İstanbul Med J 2025; 26(3): 178-84
DOI: 10.4274/imj.galenos.2025.72558

A Novel Tool for the Evaluation of Residency Training: Scale Development and Validation

Nazan Karaoğlu[1], Ümmü Doğan[2], Aşir Genç[3], Nur Demirbaş[1], Fatih Yılmaz[4], Hatice Küçükceran[1], Fatma Gökşin Cihan[1], Ümran Münire Kahraman[5]

[1]Necmettin Erbakan University Faculty of Medicine, Department of Family Medicine, Konya, Türkiye
[2]Hasköy Public Hospital, Clinic of Family Medicine, Muş, Türkiye
[3]Necmettin Erbakan University Faculty of Science, Department of Statistics, Konya, Türkiye
[4]Ministry of National Education, Konya, Türkiye
[5]Necmettin Erbakan University Faculty of Political Science, Department of Business Management, Konya, Türkiye

ABSTRACT

Introduction: The purpose of medical specialty training is to provide the physician with sufficient knowledge, skills, attitudes, and qualifications in the chosen field of medicine. In medical training, communication and social opportunities are crucial as well as appropriate goals, programs, and learning environments. The purpose of this paper is to develop a valid and reliable measurement tool to evaluate the compliance of this training with the standards.

Methods: In this survey-type research, in line with the literature, a draft scale to evaluate the training in both the department and the institution consisting of 49 items, was created.

Results: The study involved 497 residents from family medicine residency programs across Türkiye. The exploratory factor analysis revealed that the Scale for Department Evaluation of Medical Specialty Training had four factors whereas the Scale for Institution Evaluation of Medical Specialty Training had three factors. The factorial structures of both scales were confirmed by confirmatory factor analysis. The Cronbach's alpha values were 0.941 and 0.928, respectively. Moreover, both scales explain a high amount of total variance.

Conclusion: The Department Evaluation Scale for Medical Specialty Training, consisting of 25 items and four factors, and the Institution Evaluation Scale for Medical Specialty Training consisting of 20 items and three factors are valid and reliable tools to study medical specialty training.

Keywords: Departments, institutional, scale, residency, training

Introduction

Lately, there have been various studies in the literature, indicating that medical specialty training should be a process structured within certain standards, honor and ethical principles of medicine, shaped by the needs and expectations of society, and certified by objective-based measurements and evaluations (1-3).

The World Federation for Medical Education (WFME) reports three main categories, termed the "trilogy," for quality improvement: basic medical education, postgraduate medical education, and continuing professional development (2). For post-graduate medical education WFME had eight areas to improve: which are "1. mission and values, 2. curriculum, 3. assessment, 4. postgraduate doctors, 5. teachers and clinical supervisors, 6. education and training resources, 7. quality improvement and 8. governance and administration" (2). In medical education, it is necessary to create awareness of the educational environment in the institution for both trainers and trainees, along with the content of the curriculum. How the educational environment is perceived by the trainees plays a key role in determining the quality of learning processes (4). WFME highlights the educational environment as one of the main areas for the evaluation of medical education programs (2,4-10). The Türkiye Board of Medical Specialization Curriculum Development and Standardization also defines educational environments in detail, and highlights the importance of the matter (1). However, the number of publications on how it is evaluated by the trainees is quite small (8,10-13).

The purpose of this paper is to develop a valid and reliable tool that can evaluate and measure the compliance of the education received by

Address for Correspondence: Hatice Küçükceran Assoc. Prof., MD, Necmettin Erbakan University Faculty of Medicine, Department of Family Medicine, Konya, Türkiye
E-mail: drhaticeran@gmail.com **ORCID ID:** orcid.org/0000-0002-0581-9934

Cite this article as: Karaoğlu N, Doğan Ü, Genç A, Demirbaş N, Yılmaz F, Küçükceran H, et al. A novel tool for the evaluation of residency training: scale development and validation. İstanbul Med J. 2025; 26(3): 178-84

Received: 05.07.2024
Accepted: 05.05.2025
Publication Date: 06.08.2025

family medicine specialty students with the standards, the educational and social opportunities, the administrative processes, and the training environment.

