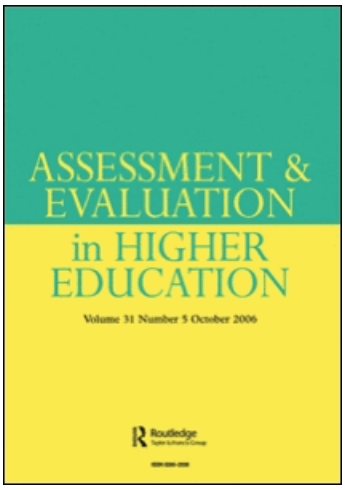
This article was downloaded by: *[B-on Consortium - 2007]*
On: *4 September 2008*
Access details: *Access Details: [subscription number 778384760]*
Publisher *Routledge*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Assessment & Evaluation in Higher Education
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713402663

A questionnaire for listening to students' voices in the assessment of teaching quality in a classical medical school

Maria Filomena Gaspar [a]; Anabela Mota Pinto [a]; Hugo Camilo F. da Conceição [a]; José António Pereira da Silva [a]
[a] University of Coimbra, Portugal

Online Publication Date: 01 August 2008

To cite this Article Gaspar, Maria Filomena, Pinto, Anabela Mota, da Conceição, Hugo Camilo F. and Pereira da Silva, José António(2008)'A questionnaire for listening to students' voices in the assessment of teaching quality in a classical medical school',Assessment & Evaluation in Higher Education,33:4,445 — 453

To link to this Article: DOI: 10.1080/02602930701562932

URL: http://dx.doi.org/10.1080/02602930701562932

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Assessment & Evaluation in Higher Education
Vol. 33, No. 4, August 2008, 445–453

Routledge
Taylor & Francis Group

RESEARCH NOTE

A questionnaire for listening to students' voices in the assessment of teaching quality in a classical medical school

Maria Filomena Gaspar*, Anabela Mota Pinto, Hugo Camilo F. da Conceição and José António Pereira da Silva

University of Coimbra, Portugal

The purpose of this study was to develop a teaching quality assessment questionnaire and assess its reliability by using it with a sample of first-year medical students. Principal components analysis with varimax orthogonal rotation resulted in the development of a 12-item, two-component tool, adequate for use in lectures and small-group sessions. The two components were named 'curriculum' and 'relationship'. The Cronbach coefficient alpha values indicated high reliability and internal consistency. According to the results obtained this teaching quality scale is a reliable measure and may be useful in identifying themes in disciplines and among teachers that may benefit from some professional development. Amongst its advantages is that it can be used with an optical reading tool.

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008

Introduction

The assessment of teaching quality and the pedagogical training of teachers are two of the priorities of medical education as a whole and, in particular, of the Undergraduate Section of the Department of Medical Education of the University of Coimbra (DEM-PG) in Portugal.

The medical teachers as teachers, researchers and managers have to assume different roles that will require specific training. As teachers, their goal has to be an increase in the quality of teaching in order to enhance the quality of students' learning. Teachers have to reflect on the curriculum they implement and/or develop, and assume responsibility regarding their self-assessment and subsequent training, as a result of the identified needs. In this manner, teachers become managers of curriculum practice and assessment/marking processes, with the purpose of implementing a quality curriculum and teaching (Dent and Harden 2005). In order to achieve this goal, teachers need to be assessed. Among the internal methods of assessment are self-assessment, assessment by students and peer-based assessment.

The definition of assessment is extremely complex, not only because it involves multiple objectives, subjects and processes, but also because it is intimately linked to the mission of the institution and learning objectives. Within this context, we share the following definition, proposed by Palomba and Banta (1999, 4): 'assessment is the systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development'. These same authors propose a set of strategies, which they term 'assessment essentials', considered to be the central

*Corresponding author. Email: ninigaspar@fpce.uc.pt

ISSN 0260-2938 print/ISSN 1469-297X online
© 2008 Taylor & Francis
DOI: 10.1080/02602930701562932
http://www.informaworld.com

ingredients needed for an assessment programme to be successful. These ingredients include listening to the 'voices' of the students, the existence of active involvement of the faculty and/or the different departments in the whole process, as well as the selection or development of appropriate assessment instruments. They emphasize, however, that it is not enough to collect data. We need to examine data, share them and act accordingly, with the purpose of improving the quality of teaching and learning. This means that assessment has to drive the actions taken at the department, faculty, teachers' and students' levels, otherwise it will not develop into a useful process in relation to the efficacy of teaching and learning.

