ORIGINAL RESEARCH

Validating the modified System for Evaluation of Teaching Qualities: a teaching quality assessment instrument

Ahmed Al Ansari[1–3]
Mona R Arekat[4]
Abdel Halim Salem[5]

[1]Medical Education, Training and Education Department, Bahrain Defense Force Hospital, Riffa, Kingdom of Bahrain; [2]General Surgery and Medical Education Department, College of Medicine and Medical Sciences, Arabian Gulf University, Manama, Kingdom of Bahrain; [3]RCSI Bahrain, Manama, Kingdom of Bahrain; [4]Internal Medicine Department, Arabian Gulf University, Manama, Kingdom of Bahrain; [5]Anatomy Department, Arabian Gulf University, Manama, Kingdom of Bahrain

Construct: We assessed the validity of the modified System for Evaluation of Teaching Qualities (mSETQ) in evaluating clinical teachers in Bahrain.

Background: Clinical teacher assessment tools are essential for improving teaching quality. The mSETQ is a teaching quality measurement tool, and demonstrating the validity of this tool could provide a stronger evidence base for the utilization of this questionnaire for assessing medical teachers in Bahrain.

Approach: This study assessed the construct validity of this questionnaire in medical schools across Bahrain using 400 medical students and 149 clinical teachers. Data were analyzed using confirmatory factor analysis (CFA). The goodness-of-fit index (GFI), comparative fit index (CFI), root mean square residual, and standardized root mean square error of approximation (RMSEA) indices were used to evaluate the model fit. The internal consistency reliability was assessed using Cronbach's alpha.

Results: The results of the CFA revealed an acceptable fit. All criteria for a good model fit were met except for the RMSEA fit index and the standardized root mean square residual (SRMR) value, which was very close to an acceptable value. Good overall reliability was found in the study ($\alpha=0.94$).

Conclusion: The overall findings of this study provided some evidence supporting the reliability and validity of the mSETQ instrument.

Keywords: clinical teachers, student-centered learning, validity, teacher evaluation

Introduction

The quality of patient care is indirectly dependent on the quality of the training that the physicians receive. This student–teacher rapport necessitates the sharing of knowledge, attitudes, skills, experiences, influences, and interactions in the relationship in an appropriate manner.[1] In the recent years of learner-centered education, teaching hospitals around the world have placed greater emphasis on evaluating their clinical teachers based on their clinical competency, teaching skills, personal qualities, involvement of teachers with the students, involvement of students in the provision of patient care, and the provision of guidance and feedback.[2–4] Past literature reviews have identified many different instruments that have been used to assess clinical teachers.[5–7] The results from these reviews indicate that approximately 30–35 instruments were available in the form of questionnaires that included 1–58 items. The assessments were mainly based on responses from student learners or residents. These assessments were used to provide formative feedback on the students' teaching efficiency, resource allocation, promotions, and performance review.[5–7] Considering the implications of these instruments,

Correspondence: Ahmed Al Ansari
Medical Education, Training and Education Department, Bahrain Defense Force Hospital, Off Waly Alahed Avenue, PO Box 28743, West Riffa, Kingdom of Bahrain
Tel +973 1776 6666 ext 6233
Email drahmedalansari@gmail.com

submit your manuscript | www.dovepress.com
Dovepress
[facebook] [twitter] [linkedin] [youtube]
http://dx.doi.org/10.2147/AMEP.S181094

Advances in Medical Education and Practice 2018:9 881–886

881

they should display high validity and reliability.[8] There are five sources of validity evidence which have been identified by the American Psychological and Educational Research Associations as follows: content, response process, internal structure, relation to other variables, and consequences.[9] These assessments are essential for continuous development and improvement in clinical teaching skills and need to be applied to a wide variety of samples at different points in the learning process.[8] The System for Evaluation of Teaching Qualities (SETQ) was developed and has been used extensively in the Netherlands to assess clinical teachers.[10–13] The SETQ consists of two sets of questionnaires containing the same items: one is for the supervisor's self-evaluation and the other collects the learners' assessments of their clinical teachers.[11] Studies of the SETQ using confirmatory factor analysis (CFA) in teaching hospitals in the Netherlands have concluded that this instrument is a highly reliable and valid tool for assessing physicians' teaching performances.[14–16] This original instrument has been modified and tested for reliability and validity by Al Ansari et al[15] in a teaching school in Bahrain. However, to increase the strength of the evidence regarding the psychometric properties of the instrument, it must be applied in several settings.[8] The modified SETQ (mSETQ) created by Al Ansari et al[15] consists of 25 items in six domains that include the teaching and learning environment, professional attitude toward and support of residents, communication of goals, evaluation of students, feedback, and promotion of self-directed learning.[15] A previous study that evaluated the validity of the mSETQ in Bahrain had a sample size that was too small. Hence, the present study aimed to use a larger sample size to strengthen the validity of this questionnaire for assessing clinical teachers in Bahrain to provide a stronger evidence base for the use of this questionnaire in this specific context.

