = 0.94 and ICC = 0.77, respectively, and good internal consistency was detected by Cronbach's α = 0.935. The MDC₉₀ values for test-retest and inter-rater reliability were 3.11 and 6.86, respectively. Clinical educators provided generally positive feedback on the acceptability, feasibility, and educational impact of the APP-T.

Conclusions The current findings provided evidence for universities and clinical educators that the APP-T has sufficient psychometric and educational properties for evaluating the clinical performance of physiotherapy students in Turkey.

Keywords Validation study, Physical therapy specialty, Professional practice, Educational measurement

*Correspondence:

Halil İbrahim Çelik
fzyoterapist70@gmail.com

¹Bilge Çocuk Special Education and Rehabilitation Center, Beysuikent, Çankaya, Ankara s06800, Turkey

²Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation, Lokman Hekim University, Ankara, Turkey

³Faculty of Physical Therapy and Rehabilitation, Hacettepe University, Ankara, Turkey



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

The process of physiotherapy education, as in most other health disciplines, includes clinical education as well as traditional academic education. The World Physiotherapy, previously known as the World Confederation of Physical Therapy, argues that clinical education should comprise no less than one-third of the physiotherapy curriculum, which emphasizes the crucial role of clinical education in training physiotherapists [1]. Clinical education enables students to efficiently integrate theoretical knowledge with physical (e.g., manual therapy techniques), cognitive (e.g., selecting appropriate assessment and intervention methods), and communication (e.g., conducting a proper interview with the clients) skills into practice. Thus, it empowers students with competence in knowledge, skills, and attitudes required to work as qualified professionals [2, 3].

Students' clinical performance can be evaluated through a variety of methods, involving oral presentations, written and practical examinations, oral presentations, objective structured clinical examinations, final theses, case history-based projects, and portfolios [1]. However, the assessment of clinical performance is commonly conducted observationally by a supervisor clinician known as a clinical educator. Although this can be extremely challenging due to the subjective nature of observational assessments, using psychometrically sound standardized assessment tools can accurately reflect the student's clinical competence [4, 5]. In a systematic review conducted in 2018 examining the psychometric evidence of assessment tools used in the evaluation of physiotherapy students' clinical performance, it was reported that the psychometric features of all 14 identified assessment tools were not sufficient. Yet, it was noted that a class 2 recommendation was made for the Assessment of Physiotherapy Practice (APP), meaning that its use could be considered [6]. This recommendation, along with the APP's relatively robust psychometric properties, influenced our decision to select this tool.

The APP is a valid and reliable assessment tool that has been nationally accepted for use in Australia and New Zealand for evaluating the clinical performance of physiotherapy students [7, 8]. According to unpublished information from the APP research team, the APP has been translated into several languages, including Portuguese, Spanish, Hebrew, Icelandic, and Thai, with ongoing efforts to translate it into French. However, the only translated version documented in the literature is Chinese [9]. Therefore, the current study aimed to translate and cross-culturally adapt the APP into Turkish and examine its psychometric (validity and reliability) and edumetric (acceptability, applicability, and educational impact) properties. The sufficiency of psychometric and educational properties would enable the Turkish version

of the APP (APP-T) to be used as a standardized assessment tool for evaluating the clinical performance of physiotherapy students in Turkey.

Methods

This methodological study was initiated concurrently at the Lokman Hekim, Hacettepe, and Gazi Universities in Ankara, Turkey in May 2023 and continued until February 2024. The study was approved by the Ethics Committee of Lokman Hekim University (April 5, 2023, decision no: 2023/4) and conducted in accordance with the Helsinki Declaration. Prior to recruitment in the study, all clinical educators and students who participated in this study signed an informed consent form.

In the first phase of the study, which was executed in two phases, the translation into Turkish and cross-cultural adaptation of the APP was conducted using the guidelines presented by Beaton et al. [10]. The protocol for phase 1 is provided as a supplement (See: supplementary file 1). In the second phase, the psychometric properties (validity and reliability) of APP-T were examined with the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) [11, 12]. Furthermore, feedback on the edumetric properties (acceptability, applicability, and educational impact) of the APP-T was obtained from clinical educators.

