Research Article

BIOMEDICAL Journal of Scientific & Technical Research

ISSN: 2574 -1241

DOI: 10.26717/BJSTR.2025.61.009545

# Construct Validation of Teaching Skills Assessment Tool for Clinical Supervision of Postgraduate Residents

**Rizwana Shahid[1]*, Shazia Zeb[2] and Farzana Fatima[3]**

[1]*Assistant Professor Community Medicine, Rawalpindi Medical University, Pakistan*

[2]*Medical Superintendent, Holy Family Hospital, Pakistan*

[3]*Assistant Professor Medical Education, Rawalpindi Medical University, Pakistan*

***Corresponding author:** Rizwana Shahid, Assistant Professor Community Medicine, Rawalpindi Medical University, Rawalpindi, Pakistan

**ARTICLE INFO**

**Received:** March 13, 2025

**Published:** March 20, 2025

**Citation:** Rizwana Shahid, Shazia Zeb and Farzana Fatima. Construct Validation of Teaching Skills Assessment Tool for Clinical Supervision of Postgraduate Residents. Biomed J Sci & Tech Res 61(1)-2025. BJSTR. MS.ID.009545.

**ABSTRACT**

**Background:** Teaching skills assessment tool has recently been designed on getting inspired with Calgary-Cambridge model to assess teaching / supervisory skills of the faculty. Its content validity and inter-rater reliability were established but one of the challenges was that its usage and application in another educational context was unknown.

**Objective:** To establish the construct validity of teaching skills assessment tool in Pakistani context.

**Methods:** A psychometric analytical study was done by getting Teaching skills Assessment tool filled in by 157 postgraduate trainees from 3 teaching hospitals affiliated with Rawalpindi Medical University, Rawalpindi through consecutive non-probability sampling during October – December 2023. Following consent data was collected pertaining to all 16 variables mentioned in the tool and was entered in SPSS version 25.0 for Exploratory Factor Analysis (EFA). Confirmatory Factor Analysis (CFA) was run in Amos 16 software that was followed by calculation of model fit indices.

**Results:** Teaching skills assessment tool was filled by total 157 postgraduate trainees. Kayser Meyer Olkin (KMO) measure confirmed adequacy of sample size. Exploratory factor analysis revealed 2 factors with eigen value >1. Factor 1 and factor 2 had 10 and 6 variables but on elimination of 2 variables from each domain due to negative loading, 8 and 4 items left in factor 1 and 2 respectively. The finalized CFA model was highly reliable with Cronbach alpha of 0.94. However, its construct validity could not be established as only $X^2$/df and Root Mean Square Residual (RMR) were within acceptable range while other model fit indices like CFI, GFI, AGFI, NFI and RMSEA were illustrative of poor model fit.

**Conclusion:** Construct validity of teaching skills assessment tool for its practical utility pertinent to Pakistani culture could not be established.

**Keywords:** Teaching Skills Assessment Tool; Exploratory Factor Analysis; Confirmatory Factor Analysis; Construct Validity; Model Fit Indices

**Abbreviations:** AFI: Adjusted Fit Index; AGFI: Adjusted Goodness of Fit Index; AMOS: Analysis of Moment Structure; BBH: Benazir Bhutto Hospital; CFA: Confirmatory Factor Analysis; CFI: Comparative Fit Index; DF: Degree of Freedom; EFA: Exploratory Factor Analysis; GFI: Goodness of Fit Index; HFH: Holy Family Hospital; H0: Null hypothesis; KMO: Kaiser-Meyer-Olkin; NFI: Normed Fit Index; PCA: Principal Component Analysis; RMR: Root Mean Square Residual; RMSEA: Root Mean Square Error of Approximation; RMU: Rawalpindi Medical University; RTH: Rawalpindi Teaching Hospital; SPSS: Statistical Program for Social Sciences; UAE: United Arab Emirates; X2: Chi-Square