Methods

Type of research: Since the purpose of this research is to develop a measurement tool to evaluate the training received by family medicine specialty students, it qualifies as survey research.

The population and sample: the study population consists of family medicine residents who use the internet and social media and are receiving family medicine specialty training in Türkiye on account of the widespread use of social media, the large number of groups involving residents and the chance to reach more participants, the researchers decided to collect the data through the internet.

As the literature recommends that the sample size be at least 10 times the number of questions in the scale, the researchers aimed to reach 490 participants (14). The researchers adopted the convenience sampling method and shared the link to the questionnaire, which briefly contained the research purpose and was created using "Google Docs," on "WhatsApp, Gmail, and Yahoo" groups. Those who followed the link and answered the questions were accepted as voluntary participants.

Ethics statement: Ethical approval was granted by Necmettin Erbakan University, Non-Interventional Clinical Research Ethics Committee (approval number: 2023/4720, date: 15/12/2023). This research was carried out in line with the Declaration of Helsinki. All participants were informed about the research and confirmed that they voluntarily participated in the research.

Data collection tools: The researchers created an item pool of 49 statements inspired by a practitioner tool used in a published thesis, main topics of national and international recommendations, and similar studies (1-3,8,9,11-13,15-17). The internal consistency coefficient of the questionnaire used in the abovementioned thesis was found to be α=0.940 for the departmental evaluation, and α=0.952 for the institutional evaluation in the specialty training; and it was considered to have high reliability (13).

The questionnaire used in this research consists of three sections. The first one collects the socio-demographic information of the participants, the second evaluates training in the department, and the third evaluates training in the institution as noted below in more detail:

Medical specialty training-department evaluation scale: This section has 25 items about the department in which the resident is receiving specialty training, including the conformity of the training to the standards, training opportunities, social relations in the department, qualifications of the trainers, training environment and academic development. The five-point Likert scale has "1- never, 2- rarely, 3- sometimes, 4- often, 5- always" response options indicating the level of agreement with the statements, has no reverse-coded items.

Medical specialty training-institution evaluation scale: This section has 24 items related to interpersonal relations, social facilities, and the educational environment of the institution in which the resident receives specialty training. The five-point Likert scale has "1- never 2- rarely, 3- sometimes, 4- often, 5- always" response options indicating the level of agreement with the statements, and has no reverse-coded items.

Statistical Analysis

The researchers used IBM SPSS AMOS (Version 23) in the data analysis. Exploratory and confirmatory factor analyses for construct validity, and Cronbach's Alpha coefficient to check internal consistency were used and explained in detail in the results section.

Results

With an average age of 29.08±4.1 years, 497 specialty students participated in this research. Among the participants, 70.21% (n=349) were females, 57.33% (n=285) were married, and 30.2% (n=150) had children. 71.4% (n=355) of the participants were satisfied with their city of residence, and 65.24% (n=324) were satisfied with their chosen specialty. The median order of preference for family medicine in the medical specialty exam was 2 (1-57), and 60.91% (n=303) of the participants stated that they would choose family medicine again if they had a second chance.

Medical Specialty Training-Analysis of Department Evaluation Scale

Construct Validity

The researchers employed exploratory and confirmatory factor analyses for construct validity. Kaiser-Meyer-Olkin (KMO) and Bartlett's tests were used to assess whether the data gathered during the pilot study were suitable for factor analysis. The results indicated that the KMO value (0.943) is acceptable (KMO >0.50), which means that the data is suitable for factor analysis (18). The results of Bartlett's test showed that there are significant relationships between the variables and the data are suitable for factor analysis [χ^2: 6587.550; standard deviation (SD): 300; $p<0.001$].