Participants

One hundred and seventy-four 4th-grade students enrolled in entry-level physiotherapy programs at three universities participated in the study. Students were assessed by 7 clinical educators with the APP-T after completing a 4-week clinical placement.

Outcome measures

The APP includes 20 items to evaluate entry-level competence across the following 7 areas of clinical practice: professional behavior, communication, assessment, analysis and planning, intervention, evidence-based practice, and risk management. Each item is scored from 0 (infrequently/rarely demonstrates performance indicators) to 4 (demonstrates most performance indicators to an excellent standard). A score of 2 (demonstrates most performance indicators to an adequate standard) shows that the minimal requirement for an entry-level physiotherapist has been reached. In addition, a global rating scale on overall clinical performance (not adequate, adequate, good, excellent) is scored by the clinical educator, yet this scoring is not added to the APP total score [7, 8].

Preparation of the clinical educators

Approximately 90 min of online training was conducted with clinical educators invited to the study. The training content was based on the manual of the original APP

[13]. In this training, detailed information about the purpose of the study, the features of the APP-T, and the assessment process was provided.

Preparation of the students

Students attending six clinical placements (developmental and early physiotherapy, neurological early rehabilitation, pediatric neuromuscular rehabilitation, rheumatological rehabilitation, prosthetic and orthotic rehabilitation, and sports physiotherapy subspecialties) were accepted to enroll in the study.

Before the placements commenced, students were informed about APP-T. In the 2nd week of each placement, mid-unit feedback for each student was provided by the clinical educators so that students who administered the APP-T in mid-unit might compare their results. At the end of the 4-week of each placement, clinical educators administered a final APP-T for an end-of-unit assessment. The original APP manual served as the foundation for the whole assessment process [13].

Post-placement feedback on the use of APP-T

After all clinical placements were finished, clinical educators were asked to complete a feedback questionnaire created by the researchers about the use of the APP-T. This questionnaire consisted of a 5-point Likert scale (1=Strongly disagree, 2=Disagree, 3=Undecided, 4=Agree, 5=Strongly agree) and focused on the acceptability, applicability, and educational impact of APP-T. The 4 and 5 points were categorized as “agree” and 1, 2, or 3 points as “disagree”.

Statistical analysis

All statistical analysis was conducted using the IBM SPSS Statistics for Windows v26.0 (SPSS Inc., Chicago, USA). Participant characteristics are presented as means and standard deviations for numerical data, and counts/percentages for categorical data. A p-value less than 0.05 was determined as an indicator of statistical significance. The psychometric properties tested in this study were as follows:

1. **Content validity:** It was assessed by considering both the perspectives of clinical educators and students. A pilot study of 10 clinical educators and 10 students was conducted to determine the relevance and comprehensiveness of APP-T. Each item was scored by the students and clinical educators on a four-point Likert scale (1 = not relevant or comprehensive, 4 = completely relevant or comprehensive) and the content validity index (CVI) was calculated [14].
2. **Construct validity:** Exploratory factor analysis was used to examine the construct validity of the APP-T. The sufficiency of the sample size and suitability of the data for the factor analysis were examined by the Kaiser Meyer Olkin (KMO) Test (the value is recommended to be higher than 0.50) and Bartlett's Test (p-value must be below 0.05), respectively. The analysis was conducted by using principal axis factoring. Scree plot, eigenvalue, and parallel analysis were employed to decide the number of factors retained. Factor loadings were accepted as significant if above 0.30 [15]. A sample size of at least 100 participants and between 4 and 10 participants per item is recommended for factor analysis [16]. Hence, our sample of 174 was sufficient to conduct this analysis.
3. **Internal consistency:** It was evaluated by Cronbach's α coefficient and item-total score correlations. For the total score, a Cronbach's α value of ≥ 0.70 is recommended [16].
4. **Test-retest reliability:** A clinical educator re-completed the APP-T within 48 to 72 h for 33 students attending the clinical placement of rehabilitation in pediatric neuromuscular diseases. It was analyzed with the intra-class correlation coefficient (ICC) and 95% confidence interval (CI) based on a single measure and a 2-way random effects model. ICC values above 0.75 were considered as excellent [17].
5. **Inter-rater reliability:** Two clinical educators, both with equal supervision roles during the placement of rehabilitation in pediatric neuromuscular diseases, independently administered the APP-T for 19 students after clinical placements were completed. As the individual item score is ordinal and not interval-ratio data, the reliability of individual APP items was evaluated using a linear-weighted Kappa (κ_w) coefficient (excellent (> 0.80), good (0.60–0.80), moderate (0.40–0.60), fair (0.20–0.40), and poor (≤ 0.20)) [18]. The APP total score is arguably, interval-ratio data, and therefore, evaluated using an ICC [17].
6. **Measurement error:** The standard error of measurement (SEM) and the minimal detectable change with 90% CI (MDC_{90}) were obtained to examine the measurement error. The ICC values were used in calculating the SEM and then the MDC_{90} . The MDC_{90} was calculated by multiplying the SEM by 1.65 and the square root of 2 to adjust for sampling from 2 different measurements [19].