## Introduction

Achievement of desired learning outcomes among students or trainees is substantially attributed to instructional strategies [1]. It is imperative for the medical teachers to be acquainted with different learning styles and ways of adapting to multiple learning situations [2]. Good Medical Practice has been elaborated by General Medical Council (GMC) as provision of ample clinical care, good teaching and training, professionalism, communicating with the patients politely and having reliable interpersonal relationships with colleagues [3]. Supervision of postgraduate trainees at their workplace is of paramount significance in their medical education [4]. Constructive feedback of the supervisors is imperative to promote learning and to enhance competencies of the trainees [5]. A systematic review by Weallans J revealed the consensus of all supervisors on the principles of giving effective feedback for clinical supervision of postgraduate trainees [6]. The supervisors are known to lead the postgraduate trainees in acquisition of clinical knowledge and core competencies by providing a conducive learning environment that in long run seems to be correlated with patient satisfaction [7]. Clinical training is a dynamic process that includes attainment of all relevant skills by engaging in lectures, bedside teaching, performance of procedures, counseling of the patients and undergoing varied assessments [8].

A rapid review by Rothwell C, et al. [9] brought to attention some of the key barriers to effective clinical supervision of trainees like lack of support and leadership from supervisors and lack of mutual trust between supervisor and supervisee [9]. Emphasizing the postgraduate residents to reflect on their clinical practice is also substantially needed for enhancement of their clinical reasoning, professionalism and ethical practice [10]. No doubt, this is an era of competency-based education and numerous tools for measuring the competencies of postgraduate trainees at their workplace have been devised, piloted and validated [11,12]. Certain tools have been designed to measure various domains of clinical supervision across the globe [13]. A tool was developed to critically appraise the clinical teaching skills of supervisors to measure the standard of clinical supervision. Its content validity and inter-rater reliability was established but its construct validity is still unknown [14]. Construct validity is the degree to which a test measures what it intends to measure [15]. Psychometric evaluation of internationally designed tools has frequently been carried out worldwide to ensure their validity in respective culture [16]. The present study is intended to do the factor analysis of a Teaching skills Assessment Tool for computing its construct validity. Construct validation of this tool would make possible its usage in our culture and context with an aim to improve the current practices where necessary.

## Subjects & Methods

A psychometric analytical study was carried out by getting Teaching skills Assessment tool filled in by 157 postgraduate trainees from 3 teaching hospitals affiliated with Rawalpindi Medical University, Rawalpindi. The teaching hospitals were Holy Family Hospital (HFH), Benazir Bhutto Hospital (BBH) and Rawalpindi Teaching Hospital (RTH). Keeping in view the ratio of minimum 5 study subjects per item, it is required to have at least sample size of 80 participants [17]. However, due to fulfilment of inclusion criteria of at least 6 months of postgraduate training, data was gathered from 157 trainees who were doing FCPS / MS /MD training in teaching hospitals affiliated with RMU. Ethical approval for this research was taken from Institutional Ethical Review Committee of Rawalpindi Medical University, Rawalpindi (Ref.No.109/IREF/RMU/2021). Null Hypothesis was that Teaching skills assessment tool is not valid in Pakistani context. Data was collected after getting informed consent from all postgraduate trainees who have completed at least 6 months of their training through consecutive non-probability sampling. Study duration was 3 months (October – December 2023). Permission for using the tool [18] to establish its construct validity was formally sought from one of its inventors Dr. Marie-Claude Audetat through email who is currently working as faculty of Centre Medical University at Geneva. Exploratory Factor Analysis (EFA) was done by using SPSS version 25.0.

This was followed by Confirmatory Factor Analysis (CFA) by means of Amos 16 software. EFA led to calculation of sample adequacy and ruled out the number of domains on Rotated component matrix displaying the items in each domain. After drawing CFA model in accordance with the results of EFA, items with negative loadings were dropped and model fit indices calculated on CFA run were CFI, GFI, AGFI, NFI, X2, RMR and RMSEA.

