

RESEARCH NOTE

Open Access



Assessment of validity and reliability of the feedback quality instrument

Sahar Amirzadeh¹ , Davood Rasouli¹ and Helen Dargahi^{1*}

Abstract

Background The purpose was to investigate the psychometric features of the Feedback Quality Instrument (FQI) in medical students, emphasizing the instrument's utility for evaluating the quality of feedback provided in clinical contexts and the importance of performing so for medical trainees.

Methods and material The Persian version of the FQI was evaluated for content validity through a focus group of medical education experts. The questionnaire's face, content, and construct validity were assessed using Confirmatory Factor Analysis, internal consistency, and inter-rater reliability. The questionnaire was revised and pilot-tested, with medical students' feedback in different clinical situations. The data was analyzed using AMOS26.

Results The content validity index equaled 0.88(> 0.79). The content validity ratio representing the proportion of participants who agreed on a selected item was 0.69(> 0.42). According to experts, item 25 is the only modified item, while items 23 and 24 are presented as one item. For reliability, Cronbach alpha was equaled to 0.98.

Conclusions The Persian version of the Feedback Quality Instrument (FQI) was valid, reliable, and fair in assessing feedback quality in medical students, providing valuable insights for other institutions. Establishing a basis for systematically analyzing how certain educator behaviors affect student outcomes is practical.

Keywords Feedback, Feedback quality instrument, Assessment tools, Psychometric evaluation, Validity, Reliability

Background

In the clinical medical education environment, feedback refers to the information that describes the student's performance in a specific activity, which aims to direct their future performance into a relevant activity [1]. Providing feedback to learners for acquiring new knowledge and skills, is important, especially in moving towards competency-based education [2]. Although feedback is a key step in obtaining clinical skills, it is often either deleted or administered in clinical education inappropriately.

This could lead to serious consequences, some of which may be beyond the training course [1]. Evidence showed that most students were dissatisfied with the feedback in clinical education [3]. Sometimes feedback is too vague and generic, making it difficult for students to apply it to improve their work. Detailed, specific, and actionable feedback is crucial for student improvement. However many feedback instances fall short in providing such quality [4]. Hewson describes how teachers in faculty development courses often indicate that their greatest need is learning how to provide effective clinical feedback. This difficulty in providing feedback may be because of an unwillingness to offend or to provoke undue defensiveness [5]. The Feedback Quality Instrument's validity and reliability are crucial for ensuring the accuracy and usefulness of feedback provided to trainees

*Correspondence:

Helen Dargahi
Helen.dargahi@yahoo.com

¹Center for Educational Research in Medical Sciences (CERMS),
Department of Medical Education, School of Medicine, Iran University of
Medical Sciences, Tehran, Iran



© 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/>.

and healthcare professionals. Research revealed a disparity between the recommended practices and what is observed and, in various instances, educators often lead feedback sessions, with learners taking on a more passive role [6]. Studies have developed and validated instruments for assessing specific aspects of feedback, or certain ones were created for particular situations like evaluating by written performance or reflection in simulation-based education [7, 8] or most cases, researchers have investigated the quality of feedback with non-validated tools [3, 9–11].

A valid and reliable feedback instrument should give educators and researchers confidence that they are accurately assessing the quality. This is essential for improving learning outcomes [7]. Additionally, specific and timely feedback linked to learning objectives can significantly enhance student performance and self-regulation [12]. Moreover, a supportive feedback environment can correlate with job satisfaction, leader-member exchange, and reduction in psychological distress [13]. The Feedback Quality Instrument (FQI) provides clinical educators with behaviors designed to encourage a learner to cooperate in performance analysis and design of functional improvement strategies [8]. Considering the lack of suitable learner-centered feedback tools focused on specific behaviors, We decided to psychometrically evaluate the FQI tool in Persian to assess the quality of teachers' feedback by learners.

Methods

Study design

The study involved translating a tool based on WHO steps, assessing face, content, and construct validity, and estimating its internal consistency and inter-rater reliability.