With the aim of constructing an instrument that proved useful in assessment—from a student perspective – of the teaching quality within a discipline (from several aspects), we developed and applied a survey in the discipline of 'Pathophysiology', which is included in the medical undergraduate course. The assessment of teaching and learning was done through the opinions of the students, in the classroom. The assessment we have conducted is included in CATs (Classroom Assessment Techniques), within the category of techniques that intend to assess students' opinions when faced with specific aspects of teaching (cf. Palomba and Banta 1999, 167–177).

A pivotal aspect to consider in controlling the quality of teaching is the relation between learning outcomes/lesson contents and the teaching strategies employed, in such a way that we are able to train doctors capable of responding to the challenges and expectations posed by the twenty-first century (European University Association 2002). Teachers are, then, the anchors of this development, and the monitoring of their pedagogical competences, which is usually not assessed, becomes an essential step in the process.

The students need to be involved in this assessment of quality because they are part of the process and the target of our working efforts. The importance of 'listening to the students' voices' in any assessment proceeding is an unquestionable strategy (cf. Crespo 1993; Van Vught 1995).

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008

Methods

Questionnaire development

Our objective was to develop a short quality assessment instrument for students to complete in relation to lectures and small-group sessions, which could be adopted in the future by all of the Faculty of Medicine of the University of Coimbra (FMUC) teachers.

Using the criteria (objectives; content; materials; teaching methods; assessment measures; presentation; preparation) proposed by Fincher et al. (2000) for assessing the quality of teaching we developed a set of 16 original statements. Those statements were conceptually grouped in the following curricular and pedagogical categories:

(1) Learning Objectives and Lesson Contents
- Item 1. The relation between content/subjects in the lesson and the general learning objectives for this discipline were clear.
- Item 2. The content/subjects taught in this lesson were interesting.
- Item 3. The contents developed in this lesson had already been taught in the context of other disciplines.
- Item 4. The contents presented are an important contribution to my training as a doctor.
- Item 5. I understood the content presented.

(2) Methods and Strategies
- Item 6. The teaching strategies the teacher used were the most adequate.
- Item 7. The audiovisual media used contributed to the quality of the teaching and learning process.
- Item 8. The lesson was well planned (organization; sequence).
- Item 9. The lesson's duration was adequate to the content and objectives.

(3) Teacher
- Item 10. The teacher promoted the active involvement of the students in the class.
- Item 11. The teacher stimulated the critical reasoning of the students.
- Item 12. The teacher promoted self-directed learning of the students outside the class.
- Item 13. The teacher created the necessary conditions for the clarifying of the students' doubts.
- Item 14. The teacher was punctual.

(4) Student Satisfaction
- Item 15. I was pleased with the lesson quality.
- Item 16. I was pleased with the quality of the teacher.

Each item containing a five-point Likert-type scale response format (1 = I totally disagree; 2 = I disagree; 3 = I agree; 4 = I agree, although not totally; 5 = I totally agree), following Palomba and Banta's (1999) suggestion of using Likert scales to discover students' attitudes and opinions.

Every item is formulated in a positive form, with the exception of item 3 that has to be recoded before proceeding to the calculation of the instrument's scores.

This instrument was conceived and developed to be used with an optical reading tool. In this way we were also hoping to sensitize the teachers to the potential of this tool in readiness for its future availability in the DEM-PG.

