Assessing the psychometric questioner for students rating teachers

[1]Razaghi Reza Sharif, [2]Mohammad Reza, [3]Sayyah Mansour and [4]Vakili, Zarichehr

[1]*Assistant Professor, MD, Infectious Disease Dept. Kashan University of Medical Sciences, Kashan, I.R. Iran*
[2]*MD, Department of Pediatrics, Kashan University of Medical Sciences, Kashan. I.R. Iran*
[3]*Affiliated Faculty Member of Education Development Center, Kashan University of Medical Sciences, Kashan, I.R. Iran*
[4]*Department of Pathology, Kashan University of Medical Sciences, I.R. Iran*
Corresponding Email: z_vakili@yahoo.com

_____

ABSTRACT

Educational institutions use different ways to evaluate their teachers. Asking students to rate their teachers is common practice. The purpose of this research was to examine the reliability of the instruments used to evaluate the instructors in a college of medicine. This cross-sectional descriptive research used questioners that evaluated instructors. The questioner was targeting different dimensions of instructors. Item analysis in addition to exploratory factor analysis was performed on 1040questioners answered by the students of the College of Medicine of Kashan University of Medical Sciences. SPSS software was used to perform the analysis. The psychometric properties of questionnaires including Cronbach alpha was determined. The result of exploratory factor analysis and item analysis indicated that three of the subscales of the questioner showed sufficient reliability to evaluate the instructors and two subscales needed further examination. This type of evaluations is necessary to ensure quality of instructors working in an institution as well as providing reliable feedback to the instructors. The result showed that while some subscales of the questioner seems to target the concept of interest; a re-evaluation of the instrument would be valuable to increase its reliability for the administrators in the colleges.

Keywords: Teacher, Evaluation, Student

_____

INTRODUCTION

There are different ways that educational institutions evaluate their teachers. Asking students to rate their teachers is one of them. Educational institutions including medical schools are concerned about the effectiveness of teachers and professors teaching different subjects at basic as well as clinical levels. Since this part of evaluation is a significant part of determining the student achievement in the institution, different parts of teacher affairs including how they test and rate their students are examined [1-5]. In some institutions including Kashan University of Medical Sciences has education development centers (EDC) to examine the performance of the instructors and professors. A common method of evaluating teachers' performance is asking students to rate their teachers [6]. The subject is so important and sometimes contradictory that for many years it has attracted researchers in different fields of education to examine the issue. Student ratings if proven reliable may be used for different purposes including giving feedback to the teachers to improve and revise their teaching method, helps the administrators make decisions about promotion, award, fellowship and scholarship, program revisions, to name a few. However, the method has been open to criticism in one hand and in other hands there has been evidences and research findings that shows the opposite. For instance, several authors have rejected the idea that student rating of

instructors are invalid or biased [7-10]. In this regard, many variables have been the concern of validation and reliability of the instrument for conducting the evaluation. More specifically, student variable aside from other variables in this regard have been examined carefully. Centra (1993) demonstrated that student age has no effect on their rating. Centra (2000) examined the student ratings in a variety of academic disciplines and concluded that despite the presence of some significant difference, they were not likely to have impact on their rating. More interesting has been the student GPA and rating. Davis (2009) concluded that there is little or no relationship between student ratings, GPA, and year in college, citing several authors [14, 15, 9].Undoubtedly, the skills and knowledge of instructors to function as an efficient teacher is a prerequisite condition that contributes to the success of student achievement. However, measuring abstract concepts such as the effectiveness of teaching is relatively complex and requires appropriate conceptual and theoretical frameworks with which to guide the assessment process [17]. Although there is no 'gold standard' or criteria with which to evaluate the quality of a teacher, the rating of teacher quality is usually measured by employing questionnaire that contain different subscales targeted to assess a predefined concept. A close look at the literature reveals that researchers unanimously agree that student ratings are multidimensional and the number of dimensions varies dependingon the form studied and the number and kind of individual items it contains. It is important to note that no single student ratings item or set of related items is useful for all purposes [18]. There have been a number of factor-analytic studies [14, 19] in which as many as 28 different dimensions were derived [7]. However, Centra (1993) and Braskamp and Ory (1994) identified six factors commonly found in student-rating forms: course organization and planning; clarity, communication skills;teacher student interaction, rapport; course difficulty, workload; grading and examination;student self-rated learning. Contrary to these categorization, Marsh's (1984, 2007) employed Students' Evaluations of Educational Quality (SEEQ) form that included nine dimensions including learning/value, enthusiasm, organization, group interaction, and individual rapport, breadth of coverage, exams/grades, assignments, and workload. Other student-rating instruments have items measuring some or all of the above dimensions.

