
A Proposed System for the Evaluation of The Teaching Performance of Lecturers by Students

See Ching Mey
School Of Educational Studies
Universiti Sains Malaysia

Kertas ini mencadangkan satu sistem penilaian prestasi pengajaran pensyarah-pensyarah oleh pelajar-pelajar. Satu soalselidik yang mengandungi 7 kategori dan 38 item telah di sediakan. Kategori-kategorinya adalah seperti berikut: (1) Organisasi, struktur atau kejelasan (2) Interaksi atau perhubungan pensyarah-pelajar (3) Kemahiran mengajar, berkomunikasi atau kebolehan bersyarah (4) Isi kandungan kursus dan kesukaran kursus (5) Tugas dan penilaian (6) Kesan terhadap pelajar (7) Penilaian secara menyeluruh.

Satu kajian perintis dijalankan di Maktab Perguruan Sultan Idris, Tanjung Malim, Perak (sekarang di kenali sebagai Institut Pendidikan Sultan Idris) untuk menyemak kesahan dan keutuhan soalselidik, mengkaji tatacara menjalankan soalselidik, menganalisa dan melaporkan maklumbalas pelajar-pelajar dan kegunaan maklumbalas itu.

Selepas itu, penyelidik akan mencadangkan satu sistem penilaian prestasi pengajaran pensyarah oleh pelajar.

Introduction

In general, the two approaches or paradigms used in the evaluation criteria for teaching is process or product. The product -criterion means that the assessment is based on the outcome s of instruction, in particular student achievement. This emphasis on the performance or achievement assessment is related to behaviourst based research among American educators. It will be sufficient here to quote:

In spite of recent enthusiasm about evaluating teachers through student achievement, there is little evidence to suggest that performance measures are really useful in doing so. Comparison of achievement measures is practical only in multiple-section courses, and it is hard to see how performance measures can provide a common yardstick for ranking faculty members in different fields. A further problem is that students in different sections of a multiple-section course apparently learn fairly similar amounts. Perhaps the most impressive thing about studies relating class achievement to class ratings of instructors is the inconsistency of results". (Kulik & McKenchie, 1975:235)

If the assessment of teaching based on student achievement or performance sounds like "payment by results" and if the above criticism are valid, it becomes necessary to turn to the process-criterion paradigm as a means of evaluation.

The process-criterion paradigm focuses on various aspects of teacher and student behaviour which are judged to be worthwhile in their own right or to be in line with student achievement. Almost invariably, the evaluation by process-criterion depends on the method of either systematic observations and/or ratings of observed or inferred behaviour.

This rating is observed or inferred behaviour of instructors can be done by administrators, peers, professional evaluators, parents, public, alumni, students and / or oneself. In this study, the student ratings in the evaluation of classroom teaching is used.

Definition, Usefulness and Limitations of Student Ratings

Crittenden and Nor (1973) defined the student ratings as a case of "person perception" where a student perceives and forms evaluative impressions of the instructors. The overall evaluation of the instructors is a combination of evaluation of individual aspects of teaching behaviour weighted by the students' estimation of the relative importance of these aspects to good teaching.

Two viewpoints surfaced in the discussion on the use of student ratings. Critics have pointed out the inadequacy and biasness of the student as a judge of teaching ability, emphasizing his inexperience and lack of maturity. Proponents, on the other hand, have emphasized that effective learning results from the student-teacher interaction and that, however biased the ratings or evaluations may be, they are valuable as a source of information concerning student reactions to the behaviour or performance of instructors.

Evidence has accumulated concerning students' ability to rate instructors, and more and more institutions have turned to student ratings for one purpose or another. Brickman (1966) reported that there is an increased trend among college students to evaluate their professors, and that students now insist on official recognition of their judgements. He argued that although students in graduate and professional schools still generally lack the attribute of mature scholarship which enables them to appraise their instructors accurately, the fact remains, however, that they do evaluate their instructors officially or unofficially.

Centra (1977a, 1977b) and Moomaw (1977) stated that evaluative information about instructional effectiveness from students is most widely used. Aleamoni (1981) in his article titled "Student Ratings of Instruction" agreed with this claim:

Student ratings are increasingly being used by faculty, students and administrators for formative and summative decisions about instructional effectiveness. In fact, student ratings tend to be the only tangible source of instructional evaluation information in the majority of colleges and universities, both here (United States) and abroad. (Aleamoni, 1981: 110).