The Exploratory Factor Analysis (EFA)

In the analysis, the researchers used varimax rotation. As a consequence, 25 items with eigenvalues greater than 1.0 were grouped under four factors, explaining 58.68% of the total variance. This value is acceptable according to the literature (19). The factor loadings above 0.30 are acceptable, as the findings show that they exceed this threshold (20,21). Eigenvalues, factor loadings, and the total variance explained are shown in Table 1. As a result of the analysis, the researchers identified the factors related to the departmental evaluation of medical specialty training, taking the items and the relevant literature into account. Accordingly, they called the factors "training" items (1-9), "communication" items (10-18), "scientific activities" items (19-23), and "assessment and evaluation" items (24,25).

The Confirmatory Factor Analysis (CFA)

To confirm the factorial structure, the researchers performed a confirmatory factor analysis. The correlation matrix in Table 2 indicates that the highest correlation is between the "training" and "communication" factors (r=0.781). The path diagram of the scale is given in Figure 1. Moreover, fit indices indicate a good fit to the factorial structure. [χ^2=756.436, df=260, χ^2/df=2.909 root mean square error

of approximation (RMSEA)=0.062, comparative fit index (CFI)=0.923, goodness of fit index (GFI)=0.893, tucker-lewis index (TLI)=0.911 adjusted GFI (AGFI)=0.866] (21-23).

Internal Consistency

The researchers adopted Cronbach's Alpha coefficient to check internal consistency. The coefficients for the factors of the scale are 0.895 for the "training" factor; 0.890 for the "communication" factor; 0.871 for the "scientific activities" factor; and 0.705 for the "assessment and evaluation" factor. As for the whole scale, it is 0.941 (Table 1). Accordingly, the scale is highly reliable (24).

Medical Specialty Training-Analysis of Institution Evaluation Scale

Construct Validity

The construct validity of the scale was tested using exploratory and confirmatory factor analyses. To check whether the data are suitable for factor analysis, the researchers employed KMO and Bartlett's tests. The results indicated that the KMO value (0.935) is acceptable (KMO >0.50), which means that the data are suitable for factor analysis (18). The results of the Bartlett's test also showed that there are significant relationships between the variables and the data are suitable for factor analysis (χ^2: 5185.997; SD: 190; $p<0.001$).

The Exploratory Factor Analysis (EFA)

In EFA, the researchers used the varimax rotation method. They included 24 items in the analysis, although some of them (items 26, 32, 33, 46) were excluded from the analysis since they had high factor loadings on more than one factor with loading differences below 0.10 (25,26). As a consequence, 20 items, with eigenvalues greater than 1.00, were grouped into three factors explaining 57.94% of the total variance. Eigenvalues, factor loadings, and the total variance explained are shown in Table 3.

As a result of the analysis, the researchers identified factors related to the

Table 1. The medical specialty training department evaluation scale-EFA and reliability analysis results

Variable	Factor load	Eigenvalue	Variance (%)	Cronbach's alpha	Corrected item-total correlation
Training		10.613	42.451	0.895	
SB3	0.743				0.743
SB4	0.741				0.701
SB2	0.728				0.706
SB1	0.725				0.701
SB13	0.616				0.664
SB12	0.605				0.672
SB8	0.599				0.558
SB9	0.597				0.646
SB25	0.540				0.546
Communication		1.492	5.969	0.890	
SB6	0.713				0.647
SB19	0.704				0.672
SB21	0.648				0.740
SB5	0.634				0.603
SB23	0.627				0.747
SB7	0.591				0.563
SB20	0.556				0.628
SB24	0.528				0.593
SB22	0.502				0.634
Scientific activities		1.333	5.333	0.817	
SB17	0.767				0.609
SB16	0.766				0.584
SB18	0.560				0.599
SB14	0.557				0.631
SB15	0.556				0.623
Assessment and evaluation		1.233	4.932	0.705	
SB10	0.799				0.545
SB11	0.795				0.545
Total			58.685	0.941	

institutional evaluation of medical specialty training, taking the items and the relevant literature into account (2). Consequently, they called the factors "environment-relationships" items (1-9), "social facilities" items (10-15), and "training environment" items (16-20).