Description of the tool

The feedback quality instrument was designed by Johnson et al. (2021). The creator allowed permission to use it. The questionnaire includes 25 questions in five main areas, consisting of 4 questions for Setting the Scene, 7 for Analyze Performance, 4 for Plane Improvement, 4 for Foster Learner Agency, and 5 for Foster Psychological Safety. The first three domains occur sequentially, and the next two domains flow during a single feedback encounter. The first domain, set the scene, is to ensure a strong start by introducing key factors that will influence the interaction from the outset. The analysis of performance helps the learner gain a better understanding of the desired performance and how their performance is measured. Items in the plan improvement, involve choosing significant learning objectives and creating efficient strategies for enhancement that are customized to the individual. The other two domains develop

throughout feedback, Foster learner agency Encouraging learner autonomy involves involvement, motivation, and active learning. Overall, the tool provides a set of explicit descriptions of helpful behaviors to guide clinical workplace feedback to educators [8].

The options are graded according to the Likert scale.

0=not done, 1=sometimes done, 2=done consistently.

The feedback quality score based on this questionnaire can be between 0 and 50.

Translation

Since the study was conducted in a Persian-speaking country, the English FQI had to be translated into Persian. The questionnaire was first translated into Persian using WHO's 4-step translation methodology. A translator with prior medical training experience and proficiency in the field and interview protocols performed the pre-translation, focusing on translating concepts and asking direct, succinct questions. Following this, a panel of experts reviewed the translation to identify and add any missing words. Subsequently, an independent translator who was unaware of the questionnaire translated the tool into English.

In the second phase, a translation was finalized following testing and interviews. As questionnaire advisors, ten participants were requested to become familiar with the tool and describe their response process. They were also questioned about any offensive difficult terms and their synonyms. The final version of the tool in Persian was then developed, incorporating additional information and the pre-test report.

Validity

After translation, a group of specialists in medical education assessed the Persian version of the FQI for content and face validity.

Face validity

Qualitative face validity was assessed based on expert comments on certain criteria, including associating the questions with the intended concepts, appropriate wording, difficulty, ambiguity, and syntax, appropriateness of question order, and the importance of questions [14]. The tool was revised. Face validity was also assessed using a 3-point Likert scale for each item. The respondents would rate the items if they completely agreed with the item's intended concept 3 and 1 if they disagreed. The impact score, the percentage of participants who rated the importance of items 2 or 3, was higher than 1.5 for all questionnaire items. This showed an acceptable level of validity.

Content validity

The CVR was calculated to assess the agreement of the expert panel on the necessity and usefulness of each item. The CVR values were calculated by applying the method proposed by Lawshe [15]. Also, CVI was calculated to assess for relevance and clarity of each item. The content validity of the items was ascertained by asking 16 medical education experts of Iran University of Medical Sciences (IUMS). Based on their feedback, the questionnaire's difficulty, non-conformity, presence of phrases, and word meaning misunderstandings were determined, through partial question modifications. The experts evaluated the questionnaire by rating each question's importance on a three-part scale (never, usually, and always). It was approved if an item received a score equal to or >0.79 [16].

Construct validity

The pre-final version was tested on 15 medical interns. As per rules of thumb, the number of subjects per variable may vary from 4 to 10. In this study, 5 subjects were calculated [17].

The Persian tool was administered to a 120 medical students to evaluate its construct validity and reliability. The students were asked to complete questionnaires based on feedback from teachers. This study occurred at Ali Asghar Hospital, an academic medical center affiliated with IUMS. Construct validity was assessed through confirmatory factor analysis. In addition, Construct validity was assessed using Chi-squared values. It is worth noting that caution should be exercised when interpreting squared values due to potential inflation associated with large sample sizes and increased freedom. Confirmatory factorial analysis is used between items to assess internal consistency [18]. AMOS version 26 was used for statistical analysis.

Reliability

For measuring the stability of tools over time, the Intra-class Correlation Coefficient [19]. Cronbach's Alpha was calculated to assess for internal consistency. The ICC was used to analyze the internal reliability in SPSS16.

Results

Participants

One hundred and ten feedbacks consisting of 25(22.7%) in morning report sessions, 31(28.2%) in clinical rounds, and 54(49.1%) in other clinical situations were recorded by medical students.