In Malaysia, the significance of student ratings in higher institutions of learning has been recognized. Nayan Arifin (1980), the deputy Vice-Chancellor of the Agricultural University of Malaysia said:

Lecturers in higher educational institutions tend to overlook modern technological developments in teaching technique. They tend to stick too much to the traditional "talk and chalk" method. Therefore, it is pertinent for institutions of higher education to embark on a programme of student evaluation of teachers. (New Straits Times, Oct.7, 1980:7).

Attempts by a number of higher institutions to carry out this programme informally have begun on a modest scale but its success and continuity is left to be seen.

The usefulness of student ratings is documented by Kent (1966) who concluded that neither blind acceptance nor mulish rejection of student rating of teaching effectiveness is an appropriate response at the present time. His most important argument is that student rating is useful as it offers a hopeful possibility for solving some of the concerns connected with the evaluation of teacher effectiveness and ways of improvement. Although Langen (1966) admitted the shortcomings of student ratings, he also indicated that the system provides a variety of data indicative of teaching quality and ways of improvement. Centra (1973) also suggested that student feedback did effect some changes in instruction over time, as instructors who had received feedback twice during the previous semester did receive better ratings than those who had received feedback once or not at all.

Marsh, Fleiner and Thomas (1975) in their findings indicated that student ratings are valid measures of instructional quality. It also provides useful feedback to improve instruction. There is little or no controversy about the use of information from student ratings to provide feedback to the instructors to improve teaching. However, the use of information for administrative purposes or as a reward system where it becomes a determinant of salary, promotion and tenure decisions is a matter of heated debate.

In the issue of the validity and reliability of student rating, Costin and his associates (1972) indicated that student ratings can provide reliable and valid information on the quality of courses and instruction. Scott (1975) also stressed that student ratings constitute one of the most credible indicators of professional performance. Furthermore, based on the findings of the meta-analysis carried out by Cohen (1981), one can safely say that student ratings of instruction are a valid index of instructional effectiveness.

At this point, there is a need to mention the possible sources of bias in student ratings. They are:

1. **Instructor Characteristics :** Coats and Swierenga (1972) stated that students do not respond directly to specific questions regarding teacher effectiveness. Rather, a kind of "halo effect" based on teacher charisma or popularity determines to a large extent how students react to questions about their teacher. The study of Naftulin, Ware and Donnelly (1973) on the Fox paradigm has received considerable attention. Their premise is that the enthusiasm or expressiveness of an instructor can significantly influence ratings regardless of the amount of substantive content presented. Although the methodology used in their study was weak (Kaplan, 1974), further studies confirmed and expanded upon the influence of instructor expressiveness on student ratings of instruction.

2. **Size of the Class:** The instructor teaching in smaller classes tends to receive higher ratings (Kulik and McKeachie, 1975; Feldman, 1978; McKeachie, 1979; Marsh, 1984). Centra (1978) reported that instructor teaching in classes of size fifteen and less are more effective in producing student learning and are also rated higher by students. More specifically, Frey (1978) found that the number of students in a class does not affect ratings on the skill items, but does influence ratings on the rapport items.
3. **Purpose of Ratings:** Students tend to give higher ratings if the purpose is for summative rather than formative evaluation (Feldman, 1979; Marsh, 1984).
4. **Non-anonymous Ratings:** If the students are requested to write their name on the rating, the ratings tend to be higher (Feldman, 1979; Marsh, 1984).
5. **Presence of Instructor:** If the instructor is present while students complete the rating form, the ratings tend to be higher (Feldman, 1979; Marsh, 1984).
6. **Expected Grades:** There tends to be a positive, but low correlation between student ratings and their expected grades (Howard and Maxwell, 1982; Marsh, 1984).

The Present System of Evaluation Lecturers in Malaysia

The present practice of evaluating lecturers in all government and semi-government institutions of learning follows the performance evaluation system that is in practice in the Malaysian Education Service and in the Malaysian Civil Service. This present system is an extension of the practice that developed in the Malaysian Civil Service. Thus, the same evaluation forms and similar regulations cover the practice of evaluation in the government and semi-government institutions of learning in the country. Furthermore, the basis of the system lies in the directions stipulated by the Government "General Orders", specifically the Service Directive, 1974 (Arahan Perkhidmatan, 1974) issued by the Public Service Department.