The Confirmatory Factor Analysis (CFA)

To confirm the factorial structure, the researchers did a confirmatory factor analysis. The correlation matrix in Table 4 indicates that the highest correlation is between the "training environment" and "social facilities" factors (r=0.572). The path diagram of the scale is given in Figure 2. In addition, the fit indices to the factorial structure indicate that there is a good fit. (χ^2=539.211; df=161; χ^2/df=3.349; RMSEA=0.069; CFI=0.925; GFI=0.903; TLI=0.912; AGFI=0.873) (21-23).

The Internal Consistency

The alpha coefficients for the factors of the scale are 0.905 for the "environment-relationships" factor, 0.786 for the "social facilities" factor, and 0.837 for the "training environment" factor. Regarding the whole scale, the reliability coefficient is 0.928 (Table 3). In brief, the scale is highly reliable (24).

Discussion

The effects of training processes and training environments on learning are of growing importance. The training environment can exert a positive or negative influence on student motivation. It has been asserted that a comprehensive understanding of the components of the educational environment, held by both students and trainers, can serve as a foundation for establishing arrangements that are conducive to learning objectives (4-10). Thus, the literature review indicates that several papers have studied the quality of training (7-9,11-13,15-17). However, the fact that almost none of those studies used a structured and tested scale has affected the results and generalizability of the studies, and prevented the formation of a common standard. That's why the present