Validity

The content validity ratio of the scale was 0.69, indicating that the necessary and important questions in the tool were mentioned. However, three items, 23, 24, and

25, were modified and merged, so item 25 is the only changed item, and 23 and 24 are presented as one item, according to experts. The content validity index for the test administered to 16 participants equaled 0.88(>0.79). All items were identified except item 24. (Fig. 1)

The instrument showed suitable construct validity. All questions play an effective role in the formation of FQI. The value of the Qi scale was 466.07, which was at the level of $P=0.000$ meaningful. Results showed the Qi Score test rejected the assumption that the analytical model is appropriate for the purpose, the model is not good for the aim. Of course, because of the Key Score test sensitivity to the large sample size, this test by itself cannot support the model properly, so the study used the distribution of the value of the key score on the side of freedom ($\chi^2/df=2.00$), which is less than 5. Could say that it was an effective model. The root means square of the approximation RMSEA was smaller than 0.001(Table 1).

The percentage index, or PCFI index, has a specified limit that is likewise greater than 0.6; in this instance, it was 0.77. IFI raised rate was equal to 0.924 (Table 1). This indicator is appropriate for values greater than 0.9.

The PNFI index, which in this case was equal to 0.725, is a measure of model life and an appropriate indicator for the percentage shown for values above 0.6. A TLI index of 0.909 was attained. (Up to 0.9)

Reliability

The ICC of 15 data was 0.984. The correlation between quantitative variables was assessed using Cronbach's alpha, representing internal consistency equaled 0.981. Cronbach's alpha and ICC for the tool subsections are presented in Table 2. Scores on test subscales strongly correlated with the total test score (correlation coefficient >0.9 ; p -value <0.001).

Total scores on feedback quality in faculty members from the medical students' perspective showed an average of 26.47(± 15.84).

Discussion

Despite the existence of valid instruments for evaluating feedback in medical education, they are underutilized in assessing feedback received by learners. This study evaluated the quality of feedback from medical students' perspectives using the FQI. This approach is rooted in Johnson's assertion that the original version of the FQI is the most comprehensive tool for assessing feedback quality [8].

A 2017 study by Bing-You emphasized the crucial role of feedback in promoting inclusive performance in medical education. The study found that the literature on feedback for learners is broad and predominantly focuses on novel or revised curricula, lacking evidence-based recommendations [2]. In a 2018 study, Bing-You