## Results

Of the 157 trainees participating in research, most (73) belonged to MS training programs while 67 and 17 respondents were enrolled in MD training programs and FCPS-II training respectively. Hospital and specialty-wise distribution of trainees in current study is illustrated below in Table 1. On Exploratory Factor Analysis (EFA), sample size of 157 was found to be adequate on viewing KMO measure as illustrated below in Table 2. Variances of the factors or domains with eigen value greater than 1 are shown below in Table 3. The scree plot shown below also depicts only 2 domains with eigen value greater than 1. Data in EFA was extracted through Principal Component Analysis (PCA) while utilizing Varimax with Kaiser normalization as rotation method. Results of Rotated Component Matrix are shown below in Table 4. On running Confirmatory Factor Analysis (CFA), the items with negative loading were eliminated and the resultant final CFA model is depicted below in Figure 1: The reliability of this CFA model following elimination of items with negative loading is 0.94. The reliability of both the components / domains in CFA model was found to be consistent and hence acceptable as shown below in Table 5. However, this model was found to be poorly fitting in Pakistani context as the resultant major indices were illustrative of poor model fit (Table 6). The value of RMR and RMSEA from Table 4 were descriptive of the acceptable fit and marginal fit of the drawn model respectively. X2/df were also illustrative of acceptable model fit.

**Table 1:** Hospital & Specialty-wise trainees participating in research from each teaching Hospital.

| Specialty | HFH | BBH | RTH | Total |
|---|---|---|---|---|
| Gynaecology / Obstetrics | 11 | 21 | 2 | 34 |
| Paediatrics | 17 | 2 | 0 | 19 |
| Medicine | 8 | 5 | 6 | 19 |
| Anaesthesiology | 8 | 4 | 0 | 12 |
| Gastroenterology | 7 | 0 | 0 | 7 |
| Dermatology | 6 | 0 | 0 | 6 |
| Psychiatry | 0 | 3 | 0 | 3 |
| Radiology | 9 | 8 | 0 | 17 |
| Nephrology | 9 | 0 | 0 | 9 |
| Surgery | 6 | 8 | 3 | 17 |
| Urology | 0 | 9 | 0 | 9 |
| Plastic surgery | 3 | 0 | 0 | 3 |
| Cardiology | 0 | 1 | 0 | 1 |
| Pediatric Surgery | 1 | 0 | 0 | 1 |
| Total | 85 | 61 | 11 | 157 |

**Table 2:** KMO and Bartlett's Test

| | |
|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | 0.843 |
| Bartlett's Test of Sphericity Approx. Chi-Square | 2611.164 |
| Df | 120 |
| Sig. | 0 |
| Determinant | = 2.70 |

**Table 3:** Factors with Eigen value >1.

| Factors | Initial Eigen values | | |
|---|---|---|---|
| | Total | % of variance | Cumulative % |
| Factor 1 | 9.616 | 60.1 | 60.1 |
| Factor 2 | 1.093 | 6.831 | 66.931 |

**Table 4:** Rotated Component Matrix.

| Items | Illustration of items as given in Teaching skills Assessment tool | Factors / Domain | |
|---|---|---|---|
| | | Factor-1 | Factor- 2 |
| Item 10 | Adapting the action plan to patient's psycho-social context or individual perspective | 0.822 | 0.142 |
| Item 14 | New learning | 0.792 | 0.465 |
| Item 13 | States strengths / what is mastered | 0.749 | 0.329 |
| Item 16 | Evaluation of supervision process | 0.739 | 0.483 |
| Item 2 | drives supervision according to the resident's needs | 0.708 | 0.193 |
| Item 15 | What has to be learnt | 0.697 | 0.597 |
| Item 4 | explores the underlying medical knowledge | 0.683 | 0.405 |
| Item 3 | discusses the case and explores clinical reasoning | 0.672 | 0.382 |
| Item 5 | The relevant psycho-social elements and patient's perspective | 0.668 | 0.244 |
| Item 1 | Welcome the resident | 0.658 | 0.366 |
| Item 8 | technical skills (technical procedures, venous puncture, stitching) | 0.322 | 0.854 |
| Item 6 | teachers or corrects history taking or clinical examination | 0.445 | 0.717 |
| Item 7 | interpretation or communication skill | 0.467 | 0.714 |
| Item 9 | discusses the development of an action plan | | 0.704 |
| Item 11 | Verifying the strategies of application | 0.493 | 0.665 |
| Item 12 | addressing own limits of knowledge | 0.515 | 0.564 |