Procedure

We applied the questionnaire during the second semester of the 2004–2005 school year, in the second-year semester discipline of Pathophysiology in the undergraduate medical course. This discipline consists of two lectures per week, each lasting one hour, on different weekdays. The lectures are available to all the students enrolled in the course, but attendance is not mandatory. Simultaneously, there is a small-group session component, of two hours' duration, where students have the opportunity to apply the contents presented in lectures. There are a total of 12 groups, each comprising about 20 students. In these lessons an active methodology is promoted, based on case studies. The main purpose for this approach is so that the students can apply the knowledge acquired in lectures to 'real' situations. In terms of Bloom's Taxonomy, the lectures are essentially focused on the basic domains of knowledge and understanding, while the small-group sessions are designed to promote the levels of application, analysis, synthesis and evaluation (Leinster 2005).The questionnaire was applied by the teacher at the end of each lesson. The answer was always given on a voluntary basis. However, none of the students refused to fill in the questionnaire.

Sixteen themes were taught in lectures and six in small-group sessions. There are themes that are taught in a single theoretical lesson (e.g. theme number 3), while others are taught in more than one lesson (e.g. theme 2 is taught in four lessons) and are, therefore, assessed more than once. There are teachers who only gave lectures (teachers' 1, 5, 6, 7, 8, 9). The others taught both types of lessons (teachers' 2, 3, 4). In the 2004–2005 school year 221

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008

students were enrolled in this discipline, 137 being female and the remaining 84 being male, with a mean age of 20 years.

Data analysis

Data analysis was conducted separately for lectures and small-group sessions. The data were analyzed using the Statistical Package for Social Sciences (SPSS for Windows Version 14.0) software.

We calculated the Kaiser–Meyer–Olkin (KMO) index and Bartlett's sphericity test to check whether our data qualified for factor analysis (see http://www.ncl.ac.uk/ucs/statistics, accessed 7 July 2006). Very small values of significance (below 0.05) for Bartlett's test indicate a high probability that there are significant correlations between variables, whereas higher values indicate the data are inappropriate for factor analysis. The KMO measure of sample adequacy provides an index (between 0 and 1) of the proportion of variance among the variables that might be common variance (i.e. that might be indicative of underlying or latent common factors). A KMO near 1.0 supports a factor analysis and anything less than 0.5 is not amenable to useful factor analysis. Kaiser (1974, cited in Dziuban and Shirkey, 1974) suggested that anything in the 0.90s was 'marvellous', in the 0.80s 'meritorious', in the 0.70s 'middling', in the 0.60s 'mediocre', in the 0.50s 'miserable' and below 0.50 unacceptable.

A principal component analysis (PCA) with varimax orthogonal rotation was performed for the 16-item teaching quality questionnaire. To identify the components we used the four criteria proposed by Konduri et al. (2006, selected by Konduri et al. based on Anastasi 1976, Hair et al. 1995, Zelnio and Simmons 1981 and Tabachnick and Fidell 2001):

(1) Kaiser's rule of an eigenvalue of at least 1.0 to retain the component.
(2) Items had to have loadings of 0.40 and above to be retained.
(3) Items with cross-loadings (loadings on more than one component) of 0.40 and above were deleted.
(4) Components with two items or less were deleted.

After the item deletion resulting from the application of the criteria mentioned above, PCA with varimax orthogonal rotation was again performed on the remaining set of items and criteria were again applied until all the criteria was respected.

The next step was to group the final set of items within their respective components and give names to the components based on their content. Internal consistency reliability estimates were calculated using Cronbach's coefficient alpha for the overall final questionnaire and for the individual identified components. According to Kline (1993), the Cronbach's coefficient alpha is the more important test for reliability assessment. We accepted alpha values higher than 0.70 (Hair et al. 1998; Rowland et al. 1991) and items with correlations (Pearson product moment correlation) with an overall scale score and the components higher than 0.20 (Streiner and Norman 1989).