The functions of the system are:

1. to assess an individual's suitability for promotion;
2. to assess an individual's suitability to perform new tasks, or a more responsible task;
3. to identify an individual's strengths and weaknesses in the performance of his job, and also to identify his interests and suitability for a particular type of job;
4. to identify the training and professional development needs of an individual; and
5. as a useful record of an individual's past and present progress in his job. (Jabatan Perkhidmatan Awam, 1974:5).

the job, and his interest in his job. In addition, the evaluator is in no position to evaluate the lecturer in this area as he is not one of the students who sits in the classroom daily and who is at the receiving end of the teaching-learning process. Lastly, the system practiced is "confidential" in that the lecturer does not know the ratings he has received and thus remains ignorant of his strengths and weaknesses. More often than not there is no counseling interview or discussion carried out by the evaluators to ensure that the lecturer overcomes his weaknesses for future improvement.

A Proposed System for the Evaluation of the Teaching Performance of Lecturers by Students.

In view of the limitations of the present system, the researcher has proposed and pilot-tested a system for the evaluation of the teaching performance of lecturers by students.

Students are an indispensable source of information about the achievement of important educational goals, the different aspects of rapport between the lecturers and themselves, and the instructional effectiveness of the lecturers. Student ratings provide a means of communication between students and lecturers. Such communication may lead to the kind of involvement by both parties in the teaching-learning process that raises the whole level of instruction. It sensitizes the students to their role in improving the teaching-learning process as this gives them a sense of efficacy that they are able to make suggestions to improve the quality of instruction to meet their needs. Furthermore, the lecturers will be able to obtain direct feedback concerning their performance in the classroom so as to improve themselves, thus working toward excellence in instruction. Student ratings also provides a system whereby lecturers who are held responsible for their performance in the classroom by the Ministry of Education, administrators, students and public at large, can provide tangible evidence of the quality of their work. Furthermore, the administrators will be able to have a systematic representation of students' judgments which can be used as a supplementary and complementary system of evaluation to the present existing system. Subsequently, the results from the two systems of evaluation will be used for decisions involving confirmation of service, salary increment, promotion, loans, transfers, offer of opportunities for further studies, and scholarships.

Table 1 shows the four parts of the proposed system. They are identification of variables related to effective teaching leading to development of questionnaire; administration of questionnaire; analysis, interpretation and report of results; and the uses of the results.

The first part, that is the contents in the trainee rating questionnaire are grouped under seven categories as listed below:

- A. Organization, Structure or Clarity of Subject Matter.
- B. Lecturer-Trainee Interaction and Rapport.
- C. Teaching Skill, Communication or Lecturing Ability.
- D. Course Content and Course Difficulty.
- E. Assignments and Evaluation.
- F. Impact on Students.
- G. Global or Overall Ratings.

It is advisable to administer the questionnaires formally in the classroom. If the lecturer is to administer the questionnaire, he is to distribute the questionnaires and read a standard set of instructions to the class, outlining the process. The lecturer concerned then leaves the room and the monitor or trainee from the class is appointed to oversee the class as he responds to the questionnaire too. When everyone has completed the questionnaires, he must gather the questionnaires, place them in a large envelope, seal the envelope before the class, check the particulars written on the envelope, and deliver the envelope immediately to the designated office.

In the case where the questionnaire is administered by the proctor, the procedure is the same except that he the proctor oversees the class, collects and delivers the questionnaire to the designated office.

When the questionnaire is administered, the students should generally be given 20-25 minutes to fill the form in their regular classroom. They should be given the impression that their frank and honest responses and comments are desired and not have the misconception that it is now their chance to "get back at" the lecturers.

In view that student evaluation arouses sensitivities, it is recommended that the implementation be in two stages. At the introduction stage, the results of the analysis are basically used to help improve the teaching performance of lecturers with non-threatening and non-punitive intentions.

Three tables and a figure will be presented to the individual lecturer concerned. The tables include trainee ratings of the lecturer concerned :percentages and scores, summary of the trainees' comments and suggestions of the lecturer concerned and comparison of teaching performance of lecturers in a department. A figure is presented, showing the comparison of rating profiles of teaching performance of the lecturer concerned with the overall average performance of the lecturers in the college/university. A counseling and discussion session is conducted for individuals whose ratings are not satisfactory so that necessary steps can be discussed and taken to improve the quality of teaching. Subsequently, the lecturer is required to put in a report or proposal to the Administrators requesting for additional training or resources needed to plan and implement the suggested instructional strategies to improve teaching.