Since teacher evolution is a common process in nearly all education setting, school of medicine in Kashan University of Medical Sciences also follows this kind of evolution partly on the bases of providing feedback to the instructors in regard to some of the dimensions being evaluated. Therefore, this research was designed to examine the questioner used by EDC by students who evaluated their instructors. .

MATERIALS AND METHODS

In this cross-sectional research, 1040 questioners containing 5 subscales including method of instruction (9 questions), subject of study (5), communication with the student(6), educational media use(2) and student evaluation(2) was employed to evaluate the instructors. The students at the College of Medicine of Kashan University of Medical Sciences completed the questioners at the end of education year before the final exams. The questioners were distributed by the administration staff prior to the start of written exams. The psychometric properties of the questionnaires were examined by performing item analysis and Cronbach's alpha was calculated for each subscales. In addition, confirmatory factor analysis was used to find out if there was sufficient evidence to use the questioner for further assessments. SPSS:16 was employed to analyze the questioners.

RESULTS

Separate analysis of data was performed for each subscale of the questioner and inters item correlation in addition to Cronbach alpha was calculated for every subscale. The result of analysis excluded 166 questioners for having missing responses and for the valid 1040 cases, the subscales of instruction showed a Cronbach alpha of 0.791. These results are presented in table 1 and 2.

Table 1:Cronbach alpha for1040 students for the subscales of method of instruction

Cases	N	%	Cronbach's Alpha
Valid	1040	86.2	
Excluded	166	13.8	.791
Total	1206	100.0	

Similar procedure was employed to analyze the subject of study question items. The result of analysis excluded 140 questioners for having missing responses and for the valid 1066 cases, the subscales of subject of study showed a Cronbach alpha of 0.734. These results are presented in table 3 and 4.

___

Inter-Item Correlation Matrix of question items for the subscales of method of instruction

	q1	q2	q3	q4	q10	q14	q15	q16	q17
q1	1.000	.420	.627	.596	.562	.540	.190	.119	.147
q2		1.000	.388	.273	.335	.295	.118	.098	.153
q3			1.000	.597	.548	.561	.226	.156	.190
q4				1.000	.505	.488	.165	.108	.080
q10					1.000	.532	.240	.139	.229
q14						1.000	.253	.181	.211
q15							1.000	.106	.124
q16								1.000	.152
q17									1.000

Table 3:Cronbach alpha for 1066 students for the subscales of subject of study

Cases	N	%	
Valid	1066	88.4	
Excluded	140	11.6	.734
Total	1206	100.0	

Table 4: Inter-Item Correlation Matrix of question items for the subscales of subject of study

	q5	q6	q7	q8	q9
q5	1.000	.448	.285	.220	.200
q6		1.000	.425	.385	.347
q7			1.000	**.780**	.213
q8				1.000	.251
q9					1.000

Further analysis was employed to examine the subscale of communication with the student. The result of analysis excluded 147 questioners for having missing responses and for the valid 1059 cases, the subscales of communication with the student Cronbach alpha of 0.601. These results are presented in table 5 and 6.

Table 5:Cronbach alphafor 1059 students forthe subscales of communication with the student

Cases	N	%	
Valid	1059	88.7	
Excluded	147	12.2	.601
Total	1206	100.0	

Table 6:Inter-Item Correlation Matrix of question items for the subscales of communication with the student

	q18	q19	q20	q21	q22	q23
q18	1.000	.117	.140	.244	.104	.193
q19		1.000	.202	.332	.098	.165
q20			1.000	.211	.130	.209
q21				1.000	.221	.370
q22					1.000	.243
q23						1.000

Table 7: Cronbach alpha for 1073 students for the subscales of educational media use

	N	%	Cronbach's Alpha
Valid Cases	1073	89.0	
Excludeda	133	11.0	.398
Total	1206	100.0	

Further analysis was performedon1073 valid cases to examine the subscale of educational media use and student evaluation of their instructors. The result of analysis showed a Cronbach'salpha of 0.398 and 0.428 for educational

media use and student evaluation of their instructors, respectively. These results are presented for educational media use in table 7, 8 and for student evaluation in table 9and 10, respectively.