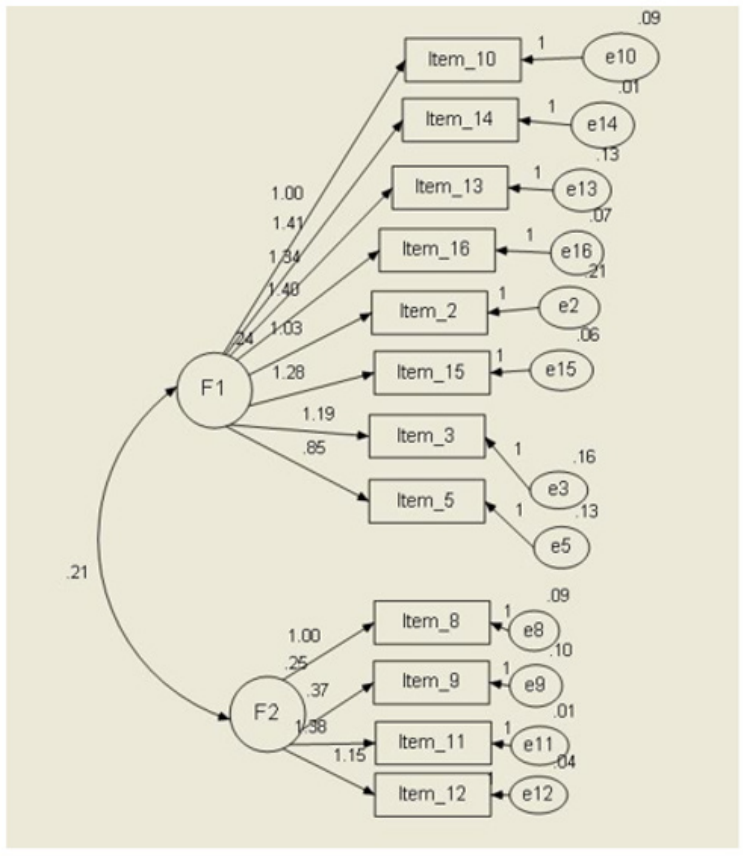


**Figure 1:** CFA model on eliminating the items with negative loadings.

**Table 5:** Reliability of the domains in CFA model.

| Factors | Factor 1 | Factor 2 |
|---|---|---|
| Cronbach alpha | 0.93 | 0.85 |

**Table 6:** Model Fit indices after dropping items with negative loadings.

| Fit indices | Indices of the current model | Yardsticks (Leach, 2008) |
|---|---|---|
| X2/df | 4.4 | <5 |
| RMR | 0.07 | <0.08 |
| AGFI | 0.63 | <0.90 |
| CFI | 0.35 | >0.93 |
| GFI | 0.75 | >0.93 |
| NFI | 0.32 | >0.93 |
| RMSEA | 0.16 | <0.08 |

However, other model fit indices like GFI, AFI and CFI were much less than acceptable fit. In the light of these indices, our null hypothesis was accepted.