Post-hoc power analyses were performed with the R Studio (packages “ICC.sample.size”, R Core Team, 2022). With a statistical significance of $\alpha=0.05$, obtained ICCs were 0.94 and 0.77 for test-retest and inter-rater reliability, and sample size=33 and 19 for test-retest and

Table 1 Characteristics of participating clinical educators and students

	Main Study	Inter-rater study	Test-retest study
Clinical educators	n=7	n=2	n=1
Age (years)	31.86±4.10	30.50±3.54	33
Clinical education experience (years)	8.86±3.78	8.00±2.83	10
Gender (female)	4 (62.50)	0 (0)	0 (0)
Gender (male)	3 (37.50)	2 (50)	1 (100)
Students	n=174	n=19	n=33
Age (years)	22.76±1.86	22.74±1.63	22.52±1.39
Gender (female)	145 (83.33)	17 (89.47)	24 (72.73)
Gender (male)	29 (16.67)	2 (10.53)	9 (27.27)
Clinical placements			
Developmental and early physiotherapy	40 (22.99)	0 (0)	0 (0)
Neurological rehabilitation	9 (5.17)	0 (0)	0 (0)
Rehabilitation in pediatric neuromuscular diseases	51 (29.31)	19 (100)	33 (100)
Rheumatological rehabilitation	19 (10.92)	0 (0)	0 (0)
Prosthetic and orthotic rehabilitation	11 (6.32)	0 (0)	0 (0)
Sports health	44 (25.29)	0 (0)	0 (0)

values are given as mean±SD or n (%)

inter-rater reliability, the post-hoc powers (1-β) were found to be 99.9%, and 71.6% for the APP-T total score.

Results

This study included 7 clinical educators (62.50% female; mean age=31.86±4.10 years) with a mean clinical education experience of 8.86±3.78 years. Each clinical placement was facilitated by one clinical educator. One additional clinical educator was involved only in one clinical placement (rehabilitation in pediatric neuromuscular diseases) that was tested for inter-rater reliability. All of the clinical educators were academicians, with 4 having a PhD degree and 3 having a Master’s degree in physiotherapy. A total of 174 physiotherapy students (83.33% female; mean age=22.76±1.86 years) participated in this study. The mean APP total and global score were 52.01±17.47 and 2.65±0.79, respectively. 93.7% of students were rated adequate, good, or excellent on the global rating scale of the APP-T. All clinical educators and students who consented to participation completed the study. Characteristics of participating clinical educators and students are presented in Table 1.