Materials and methods
The mSETQ instrument

The 25-item mSETQ instrument was previously developed as a modification of the SETQ.[15] The mSETQ is composed of the following six domains: teaching and learning environment (six items), professional attitudes toward students (four items), communication of goals (three items), evaluation (five items), feedback (four items), and promotion of self-directed learning (three items). Each of the items is a statement about which the participants have to identify the extent to which they agree by providing their rating on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree". In this study, this questionnaire was presented several times

in printed handouts, and the participants were shown a list of teachers that they could choose to evaluate. The students wrote the name of the tutor they would like to evaluate on top of each mSETQ questionnaire. They were informed they could evaluate as many (or as few) teachers as they liked using separate questionnaires.

Study population, setting, and data extraction

A total of 400 medical students from three different clinical years and 149 clinical teachers working at four different teaching hospitals in the Kingdom of Bahrain were invited to participate in the study from September 2016 until June 2017. The sample included students from the Arabian Gulf University (AGU) who were from different nationalities including Bahrain, Kuwait, Saudi Arabia, and Oman. The written informed consent to participate was obtained from the clinical teachers, and the verbal consent was obtained from students which was acceptable and approved by research ethical committee at the AGU of Bahrain. The study was conducted in accordance with the Declaration of Helsinki.[16] The study was approved by the research ethics committee at AGU, Bahrain. The students and teachers were aware that their data would be stored anonymously (using numbers rather than names) and that their data might be published. The clinical teachers were working in different departments that included internal medicine, obstetrics, gynecology, pediatrics, surgery, ophthalmology, psychiatry, and family medicine. The printed mSETQ questionnaire was distributed to the students by their clinical coordinators during their rotations in different hospitals and during their in-campus lectures. The students were requested to evaluate as many teachers as they liked (by completing several mSETQ questionnaires) and to return the completed questionnaires to their clinical coordinators. A total of 1,615 completed surveys were received. Incomplete questionnaires were excluded from the study. In total, the number of teachers included in the final analysis was 125, and 1,551 complete questionnaires were used in the analysis.

Statistical analyses

Missing data are common in much educational research, and we used pairwise deletion method to handle the missing data in the analysis. Pairwise deletion only removes the specific missing values from the analysis (not the entire case). Thus, pairwise deletion maximizes all data available by an analysis-by-analysis basis. The strength of this technique is that it increases the power of analyses.

CFA was performed to evaluate the construct validity of the questionnaire. The analysis was conducted using EQS Structural Equation Modeling software (Multivariate software Inc, Broadway, CA, USA). The descriptive and reliability analyses were conducted by IBM SPSS version 23.0 (IBM Corporation, Armonk, NY, USA). In this study, the data were negatively skewed and violated the normality assumption. Therefore, we performed CFA with the robust maximum likelihood estimation (MLE) method.[17] The robust maximum likelihood (ML) estimator provides ML parameter estimates with standard errors and a chi-squared test statistic that are robust to non-normality. The fit indices used to determine the fit of the model were the comparative fit index (CFI), goodness-of-fit index (GFI), adjusted GFI (AGFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR).[18] GFI, CFI, and AGFI values of >0.90 suggested a good model fit.[19,20] An RMSEA value between 0.08 and 0.10 indicated an average model fit, and values below 0.06 demonstrated a good fit. SRMR values below 0.08 demonstrate a good fit.[21]

The internal consistency reliability was assessed using the Cronbach's alpha for each scale. A Cronbach's alpha coefficient value of ≥0.70 was considered as an acceptable reliability.

Results

Descriptive data

The number of student raters varied from 3 to 41 for each clinical tutor. The response rates were 83.8% (N=125/149) for the clinical teachers and 100% (N=400/400) for the clinical raters.

In this study, 3.57% of the total values were missing (9.73% of the cases), and pairwise deletion method was used to handle the missing data. Table 1 describes the students' evaluation scores for each item and the overall score of each scale. The overall scores ranged from 3.76 to 3.90. The scale "promoting self-directed learning" was achieved the highest rating of an average score of 3.90 followed by "professional attitude toward students" scored 3.89 and "communication of goals" scored an average of 3.86.

