# How Can We Individualise Faculty Members' Teaching Evaluations?
# Proposing an Evaluation Model

Nikoo Yamani, Alireza Yousefy and Tahereh Changiz
Department of Medical Education, Medical Education Research Center,
Isfahan University of Medical Sciences, Isfahan, Iran

**Abstract:** The characteristics of teacher evaluations vary between educational settings and the variety of learners. This study attempts to develop a teacher evaluation model that provides an opportunity for every faculty member to design his own evaluation form based on the educational setting and types of learners. Using a comprehensive evaluation item bank, with 271 items designed in a previous study, the items were distributed in various educational settings and to different types of students. The steps through, which each faculty member could design his own questionnaire were then determined and a software was designed for this purpose. In the pilot study, 55 faculty members form the School of Medicine were randomly selected to participate and their views on this model were solicited through the use of a questionnaire. The first stage produced was 14 smaller databases, each including a number of items appropriate for a specific educational setting. A software for creating evaluation forms was then designed. In the third stage, 46 faculty members participated in the study creating their own forms using the software. The mean of their responses was $4.02\pm0.5$ out of 5, the highest scores on feeling of participation ($4.38\pm0.67$) and satisfaction from the presented model ($4.25\pm0.71$). The individualisation model of evaluation proved to be successful by providing faculty members the possibility of designing an evaluation form based on the educational setting and the types of learners. More studies are recommended to investigate the application of this model in other disciplines.

**Key words:** Evaluation, item bank, teacher evaluation, teaching effectiveness

## INTRODUCTION

The improvement of the quality of learning and teaching is a primary objective in every university and is a precondition for achieving or maintaining excellence in teaching. Evaluation as a critical element in professional practice plays an important role in this regard. It serves unique professional development purposes as well as broader purposes involving comparison. Both purposes are legitimate and both kinds of evaluation are to be encouraged by the university.

Medical schools require evaluation as part of their quality assurance procedures. It provides evidence of how well students' learning objectives are being met and whether teaching standards are being maintained (Morrison, 2003). Teacher evaluation must be concerned with one primary purpose, which is to facilitate better student learning through the improvement of teaching. In order to evaluate faculty members' teaching, teacher

evaluation tools have been used widely in most higher education institutions. Among such tools, student evaluation forms have been used and also criticised the most.

Although, students are not in a position to provide useful observations about all aspects of teaching, there are, however, a range of very important matters to do with teaching that students are in an excellent position to observe and judge reliably (Boyle *et al.*, 1997; Green *et al.*, 1998). The literature shows many criticisms about student evaluations of teachers. The most common criticism is that teacher evaluation ratings are influenced by students' course grades (Colaizzie, 1995; Huemer, 2007). Therefore, the faculty members would make their courses easier to receive higher ratings. Other criticisms indicate that some variables such as class size, superficial stylistic matters, teachers' personality, gender, age and even accent influence students' ratings (Haskell, 1997; Huemer, 2007). Moreover, student evaluations are considered as a threat

**Corresponding Author:** Nikoo Yamani, Department of Medical Edcation, Medical Education Development Center,
Medical Education Research Center, Isfahan University of Medical Sciences, Isfahan,
P.O. Box: 81745/353, Iran

to academic freedom by providing a control mechanism over curricular, course content, grading and teaching methodology (Haskell, 1997; Huemer, 2007). Despite these criticisms, there are some studies, which have shown that students' evaluations are relatively unaffected by a variety of variables mentioned in some other studies (Marsh, 1987; Seldin, 1993).

There is one thing clear that student ratings have become the most widely used and sometimes the only source of information for teaching evaluation (Copeland and Hewson, 2000; Kreiter and Lakshman, 2005; Pelsang and Smith, 2000; Seldin, 1993). Because they are easy and inexpensive to administer, give an impression of objectivity and there are few alternative to students' evaluations (Huemer, 2007). There is also, wide support in the literature for the validity of properly derived and used student ratings of teacher effectiveness, for both formative and summative purposes (Colaizzie, 1995; Copeland and Hewson, 2000; Green *et al.*, 1998). Many studies have determined that student ratings generally are both reliable and valid (Bardes and Hayes, 1995; Boyle *et al.*, 1997; Copeland and Hewson, 2000; Irby *et al.*, 1987; Seldin, 1993; Shores *et al.*, 2000).