Results

Factorial analysis—identification of questionnaire components

Lectures

We have used the questionnaires answered in the theme 3 lecture, since it was the first theme that was taught in a single lecture. The questionnaire was answered by 99

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008

students. Regarding the adequacy of our data for a factorial analysis, we have obtained a Kaiser–Meyer–Olkin index of 0.84 and a Bartlett's sphericity test with statistical significance ($p < 0.001$), which confirms the factoriality of the correlation matrix.

The initial rotated solution of three components explains a total of 59.15% of variance, with a contribution of 31.03% of component 1, 20.58% of component 2 and 7.54% of component 3. All three factors presented eigenvalues superior to 1 (4.97, 3.29 and 1.21, respectively). After applying the four criteria described to retained components explained in the data analysis section, 13 items remained. These 13 items were again subjected to a principal component analysis with varimax rotation and the four criteria applied again. This resulted in a two-component solution with the retention of 12 items. Analysis of the factorial saturations of items in each of the two components (cf. Table 1) shows us that items 1, 2, 5, 6, 7, 8, 15 and 16 saturate strongly (saturations superior to 0.66) in the same component (first). Items 10 to 13 have saturations superior to 0.70 in the second component. The two components accounted for 62.75% of the explained variance. The explained variance associated with each component was, respectively, 37.62% and 25.13%. The two components were named based on their content as follows:

(1) curriculum (eigenvalue = 4.51, 8 items);
(2) relationship 'teacher–students' (eigenvalue = 3.02, 4 items).

Small-group sessions

For the analysis of the psychometric properties of the questionnaire we used the small group-session corresponding to theme 1. The number of students who answered is much higher ($n = 188$) than those who answered in any of the lectures, since attendance at small-group sessions is mandatory, which is not the case with lectures. We obtained a Kaiser–Meyer–Olkin score of 0.90 and a Bartlett's sphericity test with statistical significance ($p < 0.001$), which confirms the factoriality of the correlation matrix.

Table 1. Lectures: item loadings for final varimax rotated two-component solution ($n = 99$).

	Component one	Component two
Component one: Curriculum		
Item 1. Relation between contents/subjects	0.709	0.230
Item 2. Contents/subjects interesting	0.680	0.280
Item 5. Understood the contents	0.755	0.109
Item 6. Teaching strategies adequate	0.788	0.142
Item 7. Audiovisual media contributed to the quality	0.774	–0.005
Item 8. Lesson was well planned	0.797	0.054
Item 15. Lesson quality	0.732	0.337
Item 16. Teacher quality	0.659	0.335
Component two: Relationship		
Item 10. Involvement of the students	0.014	0.837
Item 11. Critical reasoning of the students	0.155	0.880
Item 12. Self-directed learning of the students	0.241	0.796
Item 13. Clarifying students' doubts	0.269	0.717

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008

The initial rotated factorial solution explains 55.91% of variance, and, as in lectures, it presents three components that contribute, respectively, to explain 28.19%, 19.9% and 7.82% of variance with eigenvalues superior to one unit (4.51, 3.18 and 1.25, respectively).

After analysing saturations and applying the four criteria used to retained components, 14 items remained. These items were again subjected to a factorial analysis and the four criteria described applied again. This resulted in a two-component solution with the retention of 12 items, as for lectures. The scores, presented in Table 2, show us that the quality items (the last two items) had a cross-loading (loadings in both components) above 0.40. The fact that our purpose was to build an identical instrument for application in both lectures and small-group sessions, and that the loadings of the two last items were much higher (above 0.67) in the first component, prompted us to keep those two items in the final version of the scale.

The analysis of the factorial saturations of items in each of the two components (cf. Table 2) show us that items 1, 2, 5, 6, 7, 8, 15 and 16 saturate strongly (saturations superior to 0.52) in the same component (first). Items 10 to 13 have saturations superior to 0.62 in the second component. In this way we found the same structure for lectures and small-group sessions. The two components accounted for 59.50% of the explained variance. The explained variance associated with each component was, respectively, 34.14% and 25.36%, and the eigenvalues were, respectively, 4.51 and 3.02.