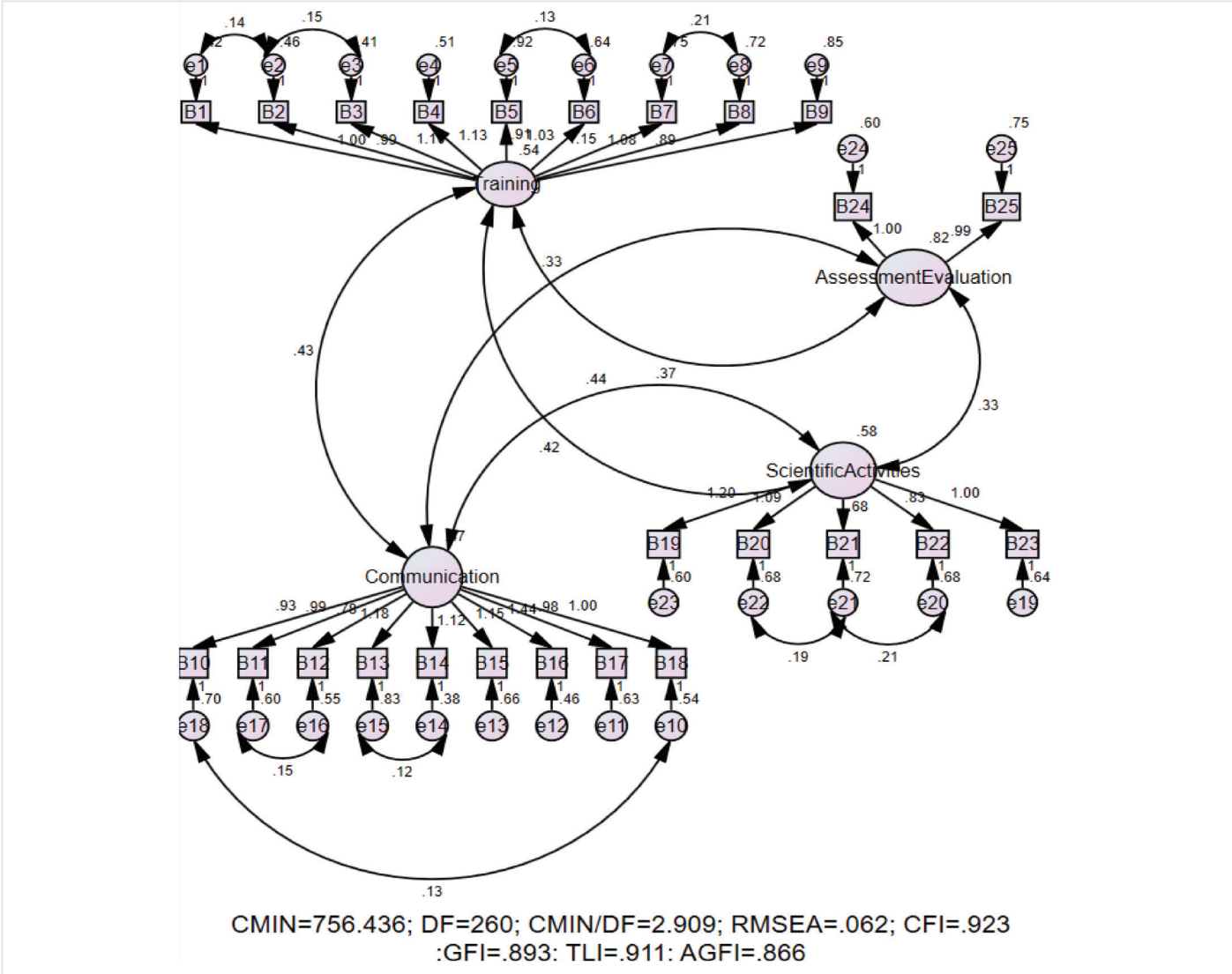

Figure 1. Confirmatory factor analysis for medical specialty training department evaluation scale
RMSA: Root mean square error of approximation, CFI: Comparative fit index, GFI: Goodness of fit index, TLI: Tucker-lewis index, AGFI: Adjusted goodness of fit index

study is of high importance, both nationally and internationally, as it addresses the lack of a measurement tool. In this paper, the researchers developed two scales that can be used separately or together to evaluate medical specialty training. The Medical Specialty Training Department Evaluation Scale, used to evaluate specialty training in the department, consists of 25 items and four factors. Similarly, the "Medical Specialty Training Institution Evaluation Scale" is a valid and reliable tool meant to evaluate the reflections of specialty training in the institution, consisting of 20 items and three factors.

The study involved 497 family medicine residents from educational institutions across Türkiye The data collected from family medicine residents, who are crucial in specialty training and have the opportunity to observe the training process from different perspectives due to their training rotations, are the strength of this paper. Moreover, some expressions, specific to particular medical disciplines, were omitted from the first study, thereby making the scale useful for all medical specialties. Alpha values of the measured scales are 0.941 and 0.928, respectively, which indicates high reliability.

Table 2. Correlations between the factors of the medical specialty education department evaluation scale

		Training	Communication	Scientific activities	Assessment and evaluation
Training	r	1			
Communication	r	0.781**	1		
Scientific activities	r	0.763**	0.751**	1	
Assessment and evaluation	r	0.676**	0.621**	0.742**	1

**Correlation is significant at the 0.01 level (2-tailed)

Table 3. The medical specialty training institution evaluation scale-EFA and reliability analysis results

Variable	Factor load	Eigenvalue	Variance (%)	Cronbach's alpha	Corrected item-total correlation
Environment-relationships		8.620	43.102	0.914	
SK4	0.844				0.773
SK3	0.829				0.759
SK5	0.819				0.782
SK2	0.703				0.712
SK9	0.651				0.619
SK1	0.615				0.641
SK8	0.605				0.699
SK7	0.599				0.685
SK6	0.519				0.635
Social facilities		1.672	8.362	0.786	
SK11	0.704				0.554
SK12	0.683				0.556
SK10	0.682				0.592
SK14	0.578				0.541
SK15	0.575				0.416
SK13	0.509				0.556
Training environment		1.295	6.478	0.837	
SK17	0.761				0.641
SK19	0.759				0.687
SK18	0.741				0.626
SK16	0.665				0.614
SK20	0.559				0.625
Total			57.943	0.928	

EFA: Exploratory factor analysis

Table 4. Correlations between the factors of the medical specialty training institution evaluation scale