Cross-cultural and content validity

The translation process was completed without any difficulties. All translations were examined by the expert committee (five academicians) and it was confirmed that semantic, idiomatic, experiential, and conceptual

Table 2 Internal consistency and factor loadings for the items of the APP-T (n=174)

Items	Mean (SD)	Item-total correlation	Cronbach’s alpha if the item deleted	Factor loadings
Item 1	2.97±1.01	0.868	0.934	0.874
Item 2	2.80±1.05	0.861	0.934	0.868
Item 3	2.94±1.09	0.896	0.933	0.902
Item 4	2.80±1.02	0.860	0.934	0.867
Item 5	2.74±1.05	0.889	0.934	0.890
Item 6	2.69±0.99	0.887	0.934	0.894
Item 7	2.61±1.03	0.895	0.933	0.902
Item 8	2.55±0.87	0.843	0.934	0.850
Item 9	2.50±0.86	0.820	0.934	0.829
Item 10	2.47±0.94	0.839	0.934	0.846
Item 11	2.50±0.97	0.865	0.934	0.873
Item 12	2.45±0.98	0.864	0.934	0.871
Item 13	2.51±0.96	0.866	0.934	0.872
Item 14	2.72±1.03	0.890	0.933	0.918
Item 15	2.37±1.02	0.893	0.934	0.901
Item 16	2.59±1.00	0.878	0.934	0.886
Item 17	2.52±1.04	0.868	0.934	0.875
Item 18	2.26±0.89	0.825	0.934	0.830
Item 19	2.37±1.08	0.827	0.934	0.835
Item 20	2.66±0.95	0.872	0.934	0.879

equivalence between the Turkish and English versions was appropriate.

Content validity was assessed by considering both the perspectives of clinical educators and students. A pilot study of 10 clinical educators and 10 students was conducted to determine the relevance and comprehensiveness of APP-T. The rate of items with 4 points for both relevance and comprehensiveness is 90%. Each item was scored and the content validity index (CVI) [14] was found to have the maximum value (1.00). Therefore, no changes were made to the items and it was decided that APP-T could be applied.

Construct validity

The Kaiser–Meyer–Olkin and Bartlett’s sphericity test confirmed that the sample size was sufficient (KMO=0.977) and the items were appropriate to conduct factor analysis (χ²=4557.84; p<0.001). Exploratory factor analysis demonstrated the presence of one dominant factor with an eigenvalue exceeding 1, explaining 76.28% of the variance (Table 2). Furthermore, the scree plot (Fig. 1) and parallel analysis supported the one-factor structure of the APP-T. When the two-factor structure was examined, the items loaded on the second factor were more loaded on the first factor, and the second factor’s eigenvalue was low at 0.44. Hence, we determined that the APP-T items represented a single dimension.

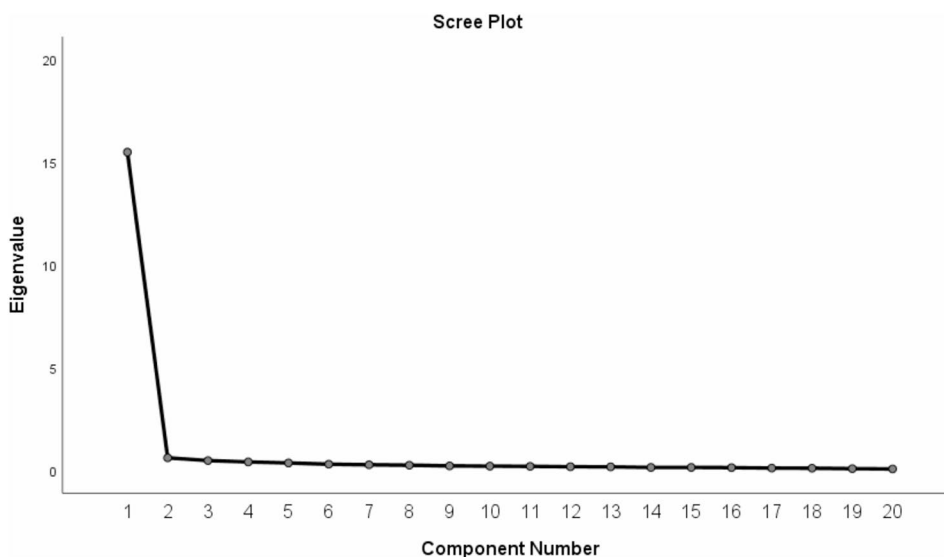


Fig. 1 Scree plot of the APP-T

Table 3 Test-retest and inter-rater reliability of the APP-T total score

	n	First score	Second score	Δ	p	ICC (%95 CI)	SEM	MDC ₉₀
Test-retest reliability	33	58.85 ± 20.89	60.23 ± 23.54	1.37 ± 7.69	0.313	0.94 (0.88–0.97)	1.34	3.11
Inter-rater reliability	19	45.56 ± 13.60	44.16 ± 12.22	-1.39 ± 8.84	0.499	0.77 (0.50–0.91)	2.96	6.86