Once the proposed system is recognized as an integral part of the evaluation system, a mandatory stage is introduced. At this stage, the results of the analysis will be included in the confidential report or file of the lecturers, so that it can be referred to when required in the application for promotion, salary increment, confirmation of service, application for loans, application for transfers and application for further studies with full/half-pay leave or provision of scholarships for further studies.

The Study

To assess the practicality of the proposed system, a test was carried out in Institut Perguruan Sultan Idris (IPSI), Tanjong Malim, Perak to detect any ambiguity in the procedure used for administering the questionnaire, to assess the procedure of analysis, interpretation and reporting of the results and to analyze the data feedback to help assess the usefulness of the information obtained.

Method and Procedure of the Study

The subjects of the study were seven lecturers from the Education Department, English Studies Department, Humanities Department, Malay Studies Department, and Mathematics and Science Department. Table 2 provides details of the number of lecturers and trainees who participated in the test.

TABLE 2: Details of Group of Lecturers and Number of Trainees Participating in the Pilot-Test

Group of Lecturers	Number of Classes	Number of Trainees
<i>Education Department</i>		
Lecturer I	1	14
Lecturer II	1	17
<i>English Studies Department</i>		
Lecturer III	1	28
Lecturer IV	1	30
<i>Humanities Department</i>		
Lecturer V	1	28
<i>Malay Studies Department</i>		
Lecturer VI	2	58
<i>Science and Mathematics Department</i>		
Lecturer VII	2	58
Total	9	233

In this study, the questionnaire in Bahasa Malaysia was used as requested by the lecturers because the trainees were more proficient in the language.

Sets of survey packages containing the questionnaires and a guide with information on how to use the questionnaires were placed under the charge of a person whose task was to distribute the questionnaires to all the participating lecturers, a week prior to study leave and examinations. He was also the person to receive the completed questionnaires from the trainee monitors. The lecturers were advised to read through the guide before administering the questionnaires. The approach whereby the questionnaires were administered by the lecturer himself was adopted and the administration of the questionnaires was carried out without any problem.

To compute the results, two commercially available software, namely a database management system and an electronic spreadsheet were utilized in combination to set up a storage system for the raw data and to produce the various necessary calculations, tables and graphs.

Analysis and Interpretation of Data

The data analysis was shown in the format to provide the individual lecturer with the feedback about his teaching performance. Three tables and a figure were provided for the individual lecturer concerned. They were the following:

1. Trainee Ratings of the Lecturer Concerned-Percentages and Scores.
2. The Comparison of Teaching Performance of Lecturers in a Department with the Overall Average Performance of the Sample of Lecturers.
3. A Summary of the Trainees' Comments and Suggestions of the Lecturer Concerned.
4. The Comparison of Rating Profiles of Teaching Performance of the Lecturer concerned with the Sample of Lecturers.

The first table displayed a presentation of the trainees ratings of the individual lecturer, using a four-point Likert scale. It showed the percentage responses to the four scale points, i.e. strongly agree, agree, disagree and strongly disagree for each item and each category. It also showed the mean or average score of each item and each category.

The ratings of one lecturer is discussed here to show the usefulness of this system. Table 3 presents the trainee ratings of Lecturer 1. He was generally above average in his teaching performance with a slightly below average score in category D. This was brought about by his low scores of the items on depth of the course covered and difficulty of course material.