However, teaching involves a complex and diverse set of processes. Therefore, any evaluation of teaching requires an approach based on a range of sources and types of information. This is particularly important for summative purposes. Beckman *et al.* (2004) believes that characteristics of teacher evaluation vary between educational settings and between different learner levels, indicating that we should consider different teaching contexts in designing a teaching evaluation system. Moreover, some research suggests that some (usually small) bias effects may be present in student ratings due to the effects of different teaching contexts (e.g., small group teaching compared to large group teaching). Systematic difference in student ratings of teaching have also been known to occur across disciplines (Boyle *et al.*, 1997).

A few suggestions have been made in order to manage these concerns. Firstly, different forms can be used for teaching contexts that are markedly different. Secondly, patterns of data from different contexts (including discipline areas) and on different variables can be monitored over time in order to detect potential sources of bias. Finally, teaching staff should gather student evaluation data from a variety of learners over time (e.g. clerkship, internship, residency,... ).

Faculty resistance is one of the main problems faced by evaluation programs. Some faculty members perceive summative evaluation as a threat to academic freedom (Green *et al.*, 1998). The literature shows that teachers are willing to accept the principles of evaluation but usually do not agree with the methods of evaluation. One of the problems is based on the belief that while teacher evaluation should only serve the purpose of improving teaching in most places it is used as a tool for decision-making as well (Shinkfield and Stufflebeam, 1995). By involving faculty members in the evaluation process, it may be possible to change their attitude towards evaluation and encourage them to pay more attention to evaluation feedback. As Chambers (Chambers *et al.*, 2003) believes that faculty members need personal involvement in the activities that are critical to their professional growth. It has been reported in the literature that faculty members' involvement in the design and construction of the evaluation system is one of the key elements of designing a faculty evaluation system (Bland *et al.*, 2002).

Considering all the above-mentioned issues, this study was designed to overcome some of the concerns of teaching evaluation by students. The purpose of the study was first, to provide the possibility for every faculty member to design different evaluation forms based on different teaching contexts (classroom, clinic or hospital, laboratory or skill laboratory, workshop) and types of learners (student, intern, resident, MS of PhD student). The study also aimed to involve faculty members in designing their own evaluation forms, which would be part of their own evaluation process.

## MATERIALS AND METHODS

In order to proceed with designing the present study, we needed a comprehensive teacher evaluation item bank that included a variety of items that covered all domains of teaching activities. This item bank had been designed in a previous study by the researcher and her colleagues. Using this comprehensive item bank, the present study was designed in 3 main stages as follows:

**Stage 1: Preparing the required item bank:** The above mentioned comprehensive item bank included 271 items, which had been divided into 13 main sub-domains. The domains and sub-domains are shown in Table 1.

During this stage, the 271 items were distributed according to the educational settings and types of students. The educational settings in this study included: Classroom (theoretical), hospital or clinic (clinical), laboratory or skill laboratory (practical) and workshop.

Table 1: The domains and sub-domains of the teacher evaluation item bank

| Domains | Sub-domains |
| --- | --- |
| Preparation for teaching | Selecting and organising the objectives |
|  | Selecting and organising the content |
|  | Learning resources |
| Organising the learning environment | Learning environment |
| Management of the teaching-learning process | Teaching methods |
|  | Educational media |
|  | Learning assignments |
|  | Student assessment methods |
| Professional characteristics of the teacher | Knowledge and skills |
|  | Professional ethics |
|  | Teacher's motivation for teaching (enthusiasm for teaching,efforts for scientific/professional promotion) |
| Learning outcomes | Reinforcement of student's motivation and attitude |
|  | Promotion of student's knowledge and skills |

The types of students included: student, intern, resident and postgraduate student (MS or PhD). This distribution took place in a group, which consisted of 7 experts in medical education who were all faculty members of the university. The Microsoft Excel software was used to facilitate the distribution of items. The expert group had to agree, on which one or more of the students' type and educational setting each item belonged to. Each item could be distributed to >1 educational setting or type of students.