Construct validity

The CFA results of the six-factor mSETQ are shown in Figure 1. The path diagram shows the standardized regression weights (factor loadings) which explains the pattern of

Table 1 Mean (SD) of students' responses and Cronbach's alpha of different scales of the mSETQ instrument

Scale	Items	Mean	SD	Overall score		Cronbach's alpha
				Mean	SD	
Teaching and learning environment	Q1	3.86	1.105	3.82	0.98	0.939
	Q2	3.81	1.065			
	Q3	3.85	1.043			
	Q4	3.81	1.082			
	Q5	3.8	1.104			
	Q6	3.77	1.128			
Professional attitude toward students	Q7	3.9	1.092	3.89	1.00	0.907
	Q8	3.96	1.09			
	Q9	3.82	1.139			
	Q10	3.87	1.104			
Communication of goals	Q11	3.86	1.075	3.86	1.02	0.933
	Q12	3.87	1.063			
	Q13	3.86	1.089			
Evaluation of students	Q14	3.77	1.106	3.78	1.03	0.961
	Q15	3.79	1.086			
	Q16	3.79	1.0790			
	Q17	3.79	1.112			
	Q18	3.76	1.11			
Feedback	Q19	3.69	1.148	3.76	1.06	0.942
	Q20	3.85	1.12			
	Q21	3.77	1.111			
	Q22	3.72	1.166			
Promoting self-directed learning	Q23	3.88	1.083	3.9	1.04	0.947
	Q24	3.89	1.091			
	Q25	3.92	1.098			

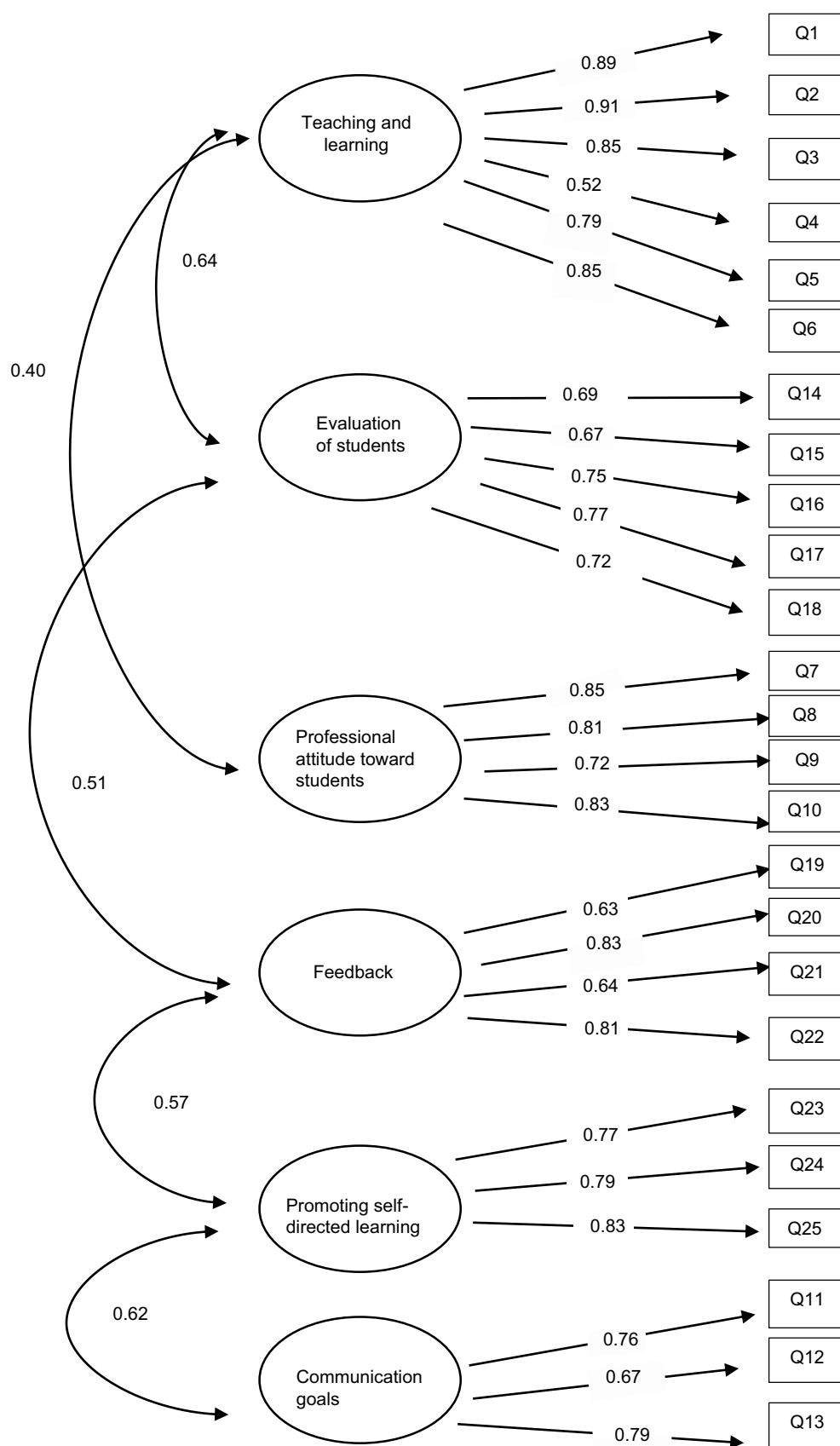**Abbreviations:** mSETQ, modified System for Evaluation of Teaching Qualities; Q, question.