Table 8:Inter-Item Correlation Matrix of question items for the subscales of educational media use

	q12	q13
q12	1.000	.249
q13	.	1.000

Table 9: Cronbach alpha for 1073 students for the subscales of educational student evaluation

Cases	N	%	Cronbach's Alpha
Valid	53	4.4	
Excluded	1153	95.6	.428
Total	1206	100.0	

Table 10:Inter-Item Correlation Matrix of question items for the subscales of student evaluation

	q24	q25	Cronbach's Alpha
q24	1.000	.272	0.398
q25	.272	1.000	

The result of explorative factor analysis also indicated that there was indeed five different subscales within the questioner with the variances greater than 5 percent including 33.09, 6.61, 5.57, or nearly approaching (4.52 and 4.37), respectively.

DISCUSSION

Assessing is an integral part of education evaluation. The purpose of this research was to examine the characteristics of evaluation forms employed in the college of medicine in Kashan University of Medical sciences. These forms are used to collect information in regard to the instructors teaching in this college. Any questioner that is not validated by psychometric procedure will not provide useful information to base prediction about the success or failure of concerned matter. Therefore, the first step is to perform an item analysis to check every single question making up the concept of interest. Assessment of abstract phenomenon requires clear cut definition based on which the result by different evaluator show close approximation. Thus, one has to rely on very logical theoretical frameworks with which to guide the assessment process [17]. The subject of students rating teachers is an important matter; however, what needs to be assessed remains a controversial issue. While some researchers include over30 subscales, there are others who limit the subscale to 5 [19]. In summary, Students' Evaluations of Teaching Effectiveness Rating Scale (SETERS) has been useful in assessing teaching effectiveness in Western Countries [25]. For the sake of making the questioner short and easy for students to complete the form, number of questions within the scale are kept to minimum after a careful item analysis is performed on the student responses. Item analysis is a necessary first step to test the reliability of an affective or cognitive test. By measuring the mean correlation between items within sub scales and overall score, Cronbach alpha is determined. George and Mallery (2003) provide the following rules of thumb: "_ > .9 – Excellent > .8 – Good > .7 – Acceptable > .6 – Questionable > .5 – Poor, and_ < .5 – Unacceptable" (p. 231).The Cronbach's alpha of 0.791, 0.734, 0.601,0.398 and 0.428 for the quality of instruction , subject of study, communication with the student, educational media use, and student evaluation of their instructors were observed for the student evaluating their teachers. According to the most reference values suggested for this purpose, the quality of instruction, subject of study, and communication with the student items were appropriate questions targeted to assess the concept of concern. However, items making up the subscales of educational media use, and student evaluation of their instructors were poor items as they also showed close variance to 5 percent which is usually a cut-off point to distinguish a factor. These two subscales seemed to need more validation to use in the future assessment.

CONCLUSION

In conclusion, the result showed that a partly valid and reliable measure to evaluate the teachers working in the college of medicine has been employed with acceptable psychometric properties tested by computer SPSS:PC software for three of the subscales of the quality of instruction, subject of study, and communication and poor internal reliability for educational media use and student evaluation of their instructors . A reevaluation of the instrument would be valuable to increase its reliability for the administrators in the College of medicine.

_____

REFERENCES

[1] Sanju Gajjar, Pradeep Kumar, Manish Rana, Rashmi Sharma, Item and test analysis to identify quality multiple choice questions (MCQS) from an assessment of medical students of Ahmedabad, Gujarat, Indian Journal of Community Medicine, Vol. 39, No. 1, January-March, 2014, pp. 17-20

[2] Sarin YK, Khurana M, Natu MV, Thomas AG, Singh T. Item analysis of published MCQs. Indian Pediatr 1998;35:1103-5 [PUBMED]

[3] Mohammad Resa Sharif, Seyed Mohsen Rahimi, Mehdi Rajabi, Mansour Sayyah, Computer Software Application in Item Analysis of Exams in a College of Medicine, ARPN Journal of Science and Technology, 2014, 4 (10), 565-569

[4] Sharif MR, Asadi MH, Sharif AR, Sayyah M, Examining the Multiple Choice Educational Examinations of College Students, Biomedical & Pharmacology Journal 2013, 6 (1), 23-27