After all items were distributed, the same experts regrouped to determine the number of items to be placed in each evaluation form and the number of items from each sub-domain to be placed in the evaluation form. In order to do this, each sub-domain was scored (out of 100) based on its importance in the effectiveness of teaching.

**Stage 2: Proposing the individualisation model:** In this stage, the steps through, which each faculty member could design his own evaluation form based on the educational setting and student types were determined. The researcher proposed these steps and they were revised and finalized during an expert group meeting. After making sure each aspect of individualisation had been considered, a software was designed to facilitate the creation evaluation forms through a user-friendly program. This software could minimise the amount of time required to create the evaluation forms as well as make the process less complicated.

**Stage 3: The pilot study:** In order to investigate the applicability of the proposed model and software, 55 faculty members were selected to create their own evaluation forms through, the software. Since, the aim of the individualisation model is to design the evaluation

forms according to different educational settings and types of students, faculty members of the School of Medicine were selected for the pilot study. For a medical university in Iran, the School of Medicine is the biggest school, which includes a variety of educational settings and types of students. In selecting the sample, we attempted to choose a range of faculty members involved in all educational settings and teaching different types of students. Fifty-five faculty members were selected and invited to take part in workshops with 10-12 faculty members each. At the beginning of each workshop, the researcher explained the aims of individualisation in evaluation and provided the necessary instructions for the faculty members to create their own evaluation forms using software provided. The steps of the proposed model will be explained in the results section. Faculty members' views towards this model of evaluation were solicited through a 21-item questionnaire using the Likert scale ranging from completely agree to completely disagree. For the analysis of data, a score was specified to each scale of the questionnaire. So that, completely agree, received a score equal to 5, completely disagree, a score equal to 1 and other scales received scores in between.

The main domains of the questionnaire included participation in evaluation, satisfaction from evaluation, the item bank as being easy to use, comprehensiveness of item bank, the role of feedback in this model of evaluation and the application of evaluation results. Face and content validity of the questionnaire was confirmed by faculty members and educational experts. Its reliability was .91 by Cronbach Alpha.

**RESULTS**

The result of the first stage of distributing the items based on educational setting and type of student was 14 smaller databases. These databases were named according to the educational settings and types of students they were designed for and included: Theory-Student, Theory-Intern, Theory-Resident, Theory-Postgraduate, Practical-Student, Practical-Postgraduate, Practical-resident Clinical-student, Clinical-intern, Clinical-resident, Workshop-student, Workshop-intern, Workshop-resident and Workshop-postgraduate.

It should be mentioned that theoretical refers to teaching in a classroom, practical refers to teaching in a laboratory or skill laboratory and clinical refers to teaching in a hospital or clinic.

In the first stage, the number of items for the evaluation form and the number of items from each sub-domain in the evaluation form were also determined. According to the expert group, the number of items

should be enough to cover all teaching domains adequately but not too many to be boring for the students. If too many items were placed in the evaluation form, the students would not pay enough attention to the items and as a result the reliability of the form would be lower than expected. It was, therefore, agreed to place 40 items in each evaluation form. By scoring each sub-domain according to its importance in teacher evaluation, the number of items from each sub-domain was determined for each educational setting as well.

The result of the second stage was suggesting the individualisation model and its steps, which led to designing a user friendly software for designing evaluation forms.

The individualisation model was suggested based on the following principles:

- Using students' points of view as they were accessible, reliable, inexpensive, practical and valid.
- Considering the characteristics of effective teaching for designing the item bank and developing teaching domains and sub-domains.
- Having faculty members' participation in different aspects of evaluation such as:

  - Determining the domains of effective teaching.
  - Designing the evaluation items.
  - Distributing the items based on educational settings and types of students.
  - Editing the evaluation forms (those designed through the individualisation software).

- Making the evaluation forms specific to different factors, such as the educational setting and type of student.
- Recognising diversity in different educational settings such as clinical, practical, theoretical and workshop.
- Recognising diversity in the educational needs of different types of students such as intern, resident, MS or PhD students.
- Providing some strategies for achieving the goals of summative evaluation as well as faculty members' formative evaluation such as specifying global items in each evaluation form.