		Environment-relationships	Social facilities	Training environment
Environment-relationships	r	1		
Social facilities	r	0.553**	1	
Training environment	r	0.572**	0.390**	1

**Correlation is significant at the 0.01 level (2-tailed)

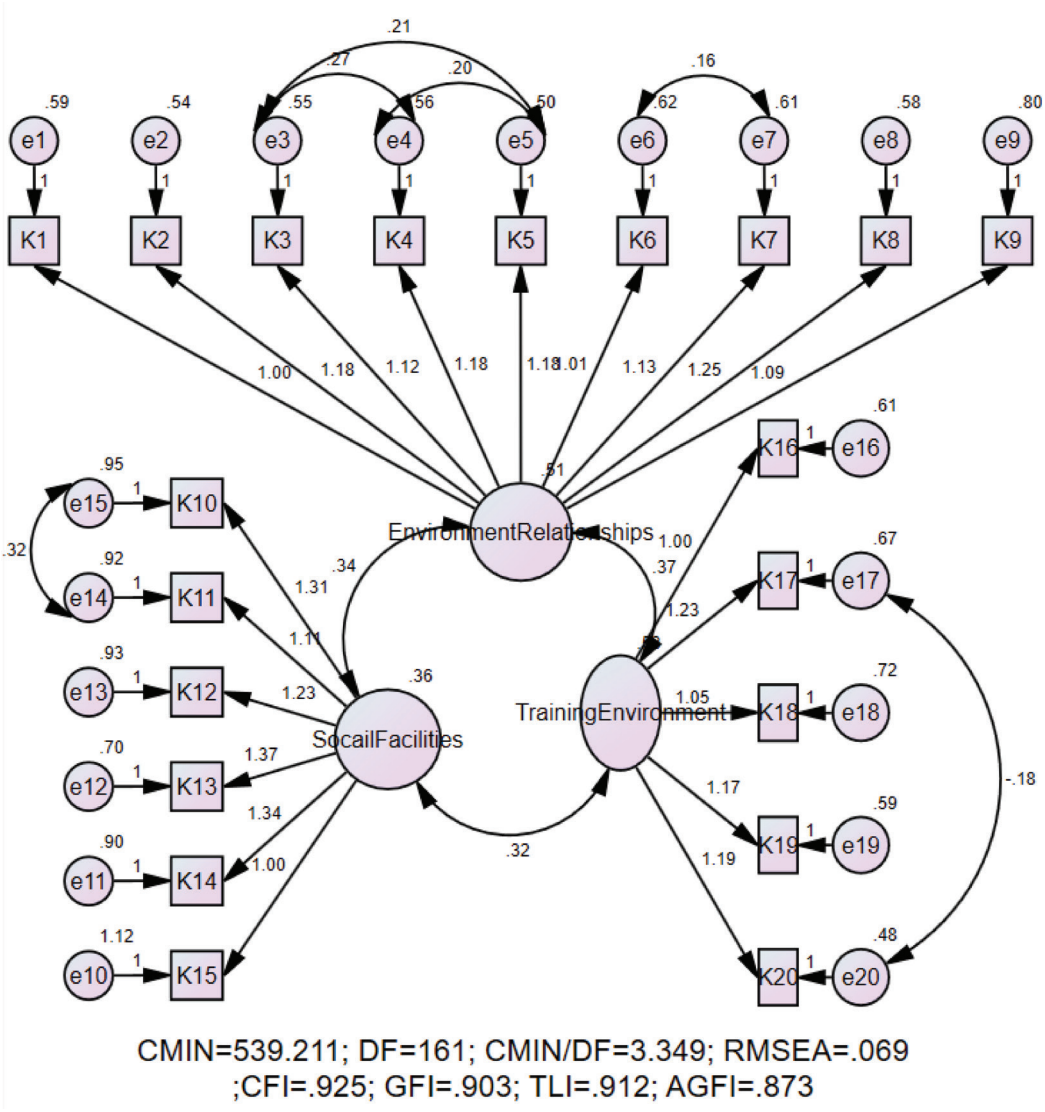

Figure 2. Confirmatory factor analysis for medical specialty training institution evaluation scale
RMSA: Root mean square error of approximation, CFI: Comparative fit index, GFI: Goodness of fit index, TLI: Tucker-lewis index, AGFI: Adjusted goodness of fit index