## Discussion

Exploratory Factor Analysis (EFA) in the current study revealed the presence of 10 items and 6 items in Factor-1 and Factor-2 respectively in accordance with the highest loading in each factor (Table 4). Results of EFA are also perceived as accurate if each resultant domain has at least 3-5 attributes [17]. Having optimum sample size in accordance with the number of items is deemed necessary to have substantially low standard error [19]. One of the absolute recommendations for running EFA was determined to have sample size of about 200-1000 [20]. However, acceptability of a sample size can also be verified by Kaiser-Meyer-Olkin test that is applied in EFA [21]. KMO measure of sample adequacy in our study is calculated to be 0.843 (Table 2). Its value ranges from 0-1 but higher KMO is illustrative of sufficient sample size. Some studies consider it better to have KMO up to 0.5 while according to others it must be greater than 0.6-0.7 [22]. One of the recommendations is to collect sample size of 30-100 for pilot testing followed by calculation of Cronbach alpha that should not be less than 7 for acceptability of attributes in any domain [23]. The Cronbach alpha in present study measured after dropping the items with negative loading revealed internal consistency of the variables in both

domains (Table 5). Although different studies provide us with varied cut-off levels regarding sample size requirement for having optimal results of EFA, it is rational to have sufficiently large sample for more accuracy and elimination of probable objections.

By mentioning eigen value greater than 1 while running EFA in present study, only 2 domains or factors were sorted out. The retention of factors while finalizing any instrument depends on multiple attributes22, the method applied in current study is kaiser's criteria of having eigen value greater than one along with scree plot (Figure 2). Although some researchers perceive kaiser's method inefficient in sorting out the factors of a tool [24], this method in present study has been coupled with scree plot that is visual representation of eigen values in all the extracted factors [22]. Some scientists undoubtedly believe in scree test as the best choice for determining the factors of a tool [25]; our results computed from both kaiser's eigen method and scree test are same. Lorenzo-Seva U, et al. [19] concluded that it is better to preferably go for exploratory factor analysis before establishing the validity of an instrument. However, EFA in our study was followed by Confirmatory Factor Analysis (CFA) by drawing model in the light of EFA results that illustrated the items or variables in each factor following elimination of items with negative loading (Figure 1) and calculation of certain set of indices to verify the fitness of model (Table 6). Only X2/df and RMR of the shown model were illustrative of model fitness while most of the indices like GFI, AGFI, CFI, NFI and RMSEA were less than the acceptable limit. Hence, these indices were demonstrative of poor fitness of the model. Absolute fit indices include GFI, AGFI, RMR, RMSEA, SRMR while relative fit indices are NFI, NNFI and CFI.