p: Paired Sample T Test, Δ: the mean difference obtained by subtracting the first score from the second score, SEM: standard error of mean, MDC: minimal detectable change, ICC: intraclass correlation coefficient, CI: confidence interval

Reliability

Total Cronbach’s α (0.935) and the item-total correlations ranged from 0.820 (item 9) to 0.895 (item 7), suggesting good internal consistency (Table 2).

There were no significant differences between the mean total scores of the APP-T obtained for both test-retest and inter-rater reliability assessments (p’s > 0.05). The ICC values for test-retest and inter-rater reliability were 0.94 and 0.77, indicating excellent reliability. The MDC₉₀ values for test-retest and inter-rater reliability were 3.11 and 6.86, respectively, indicating that if the score change between two assessment points is less than the MDC₉₀ values, the main reason for the change is measurement error rather than the actual change of the student. (Table 3). The global score of APP-T showed excellent test-retest reliability (κ_w=0.87; %95 CI=0.73–1.01) and good interrater reliability (κ_w=0.61; %95 CI=0.144–1.08).

The APP-T total scores of the students (n=174) participating in this study ranged from 8 to 80. No floor or ceiling effect was observed as there were no students with a minimum score (0) and 1 (0.57%) student with a maximum score (80).

Acceptability, feasibility, and educational impact

All 7 clinical educators provided feedback concerning the acceptability, feasibility, and educational impact of the

APP-T. All clinical educators agreed with the statements about acceptability. However, considering feasibility, 42.86% of the clinical educators agreed with the statement, “The time spent implementing the APP does not affect my performance in the clinic.” Except for this statement, all statements about the acceptability, feasibility, and educational impact of the APP-T had an agreement rate of more than 70% (Table 4).

Discussion

Clinical education is a crucial process in which students learn how to use theoretical knowledge in clinical settings, resulting in competence in the knowledge, skills, and attitudes required to work as qualified professionals. Therefore, standardized assessment tools with appropriate psychometric and edumetric properties are required to evaluate the clinical performance of physiotherapy students. In contrast to countries like Australia and New Zealand, the assessment of clinical performance of physiotherapy students in Turkey is not addressed by standardized assessment tools. The current assessment of physiotherapy students during clinical placement in Turkey is conducted through various ways such as written and practical examinations, oral presentation, case history-based project, and portfolio. Hence, to contribute towards establishing a standard for evaluating the clinical performance of physiotherapy students in Turkey, the

Table 4 Feedback of the clinical educators on the acceptability, feasibility, and educational impact of the APP-T

	Feedback items	Percentage of agreement	Mean ± SD
Acceptability	The APP-T is a valuable tool to assess students' clinical competence and improve their clinical development.	100%	4.29 ± 0.49
	The APP-T should be used to assess students' clinical competencies and ensure their clinical development.	100%	4.29 ± 0.49
Feasibility	The scoring system of the APP-T is reliable.	100%	4.00 ± 0.00
	The implementation of the APP-T is easy.	71.43%	3.86 ± 1.07
	The time required for the implementation of the APP is reasonable.	85.71%	4.29 ± 0.76
	The information about the implementation of the APP is comprehensive and adequate.	71.43%	4.14 ± 0.90
	The time spent implementing the APP does not affect my performance in the clinic.	42.86%	3.14 ± 1.35
Educational impact	The APP can help to reveal students' weaknesses in their clinical competencies and provide an opportunity to address them.	100%	4.29 ± 0.49
	After the assessment of clinical competence with the APP, it is easy and beneficial to provide feedback to the students.	100%	4.14 ± 0.38
	The assessment of clinical competence with the APP positively affects students' education.	71.43%	3.86 ± 0.69

SD: Standard deviation

current study was conducted to develop a Turkish version of the APP (APP-T) and to examine its psychometric and educational properties. The findings revealed that the APP-T has a single-factor structure and sufficient psychometric properties that confirm it is a reliable and valid tool. Furthermore, it received generally positive feedback from clinical educators regarding the acceptability, applicability, and educational impact of the APP-T.