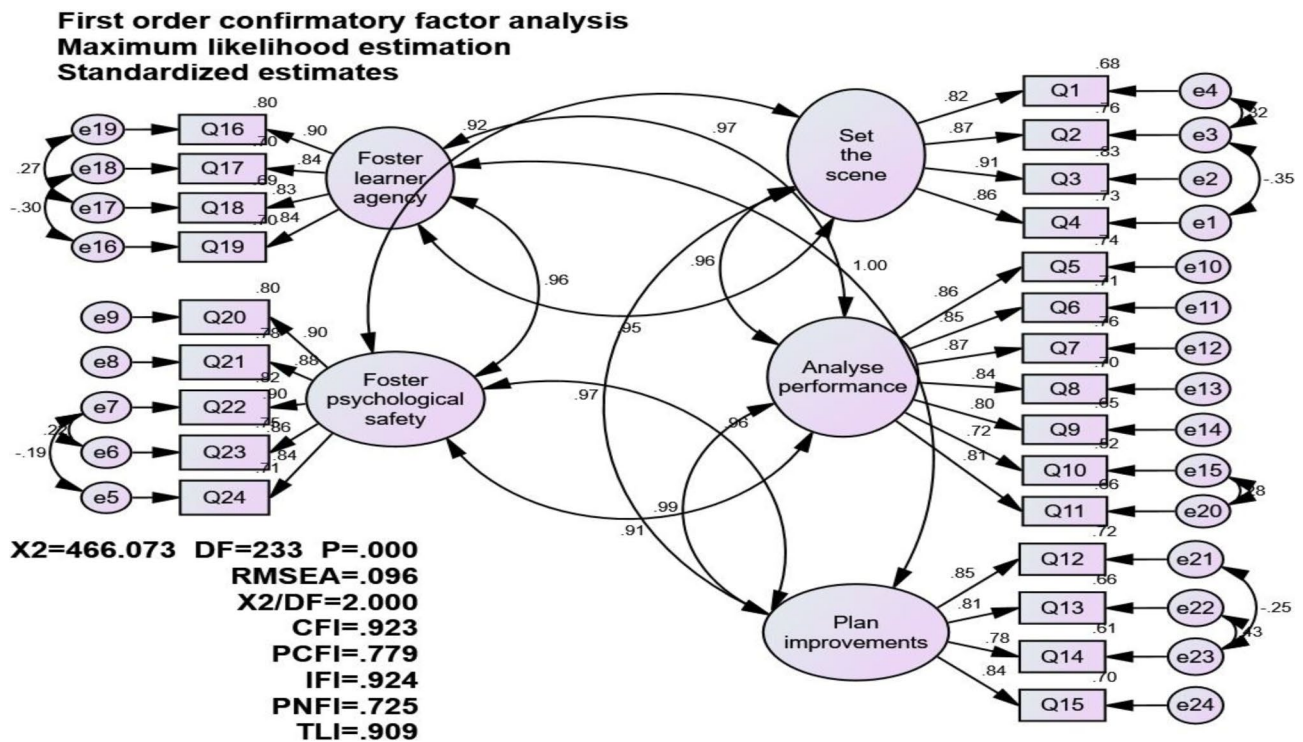


Fig. 1 Factor analysis using standardized coefficient method

Table 1 Measure of construct validity based on confirmatory factor analysis

Medical student	χ^2	χ^2/df	RMSEA	CFI	IFI
	466.073	2.00	0.096	0.923	0.924

Table 2 Internal consistency reliability, Cronbach’s alpha, and Intraclass correlation coefficient (ICC) for the domains of feedback quality

Subscale	Cronbach’s alpha	ICC
Set the scene	0.923	0.864
Analyze Performance	0.936	0.933
plane improvement	0.897	0.812
Foster learner agency	0.910	0.880
Foster psychological safety	0.944	0.940
TOTAL	0.981	0.984

et al. emphasized the importance of feedback in medical education and the need to evaluate students’ perceptions. They developed two new instruments, FEEDME-Culture and FEEDME-Provider. The findings suggest that these tools could enhance the culture of feedback and faculty development in medical education, although further validation is recommended [20]. The FQI provides clinical educators with a defined set of explicit behaviors designed to motivate learners to actively participate in analyzing their performance and devising effective strategies for improvement.

Haghani et al. conducted a descriptive cross-sectional study involving midwifery students at Isfahan University

of Medical Sciences to assess their “perceived feedback” using a researcher-made questionnaire. The study revealed that implementing basic feedback rules was not optimal, indicating a need for instructor training in effective feedback methods [21]. Although the questionnaire measured students’ perceptions of their feedback experiences, it did not indicate which specific behaviors contributed to these perceptions. This highlights the need for a more comprehensive and valid tool with proper reliability. In such scenarios, FQI could more effectively stimulate constructive criticism than subjective impressions of a feedback session’s effectiveness [17].

One of the limitations of this study is that those who completed the feedback quality instrument questionnaire may not have been representative of medical students in general. In addition, the study was conducted in a single center, which may affect the generalizability of the results; therefore, there is a need for multicenter studies with a larger sample size and include perspectives from students in other medical sciences.

Conclusions

FQI simplifies the process for educators to compare actions and expectations, providing clear explanations of standards. This helps educators align their practices with desired outcomes and has the potential to improve clinical performance. Clinical educators can enhance their feedback delivery by asking a colleague to observe their feedback sessions, with the learner’s consent, and

complete the FQI or conducting self-assessments by videotaping their feedback sessions.

Our study found that the Persian version of FQI was valid, reliable, and fair for assessing the quality of feedback provided by faculty members, as perceived by medical students at IUMS.

Abbreviations

CVR	Content Validity Ratio
CVI	Content Validity Index
ICC	Intra-class Correlation Coefficient
RMSEA	Root Mean Score of Approximation
CFI	Comparative Fit Index
IFI	Incremental Fit Index

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13104-024-06881-x>.

Supplementary Material 1: The feedback Quality Instrument (FQI)

Supplementary Material 2: Persian Version of FQI

Supplementary Material 3: Standardised estimates

Supplementary Material 4: Unstandardized estimate

Acknowledgements

The authors are thankful to the Education Development Office of Ali Asghar Hospital and the experts in medical education and medical students at Iran University of Medical Sciences, who dedicated their scarce time to participate in this study.

Author contributions

SA, DR and HD contributed to study conception and design. SA coordinated session planning and teamwork, and was in charge of the translation and data collection phase. DR conducted the data analysis. HD was a coordinator and worked as a translator for the team. All authors read and approved the final manuscript.

Funding

Not applicable.