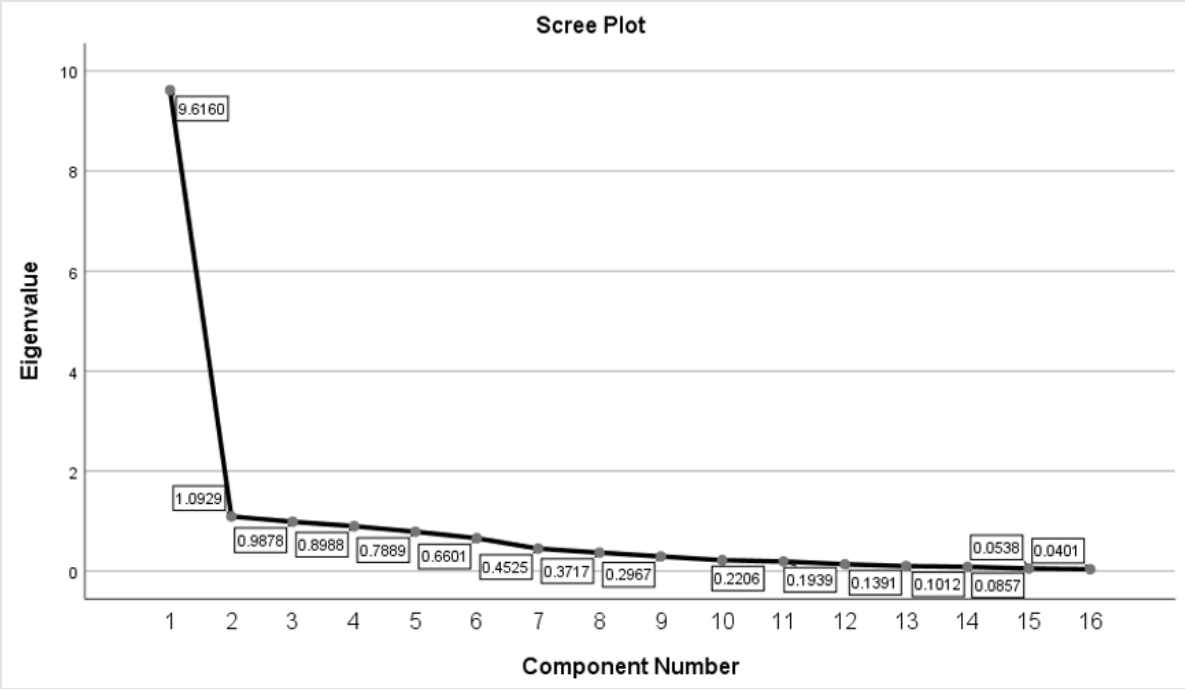


**Figure 2:** Scree plot.

These indices are compared with the given yardsticks for ultimate confirmation about the CFA model to declare it as poor fit, marginal fit or acceptable [26]. As teaching skills assessment tool subjected to factor analysis in present study was designed by researchers of Geneva, Switzerland; the model therefore could not accurately fit in Pakistani culture. Likewise, AMEET inventory was a tool designed originally by doctors of United Arab Emirates to measure educational environment by teachers; its construct validity also could not be established in our country [27]. Likewise, a recent study by Meijer H, et al. [28] revealed that construct and consequential validity of diverse assessment instruments employed in higher education illustrate drastic variation among different cohorts [28]. Like current study, Morad S, et al. [29] did confirmatory factor analysis for establishing construct validity of a tool for measuring innovative thinking capabilities [29]. Construct validity is comprehensive method that utilizes all variables of a tool with significant loadings before giving indices and ultimate conclusion [30]; tools under than factor analysis should be given due consideration in future studies as well for confirmation.

## Conclusion & Recommendation

Construct validity of teaching skills assessment tool could not be established in Pakistani context although it was proved to be highly reliable. It is suggested to use any international tool only after verification of its construct validity in any culture. Although factor analysis is a valid approach towards construct validation; other techniques should also be given due consideration for this purpose.

## Conflicts of Interest

The authors declared no conflict of interest.

## Source of Funding

The author(s) received no financial support for the research, authorship and or publication of this article.