Internal consistency examines whether the items of a measurement tool are consistent with each other or whether they measure the same concept (in this case, clinical performance). An assessment tool's internal consistency can be examined by item-total correlation and Cronbach α analysis. A Cronbach α coefficient above 0.7 and an item-total correlation coefficient above 0.3

are deemed acceptable [15]. In the current study, Cronbach's α and item-total correlation coefficients indicated that the APP had adequate internal consistency. In the Chinese version study, only the edumetric properties of the APP were examined, yet no psychometric properties were addressed. In the original studies, item analyses were not performed with Cronbach's α and item-total correlation analyses [7, 8, 20]. Therefore, our findings regarding internal consistency were not comparable with previous studies.

Exploratory factor analysis was performed since the APP was translated into Turkish language and its factor structure had not been examined on the Turkish population [21]. Our findings demonstrated that the items of the APP-T represent a single dominant dimension. For the original version, the APP was reported to have one dimension [8] and two dimensions [20]. Different study populations and different statistical analyses used for construct validity may explain the differences in the dimensionality of the APP, where Rasch analysis was used in the original version studies and factor analysis was used in the current study.

Test-retest reliability is an aspect of reliability obtained by re-administering a measurement tool to the same participants, under the same conditions and within a certain period. It, which was not addressed in the original version, was found to be excellent with an ICC value of 0.94 in the current study. However, it should be kept in mind that it is very difficult to blindly measure the same participant by the same assessor [22]. Hence, our findings should be cautiously interpreted by researchers. Furthermore, it may be of greater importance to investigate the inter-rater reliability of performance-based measurement tools like the APP, which means different educators assessing the same students. Consistent with the original version (ICC=0.92) [7], the ICC value for the inter-rater reliability of the APP-T was 0.77 in this study, indicating excellent reliability.

In this study, SEM and MDC₉₀ values, which provide information about the magnitude of measurement error, were also calculated for the reliability of the APP-T. In parallel with the original study (MDC=7.86), the MDC₉₀ for the APP-T total score was found to be 6.86. This MDC₉₀ value showed that a change equal to or higher than 6.86 points was needed to be 90% certain that the change was not resulting from measurement error. Since the scale width of the total score of the APP is 0–80, the MDC₉₀ value represents approximately 9% of this width, which is an acceptable error rate.

In the present study, the edumetric properties of APP-T were also studied. Feedback was collected to learn the views of clinical educators about the acceptability, feasibility, and educational impact of the APP-T. 100% of clinical educators agreed with the statements about APP-T

acceptability. The rate of agreement was above 70% for all statements except one regarding the feasibility of APP-T and for all statements regarding its educational impact. “The time I spend implementing the APP-T does not affect my performance in the clinic”, 42.86% of the clinical educators agreed. The reason for this may be that clinical educators have no experience in using any routine assessment process for evaluating students’ clinical performance. In line with our findings, it was reported in the study of the Chinese version that clinical educators had positive views about the edumetric properties of APP [9].

Strengths, limitations, and future directions

In Turkey, different tools are used to assess students’ clinical performance in physiotherapy entry-level programs. Therefore, we believe that the findings of the current study will encourage the national use of the APP-T as an appropriate, structured, and objective assessment tool for evaluating students’ clinical performance. The use of the APP-T will allow not only national but also international comparisons of students’ clinical performance. In addition, it will also reduce the assessment burden of clinical educators who supervise students from various university programs who prefer different assessment tools.