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008

Internal consistency

Lectures

The Cronbach coefficient alpha value for the total lectures ($n = 99$) questionnaire (12 items) was 0.89, indicating a high internal consistency. The internal consistency values for the two components are 0.89 (component 'curriculum') and 0.85 (component 'relationship'). The deletion of any item from total scale or component scales did not improve the values of internal consistency.

Table 2. Small-group sessions: item loadings for final varimax rotated two-component solution ($n = 188$).

	Component one	Component two
Component one: Curriculum		
Item 1. Relation between contents/subjects	0.674	0.170
Item 2. Contents/subjects interesting	0.590	0.353
Item 5. Understood the contents	0.534	0.162
Item 6. Teaching strategies adequate	0.796	0.280
Item 7. Audiovisual media contributed to the quality	0.731	0.039
Item 8. Lesson was well planned	0.784	0.174
Item 15. Lesson quality	0.742	0.415
Item 16. Teacher quality	0.671	0.472
Component two: Relationship		
Item 10. Involvement of the students	0.283	0.787
Item 11. Critical reasoning of the students	0.049	0.893
Item 12. Self-directed learning of the students	0.279	0.630
Item 13. Clarifying students' doubts	0.256	0.737

Table 3. Lectures: means, standard deviations and possible range for total questionnaire and its components (n = 99).

Variable	No. of items	Mean	SD	Possible range
Total score	12	47.45	5.81	12–60
Curriculum score (Component one)	8	34.07	4.09	8–40
Relationship score (Component two)	4	13.38	2.74	4–20

Corrected correlations values between each item and the total score range from 0.44 to 0.72 and from 0.64 to 0.72 for component one and from 0.61 to 0.77 as regards component two. The means, standard deviations and possible range for the total scale and for each component are presented in Table 3. Higher scores indicate higher quality.

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008

Small-group sessions

Every item obtains correlations with the scale's total score superior to 0.44 (range from 0.44 to 0.78). Cronbach's coefficient alpha value for the total small-group sessions (n = 188) questionnaire (12 items) was 0.89, indicating high internal consistency. The deletion of any item from the total scale did not improve the values of alpha.

As regards the two sub-scales, corresponding to the two components found, for the 'curriculum' sub-scale the Cronbach coefficient alpha value was 0.88, with item correlations with the sub-total score ranging from 0.44 to 0.78. The alpha values did not improve with the deletion of any item. For the 'relationship' sub-scale the alpha value was 0.79 improving to 0.82 with the deletion of item 12, 'The teacher promoted self-directed learning of the students outside the class'. We have opted to maintain the item in the questionnaire. This decision was due to the fact that the item presented a strong correlation (0.53) with the score of the total scale and with its own sub-total score sub-scale (0.52). Additionally our goal was to construct a questionnaire that could be used to assess lectures and small-group sessions' teaching quality in order to compare the results of each type of lesson.

The means, standard deviations and possible range for the total scale and for each component are presented in Table 4. Higher scores indicate higher quality.

Conclusions

The 1980s saw a growth in the promotion of assessment and quality of higher education. Van Vught (1995) defines the assessment of quality as the mechanisms and general processes that were, or still are, used with the intention of assuring that the products and processes of the institutions of higher education demonstrate 'quality', whether this quality is extrinsic (capacity to respond to the changes in the society where they belong), or intrinsic

Table 4. Small-group sessions: means, standard deviations and possible range for total questionnaire and its components (n = 188).

Variable	No. of Items	Mean	SD	Possible range
Total score	12	53.16	5.49	12–60
Curriculum score (Component one)	8	35.56	3.88	8–40
Relationship score (Component two)	4	17.60	2.24	4–20

(values and basic ideals that form the core of higher education: the search for truth and the production of knowledge). Throughout this article we have tried to provide a framework and justify the usefulness of an instrument for assessing teaching quality through listening to students' voices, within a broader process of quality enhancement in higher education that focuses on the pedagogical development of teachers.