## References

1. Kim S, Raza M, Seidman E (2019) Improving 21st-century teaching skills: The key to effective 21st-century learners. Research in Comparative and International Edu 14(1): 99-117.
2. Ralhan S, Bhogal P, Bhatnagar G, Young J, Green M (2012) Effective teaching skills – How to become a better medical educator. BMJ 344: e765.
3. Papnikitas A, Lunan C (2018) Inside general practice ethics: guidelines 'and' 'of' or 'for' good clinical practice. London Journal of Primary Care 10(2): 34-38.
4. Hore CT, Lancashire W, Fassett RG (2009) Clinical supervision by consultants in teaching hospitals. Med J Aust 191: 220-222.
5. Brown N, Cooke L (2009) Giving effective feedback to psychiatric trainees. Adv Psyhiatr Treat 15: 123-128.
6. Weallans J, Roberts C, Hamilton S, Parker S (2022) Guidance for providing effective feedback in clinical supervision in postgraduate medical education: a systematic review. Postgraduate Medical Journal 98(1156): 138-149.
7. Keshavarzi MH, Azandehi SK, Koohestani HR, Baradaran HR, Hayat AA, et al. (2022) Exploration the role of a clinical supervisor to improve the professional skills of medical students: a content analysis study. BMC Med Educ 22: 399.
8. Sharifi B, Ghafarian Shirazi H, Momeninejad M, Saniee F, et al. (2012) A survey of the quality and quantity of clinical education from the viewpoint of medical students. Pars Jahrom Univ Med Sci 10(2): 57-64.
9. Rothwell C, Kehoe A, Farook SF, Illing J (2021) Enablers and barriers to effective clinical supervision in the workplace: a rapid evidence review. BMJ Open 11(9): e052929.
10. Wilson HMN, Davies JS, Weatherhead S (2016) Trainee therapists' experiences of supervision during training: Meta-synthesis. Clin Psychol Psychother 23(4): 340-351.
11. Sultan AS, Ali R, Shakil S, Khan RN (2020) Workplace based assessment: Tools to assess competencies in a clinical setting. J Pak Med Assoc 71(Suppl 1) (1): S89-S93.
12. Shahid R, Khalid R, Umar M, Khan S (2023) 360-degree evaluation of postgraduate residents at Rawalpindi Medical University: Appraising the achievement of core competencies by comparing the scores of 1st and 2nd cycle of evaluation. JRMC 27(1): 187-194.
13. White E (2018) Measuring clinical supervision: how beneficial is yours and how do you know? J Adv Nurs 74: 1437-1439.
14. Sommer J, Lanier C, Perron NJ, Nendaz M, Clavet D, et al. (2016) A teaching skills assessment tool inspired by Calgary-Cambridge model and the patient-centred approach. Patient Education and Counseling 99: 600-609.
15. Messick S (1995) Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. Am Psychol 50(9):741-749.
16. Ugulu I (2013) Confirmatory factor analysis for testing validity and reliability of traditional knowledge scale to measure university students' attitudes. Educ Res Rev 8(16): 1399-1408.
17. Tucker LR, MacCallum RC (1997) Exploratory factor analysis.
18. Sommer J, Lanier C, Perron NJ, Nendaz M, Clavet D, et al. (2016) A teaching skills assessment tool inspired by the Calgary-Cambridge model and the patient-centered approach. Patient Education and Counseling 99: 600-609.
19. Lorenzo-Seva U, Ferrando PJ (2024) Determining sample size requirements in EFA solutions: A simple empirical proposal. Multivariate Behavioral Research 59(5): 899-912.
20. Goretzko D, Pham TTH, Buhner M (2021) Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. Current Psychology 40(7): 3510-3521.
21. Watkins MW (2018) Exploratory Factor Analysis: A guide to best practice. Journal of Black Psychology 44(3): 219-246.
22. Taherdoost H, Sahibuddin S, Jalaliyoon N (2022) Exploratory Factor Analysis: Concepts and Theory. Advances in Applied in Pure Mathematics 27: 375-382.
23. Kishore K, Jaswal V, Kulkarni V, De D (2021) Practical Guidelines to Develop and Evaluate a Questionnaire. Indian Dermatol Online J 12(2): 266-275.
24. Hayton JC, Allen DG, Vida S (2004) Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. Organizational Research Methods 7(2): 191-205.
25. Costello AB, Osborne J (2005) Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis. Practical Assessment, Research & Evaluation 10(7): 1-9.
26. Shahid R (2020) Use of factor analysis for best practices in behavioral sciences (Editorial). Pak J Med Res 59(2): 43-44.
27. Shahid R, Khan RA, Yasmeen R (2019) Establishing construct validity of AMEET (assessment of medical educational environment by the teachers) inventory. J Pak Med Assoc 69(1): 34-43.
28. Meijer H, Brouwer J, Hoekstra R, Strijbos JW (2022) Exploring Construct and Consequential Validity of Collaborative Learning Assessment in Higher Education. Small Group Research 53(6): 891-925.
29. Morad S, Ragonis N, Barak M (2021) The validity and reliability of a tool for measuring educational innovative thinking competencies. Teaching and Teacher Education 97: 103193.
30. Drew W, Rosenthal Robert (2003) Quantifying construct validity: Two simple measures. Journal of Personality and Social Psychology 84 (3): 608-618.

**ISSN: 2574-1241**

DOI: 10.26717/BJSTR.2025.61.009545

**Rizwana Shahid.** Biomed J Sci & Tech Res



**Submission Link**: https://biomedres.us/submit-manuscript.php