Keeping the above mentioned principles in mind, the following steps were taken through, which each faculty member could design his own evaluation form:

Choosing the educational setting (clinic, theory, practice, workshop) and the type of students who would be evaluating him in that educational setting, Selecting

Table 2: The mean and standard deviation of faculty members' viewpoints towards individualization model based on the domains of the questionnaire

| Questionnaire domains | Mean | S.D. |
|---|---|---|
| Participation in evaluation process | 4.38 | 0.67 |
| Satisfaction of evaluation model | 4.25 | 0.71 |
| Easy to use item bank | 4.05 | 0.55 |
| having a comprehensive item bank | 3.83 | 0.69 |
| The effective role of feed back in this model | 4.11 | 0.67 |
| Application of evaluation results | 3.73 | 0.69 |
| Total mean | 4.02 | 0.50 |

the evaluation items under each domain in the item bank provided, Selecting an open-ended question, which has been provided through the software, Typing an open-ended question if desired and if faculty member needs feedback on specific aspect, Providing suggestions and comments by typing in the provided box, Editing the previously created forms and Receiving feedback and comments based on the forms filled by the students.

Based on the above-mentioned steps an evaluation software was designed through, which faculty members could easily take those steps and make their own forms.

In the third stage, from 55 faculty members who were invited to participate in the pilot study, 46 attended the workshop. The result was 46 different evaluation forms for a specific educational setting. The pilot study proved that the individualisation model was practical, successful and easy to apply. Some faculty members made about 3 different forms during the pilot study, which shows that faculty members enjoyed participating in creating their own evaluation forms.

Based on the results of the questionnaire, the mean and standard deviation of faculty members' views regarding the individualisation model of evaluation were 4.02±.5 out of the total score of 5. The mean and standard deviation of faculty members' views of the different domains of the questionnaire are presented in Table 2. This table shows that among the 6 domains of the questionnaire, feeling of participation and feeling of satisfaction had the highest scores. The lowest score belonged to application of evaluation results.

**DISCUSSION**

This study tried to propose an evaluation model, which not only considers different educational settings and types of students in designing evaluation forms, but which also increases faculty members' participation in the evaluation process. The latter, in turn could increase faculty members' satisfaction with evaluation, which could lead them to paying more attention to evaluation feedback and moving towards improvement.

In an evaluation system, the evaluation tool is of much importance because improving faculty members' performance as well as faculty development programs is not possible without a fair evaluation system using defined processes and appropriate tools (Colaizzie, 1995). Other studies have also emphasised the importance of the evaluation tool such as Kreiter and Lakshman (2005), who believes that student evaluation of teaching is one of the most accessible sources of information and plays an important role in managerial decision-making. Therefore, using a valid and reliable evaluation tool is very important (Kreiter and Lakshman, 2005). The main focus of the present study was on the evaluation tool, to make it more specific and more applicable. In developing the individualisation model by considering the evaluation tool as the central point for quality improvement, 3 main principles were considered, including considering the characteristics of effective teaching, faculty members' participation in their own evaluation and using students' viewpoints as a reliable source for teacher evaluation. The focus of this model was to provide information for different evaluation purposes. In other words, by considering different educational settings and types of students, we tried to propose a model, which not only can achieve the formative purpose of evaluation by providing feedback for faculty improvement, but that could also help managerial decision-making as summative evaluation. As Green *et al.* (1998) mentions in his study, formative evaluation is a desirable method for assessing faculty members and encouraging their improvement. However, he believes performing summative evaluation for decision-making is unavoidable and some aspects of formative evaluation have to be added to the process of summative evaluation (Green *et al.*, 1998).

The individualisation model requires many steps to be taken in the creation of an individualised evaluation form. In order to make this process easier and more applicable, the researcher had access to experts who designed a software, which not only makes the process easier and simpler but also requires less time for designing the forms. Moreover, the software includes a variety of features such as the possibility of editing the item bank and evaluation forms as well as receiving feedback from faculty members who use this system. The literature has also shown that there have been moves towards computerising the evaluation system used in many universities and putting the system online, which can increase participation in evaluation up to 81-92% (Afonso *et al.*, 2005).