The results obtained confirm the good global internal quality of the questionnaire and the components found correspond to nodal aspects of the quality of the teaching: curriculum and the relationship between teacher and students. Additionally it is a reliable instrument to use in both lectures and small-group sessions and may be useful in identifying areas of change that may be presented within disciplines and with teachers from different disciplines.

Using the data obtained from the questionnaire, the discipline's coordinator took three actions. First he withdrew the worst-ranked teacher from teaching theoretical classes. Then he organized a meeting with all the teachers to discuss the results and the strategies that should be followed to improve teaching. The best-ranked teacher was made responsible for coordinating those changes. Finally, the teachers with low rankings were advised to take a medical education course run by the DEM-PG.

The validation process of assessment instruments is a never-ending task. So, future research will need to be performed to establish the validity of the teaching quality assessment questionnaire we constructed and studied.

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008

Notes on contributors

Maria Filomena Gaspar is Auxiliary Professor at the Faculty of Psychology and Educational Sciences, University of Coimbra, Portugal, a member of the Center of Psicopedagogia (FEDER/POCI2010-SFA-160-490, University of Coimbra), and guest member of the Undergraduate Section of the Department of Medical Education (Faculty of Medicine) in the school year 2005–2006.

Anabela Mota Pinto is Associate Professor of the Faculty of Medicine, University of Coimbra, Portugal, and a member of the Undergraduate Section of the Department of Medical Education (Faculty of Medicine).

Hugo Camilo F. da Conceição is an Educationist in the Undergraduate Section of the Department of Medical Education of the Faculty of Medicine, University of Coimbra, Portugal.

José António Pereira da Silva is Associate Professor with Aggregation of the Faculty of Medicine of the University of Coimbra, Portugal, and Coordinator of the Undergraduate Section of the Department of Medical Education (Faculty of Medicine).

References

Crespo, V. 1993. *Uma universidade para os anos 2000: o ensino superior numa perspectiva de futuro.* Mem Martins: Inquérito.

Dent, J., and R. Harden. 2005. *A practical guide for medical teachers.* 2nd edn. Edinburgh: Churchill Livingstone/Elsevier.

Dziuban, C., and E. Shirkey. 1974. When is a correlation matrix appropriate for factor analysis? *Psychological Bulletin* 81, no. 6: 358–361.

European University Association. 2002. *Evaluation of the 5 established faculties of medicine in Portugal.* Draft final report.

Fincher, R.-M., D. Simpson, S. Mennin, G. Rosenfeld, A. Rothman, M. McGrew, P. Hansen, P. Mazmanian, and J. Turnbull. 2000. Scholarship in teaching: an imperative for the 21st century. *Academic Medicine* 75, no. 9: 887–894.

Hair, F., R. Anderson, R. Tatham, and W. Black. 1998. *Multivariate data analysis with readings.* 5th edn. Englewood Cliffs, NJ: Prentice Hall.

Kline, P. 1993. *An easy guide to factor analysis.* London: Routledge.
Konduri, N., G. Gupchup, M. Borrego, and Worley-Louis, M. 2006. Assessment of the reliability and validity of a stress questionnaire. *College Student Journal* 40, no. 1: 78–90.
Leinster, S. 2005. The undergraduate curriculum. In *A practical guide for medical teachers,* eds. J. Dent and R. Harden, 2nd edn. Edinburgh: Churchill Livingstone/Elsevier.
Palomba, C., and T. Banta. 1999. *Assessment essentials: planning, implementing and improving assessment in higher education.* San Francisco, CA: Jossey-Bass.
Rowland, D., D. Arkkelin, and L. Crisler. 1991. *Computer-based data analysis: using SPSS in the social and behavioral sciences.* Chicago: Nelson-Hall.
Streiner, D., and G. Norman. 1989. *Health measurement scales.* Oxford, UK, Oxford University Press.
Van Vught, F. 1995. The new context for academic quality. In *Emerging patterns of social demand and university reform: through a glass darkly,* eds. D. Dill, and B. Sporn. Oxford: IAU Press/ Pergamon.

Downloaded By: [B-on Consortium - 2007] At: 11:34 4 September 2008