The WFME emphasizes that in postgraduate medical education it is essential to define the requirements of education and to follow a systematic educational programme that defines the general and discipline-specific components of education. However, in a limited number of studies conducted in Türkiye or in specific hospitals, residents are observed to complain about the limitations of educational and research opportunities, and that service delivery is prioritized over their education (11,12). However, one of the most basic standards set by the WFME and the Board of Specialization in Medicine Curriculum Creation and Standard Setting System (TUKMOS in Turkish), is to prioritize the education of residents and to ensure that the integration of health services is structured according to education (1,2). In fact, one study showed that residents' satisfaction with their education increased with their participation in scientific activities (17). A much larger study in Japan also found that participation in scientific activities increased satisfaction with education (27). We believe that the present study will make an important contribution to this field by providing a scale related to the quality of education, which is emphasized by both the WFME and TUKMOS, but has not yet been measured in our country with a structured data collection tool.

Study Limitations

The voluntary participation, data collection via internet, and study population being limited to only family medicine residents may introduce bias. The scale and items developed in this study cannot be discussed in detail. This is due to the fact that there is no other scale related to the subject for comparison.

Conclusion

In conclusion, the scales are valid and reliable, and can be used to evaluate the training received by the residents. In addition, when

the scales are used together, (department and institution) they provide a thorough evaluation. Moreover, the scales can also be used independently to eliminate the positive or negative effects of the institution on the department or vice versa. It is thought that both scales will become more effective and efficient as they are used in different fields of expertise and institutions. The scales, as used in new research, will be able to evaluate the educational processes more objectively.

Ethics

Ethics Committee Approval: Ethical approval was granted by Necmettin Erbakan University, Non-Interventional Clinical Research Ethics Committee (approval number: 2023/4720, date: 15/12/2023).

Informed Consent: All participants were informed about the research and confirmed that they voluntarily participated in the research.