According to Bland *et al.* (2002), experience, faculty members' support is one of the key elements of faculty evaluation system design. In our pilot study, almost 83% of faculty members participated, showing their desire to participate in their own evaluation. This in turn could increase their motivation to receive feedback and act on it with regard to improvement. Research has shown faculty members tend to resist evaluation and suggests that by having faculty members' participation in different aspect of evaluation, it may be possible to remove this resistance (Green *et al.*, 1998; Neal, 1988) and provide ground for paying more attention to the evaluation results. During the pilot study, many participants showed an interest in studying different evaluation items and admitted that they can obtain some information regarding the teaching-learning process by reading the items. Thus, the individualisation model can increase faculty members' knowledge regarding effective teaching, by having them read different evaluation items.

Faculty members' views towards the individualisation model showed that feeling of participation in evaluation process had the highest score. As other studies have shown, faculty members' attitude towards evaluation affects their use of the evaluation results. This means that those with a positive attitude take more advantage of the evaluation results for the improvement of their performance, compared to those who lack such an attitude.[17] Therefore, the individualisation model is able to increase the effectiveness of teacher evaluation by improving the attitude of faculty members towards evaluation. Moreover, other studies have shown that participation of faculty members in the evaluation process reduces their resistance against evaluation (Green *et al.*, 1998).

Some limitations of the present study include the fact that there are many different evaluation forms, making the evaluation process and the analysis of evaluation forms more complicated. In addition, some faculty members resist using the system and designing their evaluation forms and it would be difficult to compare faculty members based on their evaluation results. However, compared to the advantages of this system, it seems that such limitations can be resolved. In order to overcome the inability to compare faculty members, several strategies such as using global items, using fixed items in the questionnaire and weighing evaluation items can help. The literature also shows that using a global item concerning course effectiveness is an appropriate criterion (Stringer and Irwing, 1998).

## CONCLUSION

While teaching-learning contexts vary in many respects, it is difficult to imagine contexts where systematic and properly collected information from

students would not improve the quality of teaching and learning. In general, using different forms for student evaluations of teaching provides important information, which can assist with the enhancement and profiling of teaching quality at various levels.

As with all processes regarding the improvement of teaching and learning, ongoing critical review and adjustment of the process is necessary. The teacher evaluation office should monitor these processes over time with a view to continuous improvement.

For ensuring the usefulness of this model in the improvement of teaching, first the clarity of feedback to faculty members, second, effective follow-up change in teaching and third availability and use of an expert consultant in the area of teaching and learning are recommended. However, it should be mentioned that when working towards the development of best practices there are many factors to consider and balance out.

## REFERENCES

Afonso, N.M., L.J. Cardozo, O.A. Mascarenhas, A.N. Aranha and C. Shah, 2005. Are anonymous evaluations a better assessment of faculty teaching performance? A comparative analysis of open and anonymous evaluation processes. Fam. Med., 37: 43-47. PMID: 15619155.

Bardes, C.L. and J.G. Hayes, 1995. Are the teachers teaching? Measuring the educational activities of clinical faculty. Acad. Med., 70: 111-114. PMID: 7865035.

Beckman, T.J., A.K. Ghosh, D.A. Cook, P.J. Erwin and J.N. Mandrekar, 2004. How reliable are assessments of clinical teaching? A review of the published instruments. J. Gen. Int. Med., 19: 971-977. PMID: 15333063. DOI: 10.1111/j.1525-1497.2004.400 66.x. http://www3.interscience.wiley.com/cgi-bin/full text/118783928/HTMLSTART.

Bland, C.J., L. Wersal, W. VanLoy and W. Jacott, 2002. Evaluating faculty performance: A systematically designed and assessed approach. Acad. Med., 77: 15-30. PMID: 11788318. http://ovidsp.tx.ovid. com/spb/ovidweb.cgi?&S=MLAIFPNNHHDDJEOP NCHLMDGJHHJBAA00&Link+Set=S.sh.46%7c1% 7csl_10.