[5] SayyahM, VakiliZ, AlaviNM, BigdeliM, SoleymaniA, AssarianM, .An Item Analysis of Written Multiple-Choice Questions: Kashan University of Medical Sciences, Nursing and Midwifery Studies, 2012, 1 (2), 83-7

[6] Siabani S., Moradi M., Siabani H,. education problems in Kermanshah University of Medical Sciences from the prospect of medicam student, 2007, Behbopd Science quarterly research 13 summer No. 2 p-162-67

[7] Feldman, K. A. (2007). Identifying exemplary teachers and teaching: Evidence from student ratings. In R. P. Perry & J. C. Smart (Eds.), The Scholarship of teaching and learning in higher education: An evidence-based perspective (pp. 93-129). Dordrecht, The Netherlands: Springer.

[8] Kulik, J. A. (2001). Student ratings: Validity, utility, and controversy. In M. Theall, P. C. Abrami, & L. A. Mets (Eds.), The student ratings debate: Are they valid? How can we best use them? New Directions for Institutional Research, No. 109 (pp. 9-25). San Francisco: Jossey-Bass.

[9] Svinicki, M., & McKeachie, W. J. (2011). McKeachie's teaching tips: Strategies, research, and theory for college and university teachers (13th ed.). Belmont, CA: Wadsworth.

[10] Theall, M., & Feldman, K. A. (2007). Commentary and update on Feldman's (1997) "Identifying exemplary teachers and teaching: Evidence from student ratings." In R. P. Perry & J. C. Smart (Eds.), The teaching and learning in higher education: An evidence-based perspective (pp. 130-143). Dordrecht, The Netherlands: Springer

[11] Centra, J. A. (1993). Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness. San Francisco: Jossey-Bass

[12] Centra, J. A., & Gaubatz, N. B. (2000). Is there a gender bias in student evaluations of teaching? Journal of Higher Education, 70, 17-33.

[13] Davis, B. G. (2009). Tools for teaching, (2nd ed.). San Francisco: Jossey-Bass

[14] Abrami, P. C., & d'Apollonia. (1990). The dimensionality of ratings and their use in personnel decisions. In M. Theall & J. Franklin (Eds.), Student ratings of instruction: Issues for improving practice: New Directions for Teaching and Learning, No. 43 (pp. 97-111). San Francisco: Jossey-Bass.

[15] Marsh, H. W. (2001). Distinguishing between good (useful) and bad workloads on student evaluations of teaching. American Educational Research Journal, 38, 183-212.

[16] McKeachie, W. J. (1997). Student ratings: The validity of use. American Psychologist, 52, 1218-1225.

[17] Bullock AD, Belfield CR, Butterfield S, Morris ZS, Ribbins PM, Frame JW. A, Continuing education courses in dentistry: assessing impact on practice, Issue , Medical Education, Volume 33, Issue 7, pages 484–488, July 1999

[18] Benton, S. L., Webster, R., Gross, A., & Pallett, W. (2010a). IDEA Technical Report No. 16: An analysis of IDEA Student Ratings of instruction using paper versus online survey methods. Manhattan, KS: The IDEA Center.

[19] Hoyt, D. P., & Lee, E. (2002a). IDEA Technical Report No. 12: Basic data for the revised IDEA system. Manhattan, KS: The IDEA Center.

[20] Feldman, K. A. (2007). Identifying exemplary teachers and teaching: Evidence from student ratings. In R. P. Perry & J. C. Smart (Eds.), The Scholarship of teaching and learning in higher education: An evidence-based perspective (pp. 93-129). Dordrecht, The Netherlands: Springer.

[21] Centra, J. A. (1993). Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness. San Francisco: Jossey-Bass.

[22] Braskamp, L. A, & Ory, J. C. (1994). Assessing faculty work: Enhancing individual and institutional performance. San Francisco: Jossey-Bass.

[23] Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. Journal of Educational Psychology, 76, 707-754.

---

[24] Marsh. H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), The Scholarship of teaching and learning in higher education: An evidence-based perspective (pp. 319-383). Dordrecht, The Netherlands: Springer.

[25] Bamidele A., and Olajide A., The African Symposium , Volume 12, No. 2, December 2012

[26] George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide andreference. 11.0 update (4th ed.). Boston: Allyn & Bacon.