Footnotes

Authorship Contributions: Concept - N.K., Ü.D., A.G., N.D., F.Y., H.K., F.G.C.; Design - N.K., Ü.D.; Data Collection or Processing - Ü.D., N.D., F.G.C.; Analysis or Interpretation - N.K., A.G., F.Y., H.K., Ü.M.K.; Literature Search - N.K., Ü.D., N.D.; Writing - N.K., Ü.D., A.G., N.D., F.Y., H.K., F.G.C., Ü.M.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. The Board of Specialization in Medicine Curriculum Creation and Standart Setting System. Available at: https://tuk.saglik.gov.tr/Eklenti/34104/0/ailehekimligimufredatv24pdf.pdf. Accessed June 14, 2024.
2. World Federation of Medical Education (WFME). Postgraduate Medical Education. WFME Global Standards for Quality Improvement of Medical Education. Available at: https://wfme.org/wp-content/uploads/2023/03/WFME-Standards-For-Postgraduate-Medical-Education_2023.pdf Accessed June 14, 2024.
3. Royal College of Physicians and Surgeons of Canada (CanMED). Available at: http://www.royalcollege.ca/rcsite/canmeds/canmeds-framework-e. Accessed June 14, 2024.
4. Roff S, McAleer S. What is educational climate? Med Teach. 2001; 23: 333-4.
5. Genn JM. AMEE Medical Education Guide No. 23 (Part 1): Curriculum, environment, climate, quality and change in medical education-a unifying perspective. Med Teach. 2001; 23: 337-44.
6. Genn JM. AMEE Medical Education Guide No. 23 (Part 2): Curriculum, environment, climate, quality and change in medical education - a unifying perspective. Med Teach. 2001; 23: 445-54.
7. Llera J, Durante E. Correlation between the educational environment and burn-out syndrome in residency programs at a university hospital. Arch Argent Pediatr. 2014; 112: 6-11.
8. Alanyalı FM, Geroglu B, Kurnaz MA, Can H, Ongel K. Opinions about residency training program of family physician residents at training and research hospitals in Izmir and their knowledge about family medicine. TJTFP. 2014; 5: 1-7.
9. Al Helal AH, Al Turki Y. Family medicine residents' educational environment and satisfaction of training program in Riyadh. J Family Med Prim Care. 2019; 8: 1330-6.
10. Akdeniz M, Yaman H, Senol Y, Akbayin Z, Cihan FG, Celik SB. Turkish Vasco da Gamma. Family practice in Turkey: views of family practice residents. Postgrad Med. 2011; 123: 144-9.
11. Uzuner A, Topsever P, Unluoglu I, Caylan A, Dagdeviren N, Uncu Y, et al. Residents' views about family medicine specialty education in Turkey. BMC Med Educ. 2010; 15: 29.
12. Tan MN, Ozçakar N, Kartal M. Resident doctors' professional satisfaction and its effect on their lives. MarmaraMedJ. 2012; 25: 20-5.
13. Dogan U, Karaoglu N. The opinions of family medicine residents about education and working conditions in Turkey. Konuralp Medical Journal. 2022; 14: 309-16.
14. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. Educ Psychol Meas. 2013; 76: 913-34.
15. Metsemakers JF. Family medicine training in Turkey: some thoughts. Türk Aile Hek Derg. 2012; 16: 23-34.
16. Oandasan IF, Archibald D, Authier L, Lawrence K, McEwen LA, Palacios M, et al. Giving curriculum planners an edge: using entrance surveys to design family medicine education. Can Fam Physician. 2015; 61: 204-10.
17. Aysan E, Koroglu G, Turkeli V, Özgönül A, Özyaşar A, Gülümser Ç, et al. Resident physicians in Turkey: results of a survey of 1069 residents from 11 provinces. Turk J Med Sci. 2008; 38: 35-42.
18. Guo Y, Zu L, Chen D, Zhang H. Research on influencing factors of satisfaction with the use of public health internet platform: Evidence from Ding Xiang Doctor (DXY) Internet Medical Platform. Int J Environ Res Public Health. 2023; 20: 2276.
19. Hair JF, Black WC, Babin BJ, Anderson RE, eds. Multivariate data analysis. (8th edition), Cengage Learning, EMEA; 2019.
20. Field A. Discovering statistics using SPSS (2nd Edition). London: Sage Publications Ltd. 2005. Available at: https://www.researchgate.net/publication/247826587
21. Shek DT, Yu L. Confirmatory factor analysis using AMOS: a demonstration. International Journal on Disability and Human Development. 2014; 13: 191-204.
22. Cole DA. Utility of confirmatory factor analysis in test validation research. JCCP. 1987; 55: 584.
23. Sarmento RP, Costa V. Confirmatory factor analysis - a case study [Internet]. arXiv preprint arXiv:1905.05598. 2019 [cited 2025 Jul 23]. Available from: https://arxiv.org/abs/1905.05598
24. Sarmento R, Costa V. Comparative approaches to using R and Python for statistical data analysis. 2017
25. Viapiana VF, Filho EJM, Fonseca RP, Giacomoni CH, Stein LM. Development of the arithmetic subtest of the school achievement test second edition. Psicologia: Reflexão e Crítica. 2016; 29: 1-10.
26. Bishop SL, Hus V, Duncan A, Huerta M, Gotham K, Pickles A, et al. Subcategories of restricted and repetitive behaviors in children with autism spectrum disorders. J Autism Dev Disord. 2013; 43: 1287-97.
27. Takahashi O, Ohde S, Jacobs JL, Tokuda Y, Omata F, Fukui T. Residents' experience of scholarly activities is associated with higher satisfaction with residency training. J Gen Intern Med. 2009; 24: 716-20.