Boyle, P., M. Pettigrove and C. Atkinson, 1997. Australian National University Student Evaluation of Teaching (ANUSET): A guide for academic staff. Australian National University. http://www.anu.edu. au/cedam/publications/setdocs/anusetguide1997.p dfss.

Chambers, D.W., H. Boyarsky, B. Peltier and F. Fendler, 2003. Development of a mission-focused faculty evaluation system. J. Dent. Educ., 67: 10-22. PMID: 12540101.

Colaizzie, J.L., 1995. Evaluation of faculty members. Am. J. Pharm. Educ., 59: 75-84. http://www.ajpe.org/legacy/ ajpe_metaview.asp?ID=1251.

Copeland, H.L. and M.G. Hewson, 2000. Developing and testing an instrument to measure the effectiveness of clinical teaching in an academic medical center. Acad. Med., 75: 161-166. PMID: 10693849.

Green, M.E., C.L. Ellis, P. Fremont and H. Batty, 1998. Faculty evaluation in Departments of Family Medicine: Do our universities measure up? Med. Edu., 32: 597-606. PMID: 10211249. DOI: 10.1046/j. 1365-2923.1998.00242.x.

Haskell, R.E., 1997. Academic freedom, tenure and student evaluation of faculty: Galloping polls in the 21st century. Education Policy Anal. Arch., pp: 5. http://epaa.asu.edu/epaa/v5n6.html.

Huemer, M., 2007. Student Evaluation: A critical review. http://home.sprynet.com/~owl1/ sef.htm.

Irby, D.M., G.M. Gillmore and P.G. Ramsey, 1987. Factors affecting ratings of clinical teachers by medical students and residents. J. Med. Edu., 62: 1-7. PMID: 3795239.

Kreiter, C.D. and V. Lakshman, 2005. Investigating the use of sampling for maximising the efficiency of student-generated faculty teaching evaluations. Med. Edu., 39: 171-175. PubMed: 15679684. DOI: 10.1111/j.1365-2929.2004.02066.x. http://www3.interscience.wiley.com/journal/ 118738018/abstract.

Marsh, H.W., 1987. Student evaluations of university teaching: research findings, methodological issues and direction for future research. Int. J. Edu. Res., 11: 253-388. http://eric.ed.gov/ERICWebPortal/ custom/portlets/recordDetails/detailmini.jsp?_nfpb= true&_&ERICExtSearch_SearchValue_0=ED338629 &ERICExtSearch_SearchType_0=no&accno=ED33 8629.

Morrison, J., 2003. ABC of learning and teaching in medicine: Evaluation. Br. Med. J., 326: 385-387. PMID: 12586676. http://www.bmj.com/cgi/content/ short/326/7385/385.

Neal, J.E., 1988. Faculty Evaluation: Its purposes and effectiveness, ERIC Clearing house of higher education. ED308800.

Pelsang, R.E. and W.L. Smith, 2000. Comparison of anonymous student ballots with student debriefing for faculty evaluations. Med. Edu., 34: 465-467. PMID: 10792688. DOI: 10.1046/j.1365-2923.2000. 00565.x http://www3.interscience.wiley.com/journal/ 119010627/abstract.

Seldin, P., 1993. The use and abuse of student ratings of professors. Chron High Edu., 39: 40. http://trc.uc davis.edu/TRC/ta/TAdevel/seldin.pdf.

Shinkfield, A.J. and Stufflebeam, 1995. Teacher evaluation: Guide to effective practice. 1st Edn. (Boston, Kluwer Academic Publisher), pp: 9-24. ISBN: 0-7923-9674-X.

Shores, J.H., M. Clearfield and J. Alexander, 2000. An index of students' satisfaction with instruction. Acad. Med., 75: 6-8. PMID: 11031190.

Stringer, M. and P. Irwing, 1998. Students' evaluation of teaching effectiveness: A structural modeling approach. Br. J. Edu. Psychol., 68: 409-426. http://cat. inist.fr/?aModele=afficheN&cpsidt=2376091.