The current study has several limitations that should be considered in future studies. The study was conducted in a single province, Ankara, which may limit the generalizability of the findings. Thus, studies conducted in multiple provinces are necessary to ensure sample diversity and increase sample representativeness. Some psychometric properties of the APP-T including concurrent validity, predictive validity, and responsiveness were not examined. Furthermore, post-placement feedback on the use of APP-T was based only on close-ended statements. Nevertheless, the current study has shed light on further methodological studies of APP-T.

Conclusions

In Turkey, there is no standardized assessment tool for evaluating students’ clinical performance in physiotherapy entry-level programs. Therefore, this study addresses an important gap by conducting a psychometric and edumetric examination of the Turkish version of the APP. Our findings provide evidence to universities and clinical educators that the Turkish version of the APP has sufficient psychometric and educational properties for evaluating the clinical performance of physiotherapy students.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12909-024-06180-w>.

Supplementary Material 1

Acknowledgements

The authors acknowledge the assistance of Hacettepe, Gazi, and Lokman Hekim Universities and thank the clinical educators and students who participated.

Author contributions

All authors contributed to the study’s conception and design. The first draft of the manuscript was written by Halil I Celik and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

No funding was received for conducting this study.

Data availability

Availability of data and materials: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the Ethics Committee of Lokman Hekim University (April 5, 2023, decision no: 2023/4). All clinical educators and students provided written informed consent to participate in the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 7 April 2024 / Accepted: 11 October 2024

Published online: 17 October 2024

References

1. Physiotherapy W. Physiotherapist education framework. World Physiotherapy London; 2021.
2. Vissers D, et al. Introducing competency-based education based on the roles that physiotherapists fulfil. *J Novel Physiotherapy Phys Rehabilitation*. 2014;1(3):1–6.
3. Gruppen LD, Mangrulkar RS, Kolars JC. The promise of competency-based education in the health professions for improving global health. *Hum Resour Health*. 2012;10(1):43.
4. Fahy A, et al. Evaluating clinical competence assessment. *Nurs Standard*. 2011;25(50):42–9.
5. Kogan JR, et al. Opening the black box of clinical skills assessment via observation: a conceptual model. *Med Educ*. 2011;45(10):1048–60.
6. O’Connor A, et al. Clinical performance assessment tools in physiotherapy practice education: a systematic review. *Physiotherapy*. 2018;104(1):46–53.
7. Dalton M, Davidson M, Keating JL. The assessment of physiotherapy practice (APP) is a reliable measure of professional competence of physiotherapy students: a reliability study. *J Physiother*. 2012;58(1):49–56.
8. Dalton M, Davidson M, Keating J. The Assessment of Physiotherapy Practice (APP) is a valid measure of professional competence of physiotherapy students: a cross-sectional study with Rasch analysis. *J Physiother*. 2011;57(4):239–46.
9. Hu J, et al. Acceptance of the ‘Assessment of Physiotherapy Practice (Chinese)’ as a standardised evaluation of professional competency in Chinese physiotherapy students: an observational study. *BMC Med Educ*. 2020;20(1):108.
10. Beaton DE, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186–91.
11. Mokkink LB, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737–45.

12. Mokkink LB, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010;10:1–8.
13. Dalton M et al. *Development of the APP (Assessment of Physiotherapy Practice) instrument.* 2009.
14. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health.* 2007;30(4):459–67.
15. Field A. *Discovering statistics using IBM SPSS statistics.* SAGE Publications Limited; 2024.
16. Terwee CB, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34–42.
17. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86(2):420.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159–74.
19. Beckerman H, et al. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res.* 2001;10(7):571–8.
20. Reubenson A, Ng L, Gucciardi DF. The Assessment of Physiotherapy Practice tool provides informative assessments of clinical and professional dimensions of student performance in undergraduate placements: a longitudinal validity and reliability study. *J Physiother.* 2020;66(2):113–9.
21. HC de Vet, et al. Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Qual Life Res.* 2005;14(5):1203–18. discussion 1219–21, 1223–4.
22. Norman, Geoffrey, and John Cairney. *Health Measurement Scales: A Practical Guide to Their Development and Use.* Aust NZJ Public Health, Vol. 117, 2015. <https://doi.org/10.1093/acprof:oso/9780199231881.003.0006>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.