Figure 1 Path analysis of different scales of the teaching quality, teaching and learning environment, professional attitude toward children, communication of goals, evaluation of students, feedback, and promoting self-directed learning.
Abbreviation: Q, question.

item–factor relationships, and all the factor-loading values are >0.40. The results of the CFA suggested an acceptable model fit. All values met the criteria for a good model fit (ie, GFI, CFI, and AGFI >0.90). However, a cutoff value close to 0.08 for SRMR and a cutoff value close to 0.06 for RMSEA are needed for a good model fit. This study reported an RMSEA (0.047) and SRMR (0.079). The chi-squared value was significant: c^2 (385)=670.76, P<0.0001. Table 2 summarizes the values of the fit indices.

Reliability

The overall reliability was measured using the internal consistency reliability (Cronbach's alpha). The overall reliability of the instrument was 0.94 showing excellent reliability. The subscale reliabilities are described in Table 1.

Discussion

The present multicentre study is an attempt to assess the construct validity and internal structure of the mSETQ questionnaire that was used to evaluate the clinical teaching quality of clinical teachers in four different teaching hospitals in Bahrain. Four hundred medical students rated 125 clinical teachers using the questionnaire. All teachers included were involved with clinical teaching of the student raters. The results of this study indicate an overall teaching score between 3.76 and 3.90, which is considered satisfactory. Regarding the individual components, similar scores were observed across all components with teaching and learning environment, communication, and promoting self-directed learning exhibiting the highest values. According to our CFA, the mSETQ with 26 items in six factors was identified as a reliable and valid tool for assessing clinical teaching which showed an acceptable model fit.

The response rate is a major criterion that determines the overall outcome of questionnaire-based research. Previous studies have revealed that the average response rate for data collected from individuals is 40–80%.[22–24] The acceptable response rate is reported to be between 60% and 80%. Considering this rate, the questionnaire tool used should provide high response rates with limited numbers of evaluations, which is an indication of the feasibility of use. This is the greatest advantage of the mSETQ; the response rate of clinical teachers was 83.8% and that of the student raters was 100%.

Limitations

In this study, the learners were allowed to choose their objects of evaluation; this is a limitation as it could lead to selection bias. Future studies should avoid selection bias and ensure appropriate randomization by not allowing the learners to select their objects of evaluation. It is possible that allowing students to choose their objects of evaluation influenced the ceiling effect found in this study. The finding that most teachers scored highly in their evaluations in this study suggests that some alterations must be made to the methodology of the study. Moreover, another limitation of the current study is that the self-evaluation component of SETQ was not incorporated and utilized in the mSETQ. Therefore, future studies should replicate this study and include the self-evaluations as well as aim to reduce selection bias and the ceiling effect.

Conclusion

We provided some evidence supporting the reliability and the validity of the mSETQ tool for evaluating the clinical teaching quality of teachers. This system not only evaluates the current situation but also provides the teachers a scope for improvement in the future based on the feedback they receive. This feedback leads to the identification of learning needs and the development of innovative strategies for teaching and learning. Validity is concerned with meaningful associations, inferences, and their interpretations, and it is of utmost importance to validate any form of an assessment before interpreting it inferences. This notion is important considering the influence provided by the inferences. The validation of the mSETQ before its use in the current study will provide a reliable and meaningful interpretation of clinical teaching quality when the questionnaire is used for assessments in the future. This paper recommends the use of this questionnaire for assessing heterogeneous groups of clinical teachers from different departments in clinical education studies.