Data availability

data is provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate

The study adhered to Helsinki's ethical standards for human subjects, with all authors reading and approving the final manuscript. The study received ethical approval from the Iran University of Medical Sciences' institutional ethics board under the registration IR.IUMS.FMD.REC.1402.378, and all students provided informed consent for their participation.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References

1. Ende J. Feedback in clinical medical education. *JAMA*. 1983;250(6):777–81.
2. Bing-You R, Hayes V, Varaklis K, Trowbridge R, Kemp H, McKelvy D. Feedback for learners in medical education: what is known? A scoping review. *Acad Med*. 2017;92(9):1346–54.
3. Tayebi V, Tavakoli H, Armat M, Nazari A, Tabatabaee Chehr M, Rashidi Fakari F, et al. Nursing students' satisfaction and reactions to oral versus written feedback during clinical education. *J Med Educ Dev*. 2014;8(4):2–10.
4. Suraworachet W, Zhou Q, Cukurova M. Impact of combining human and analytics feedback on students' engagement with, and performance in, reflective writing tasks. *Int J Educational Technol High Educ*. 2023;20(1):1.
5. Chowdhury RR, Kalu G. Learning to give feedback in medical education. *Obstetrician Gynaecologist*. 2004;6(4):243–7.
6. Johnson CE, Keating JL, Farlie MK, Kent F, Leech M, Molloy EK. Educators' behaviours during feedback in authentic clinical practice settings: an observational study and systematic analysis. *BMC Med Educ*. 2019;19(1):129.
7. Schlegel C, Woermann U, Rethans J-J, van der Vleuten C. Validity evidence and reliability of a simulated patient feedback instrument. *BMC Med Educ*. 2012;12:1–5.
8. Johnson CE, Keating JL, Leech M, Congdon P, Kent F, Farlie MK, Molloy EK. Development of the Feedback Quality Instrument: a guide for health professional educators in fostering learner-centred discussions. *BMC Med Educ*. 2021;21:1–17.
9. Safaei Koochaksaraei A, Imanipour M, Geranmayeh M, Haghani S. Evaluation of Status of Feedback in Clinical Education from the viewpoint of nursing and midwifery professors and students and relevant factors. *J Med Educ Dev*. 2019;11(32):43–53.
10. Moaddab N, Mohammadi E, Bazrafkan L. The Status of Feedback Provision to Learners in Clinical Training from the residents and medical students' perspective at Shiraz University of Medical Sciences, 2014. *Interdisciplinary J Virtual Learn Med Sci*. 2015;6(1):58–63.
11. Monadi Ziarat H, Hashemi M, Fakhrazadeh L, Akbari Nasagi N, Khazni S. Assessment of Efficacy of oral feedback on trainee's satisfaction in nursing education. *J Nurs Educ*. 2018;7(1):30–7. 2.
12. Arabi E, Garza T. Training design enhancement through training evaluation: effects on training transfer. *Int J Train Dev*. 2023;27(2):191–219.
13. Momotani H, Otsuka Y. Reliability and validity of the Japanese version of the Feedback Environment Scale (FES-J) for workers. *Ind Health*. 2019;57(3):326–41.
14. Lenz ER. Measurement in nursing and health research. Springer publishing company; 2010.
15. Lawshe CH. A quantitative approach to content validity. *Pers Psychol*. 1975;28(4):563–75.
16. Terwee CB, Prinsen CA, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27:1159–70.
17. Kline P. Handbook of psychological testing. Routledge; 2013.
18. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess*. 2003;80(1):99–103.
19. Stone AT, Bransford RJ, Lee MJ, Vilela MD, Bellabarba C, Anderson PA, Agel J. Reliability of classification systems for subaxial cervical injuries. *Evidence-based spine-care J*. 2010;1(03):19–26.
20. Bing-You R, Ramesh S, Hayes V, Varaklis K, Ward D, Blanco M. Trainees' perceptions of feedback: validity evidence for two FEEDME (feedback in medical education) instruments. *Teach Learn Med*. 2018;30(2):162–72.
21. Haghani F, Rahimi M, Ehsanpour S. An investigation of Perceived Feedback in Clinical Education of Midwifery students in Isfahan University of Medical Sciences. *Iran J Med Educ*. 2014;14(7):571–80.

Publisher's Note

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

Received: 6 February 2024 / Accepted: 30 July 2024

Published online: 16 August 2024