Acknowledgment

Table 2 CFA results: the model fit indices

Indices	Coefficient
χ^2/df (c_{min}/df)	χ^2 (385)=670.76, P<0.0001
CFI	0.949
GFI	0.934
AGFI	0.923
RMSEA	0.047
SRMR	0.079

Abbreviations: AGFI, adjusted goodness-of-fit index; CFA, confirmatory factor analysis; CFI, comparative fit index; GFI, goodness-of-fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. *Med Educ*. 2015;49(10):961–973.
2. Woolley S, Emanuel C, Koshy B. A pilot study of the use and perceived utility of a scale to assess clinical dental teaching within a UK dental school restorative department. *Eur J Dent Educ*. 2009;13(2):73–79.
3. Mookherjee S, Monash B, Wentworth KL, Sharpe BA. Faculty development for hospitalists: structured peer observation of teaching. *J Hosp Med*. 2014;9(4):244–250.
4. Biery N, Bond W, Smith AB, Leclair M, Foster E. Using Telemedicine Technology to Assess Physician Outpatient Teaching. *Fam Med*. 2015;47(10):807–810.
5. Fluit CR, Bolhuis S, Grol R, Laan R, Wensing M. Assessing the quality of clinical teachers: a systematic review of content and quality of questionnaires for assessing clinical teachers. *J Gen Intern Med*. 2010;25(12):1337–1345.
6. Lombarts MJ, Heineman MJ, Arah OA. Assessing the quality of clinical teachers. *J Gen Intern Med*. 2011;26(1):14.
7. Beckman TJ, Ghosh AK, Cook DA, Erwin PJ, Mandrekar JN. How reliable are assessments of clinical teaching? A review of the published instruments. *J Gen Intern Med*. 2004;19(9):971–977.
8. Spencer J. Learning and teaching in the clinical environment. ABC of learning and teaching in medicine. *Br Med J*. 2003;326:591–594.
9. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ*. 2003;37(9):830–837.
10. Boerebach BC, Arah OA, Busch OR, Lombarts KM. Reliable and valid tools for measuring surgeons' teaching performance: residents' vs self evaluation. *J Surg Educ*. 2012;69(4):511–520.
11. Lombarts KM, Bucx MJ, Arah OA. Development of a system for the evaluation of the teaching qualities of anesthesiology faculty. *Anesthesiology*. 2009;111(4):709–716.
12. van der Leeuw R, Lombarts K, Heineman MJ, Arah O. Systematic evaluation of the teaching qualities of Obstetrics and Gynecology faculty: reliability and validity of the SETQ tools. *PLoS One*. 2011;6(5):e19142.
13. van der Leeuw R, Lombarts K, Heineman MJ, Arah O. Systematic evaluation of the teaching qualities of Obstetrics and Gynecology faculty: reliability and validity of the SETQ tools. *PLoS One*. 2011;6(5):e19142.
14. Boerebach BC, Lombarts KM, Arah OA. Confirmatory Factor Analysis of the System for Evaluation of Teaching Qualities (SETQ) in Graduate Medical Training. *Eval Health Prof*. 2016;39(1):21–32.
15. Al Ansari A, Strachan K, Hashim S, Otoom S. Analysis of psychometric properties of the modified SETQ tool in undergraduate medical education. *BMC Med Educ*. 2017;17(1):56.
16. World Medical Association. Declaration of Helsinki. *Bull World Health Organ*. 2001;79(4).
17. Bentler P. EQS Structural Equations Program Manual. Encino (Calif.): Multivariate Software; 1995.
18. Harrington D. *Confirmatory Factor Analysis*. Oxford: Oxford University Press; 2009.
19. Kline RB. *Principles and Practice of Structural Equation Modeling*. 3rd ed. New York: Guilford Publications; 2015.
20. Byrne B. *Structural Equation Modeling with AMOS*. Mahwah: Lawrence Erlbaum Associates; 2001.
21. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling A Multidisciplinary J*. 1999;6(1):1–55.
22. Fekete C, Segerer W, Gemperli A, Brinkhof MW; SwiSCI Study Group. Participation rates, response bias and response behaviours in the community survey of the Swiss Spinal Cord Injury Cohort Study (SwiSCI). *BMC Med Res Methodol*. 2015;15(1):80.
23. Eskander A, Freeman J, Rotstein L, et al. Response rates for mailout survey-driven studies in patients waiting for thyroid surgery. *J Otolaryngol Head Neck Surg*. 2011;40(6):462–467.
24. Christensen AI, Ekholm O, Kristensen PL, et al. The effect of multiple reminders on response patterns in a Danish health survey. *Eur J Public Health*. 2015;25(1):156–161.