Table 3
Trainee Ratings of Lecturer I - Percentage and Scores

Categories and Items	Percentage Response				Score
	SA (4)	A (3)	D (2)	SD (1)	Mean
<i>A. Organization, Structure or Clarity</i>					
1 Preparation	42.86	57.14	0.00	0.00	3.43
2 Content presentation	57.14	42.86	0.00	0.00	3.57
3 Pace	07.14	57.14	35.71	0.00	2.71
4 Explanation	50.00	50.00	0.00	0.00	3.50
5 Interpretation	28.57	64.29	07.14	0.00	3.21
6 Use of time	57.14	35.71	07.14	0.00	3.50
	40.48	51.19	08.33	0.00	3.32
<i>B. Lecturer - Trainee Interaction or Report</i>					
7 Helpfulness and concern	64.29	35.71	00.00	0.00	3.64
8 Interest in individuals	35.71	57.14	07.14	0.00	3.29
9 Sensitivity to trainee's understanding material	14.29	71.43	14.29	0.00	3.00
10 Availability for consultation	50.00	35.71	14.29	0.00	3.36
11 Trainees' freedom to question and express	35.71	57.14	07.14	0.00	3.29
12 Ability to answer questions	78.57	21.43	00.00	0.00	3.79
	46.43	46.43	07.14	0.00	3.40
<i>C. Teaching Skill, Communication or Lecturing Ability</i>					
13 Knowledge of subject matter	42.86	57.14	00.00	0.00	3.43
14 Command of language	14.29	78.57	07.14	0.00	3.07
15 Enthusiasm in subject matter	50.00	50.00	00.00	0.00	3.50
16 Adaptation to trainees' level of comprehension	21.43	64.29	14.29	0.00	3.07
17 Usage of examples and illustrations	42.86	50.00	07.14	0.00	3.36
18 Relation to practical situations	21.43	64.29	14.29	0.00	3.07
19 Encouragement towards independent thinking	28.57	64.29	07.14	0.00	3.21
20 Development of new viewpoints and appreciations	07.14	85.71	07.14	0.00	3.00
21 Sense of honour	57.14	35.71	07.14	0.00	3.50
22 Summary and emphasis	71.43	21.43	07.14	0.00	3.64
23 Conducive atmosphere	14.29	78.57	07.14	0.00	3.07
	33.77	59.09	07.14	0.00	3.27

Table F1 (cont'd)

Categories and Items	Percentage Responses				Score
	SA (4)	A (3)	D (2)	SD (1)	
<i>D. Course Content and Course Difficulty</i>					
24 Depth	7.14	64.29	28.57	0.00	2.79
25 Relevance	21.43	71.43	7.14	0.00	3.14
26 Challenging	28.53	57.14	14.29	0.00	3.14
27 Difficulty	00.00	42.86	57.14	0.00	2.43
	14.29	58.93	26.79	0.00	2.88
<i>E. Assignments and Evaluation</i>					
28 Advice	28.57	71.43	0.00	0.00	3.29
29 Reflection of important aspects	42.86	57.14	0.00	0.00	3.34
30 Instructions	35.71	57.14	7.14	0.00	3.36
31 Feedback	50.00	35.71	14.29	0.00	3.36
32 Information of process	21.43	57.14	21.43	0.00	3.00
	35.71	55.71	8.57	0.00	3.27
<i>F. Impact on Trainees</i>					
33 Amount learnt	35.71	64.29	0.00	0.00	3.37
34 Interest stimulated	21.43	78.57	0.00	0.00	3.21
35 Motivation	57.14	35.71	7.14	0.00	3.50
	38.09	59.52	2.38	0.00	3.36
<i>G. Global or Overall Ratings</i>					
36 Interesting	37.71	64.29	0.00	0.00	3.36
37 Worthwhile	71.43	28.57	0.00	0.00	3.71
38 Good Model	78.57	21.43	0.00	0.00	3.79
	61.90	38.10	0.00	0.00	3.62

Number of trainees = 14

Table 4 enabled the individual lecturer to compare the relative standing of his teaching performance in his department with that of the sample of lecturers in MPSI. The mean score of each item and each category was used for the comparison. In this study however, the mean score of lecturer1 and lecturer 2 in the same department and the overall mean scores of seven lecturers were unweighed by sample size of the trainees, assuming that the size of class did not influence the ratings. As for cases like lecturer V, lecturer VI and lecturer VII who were solitary lecturers evaluated in their individual departments, comparison was made only between the relative standing of their teaching performance with that of the sample of lecturers in the College. Lecturer's 1's performance was above average in all the categories except in category D. Compared to the average performance of the seven lecturers in category D, he fell short in his performance but compared to the average of the two lecturers in his department, he was above average. It showed that the courses offered by the Education Department was too deep and difficult for the students.

Table 4
Comparison of Teaching Performance of Lecturers in a Department
with Overall Average Performance of Sample of Lecturers

Categories and Items	Mean Likert Score		
	Department Lecturer 1	Average of two Lecturer	Sample of Lecturers (n=7)
<i>A. Organization, Structure or Clarity</i>			
1 Preparation	3.43	3.19	3.20
2 Content presentation	3.57	3.08	3.15
3 Pace	2.71	2.53	2.84
4 Explanation	3.50	3.19	3.09
5 Interpretation	3.21	2.90	2.96
6 Use of time	3.50	3.02	3.17
	3.32	2.99	3.07
<i>B. Lecturer- Trainee Interaction or Rapport</i>			
7 Helpfulness and concern	3.64	3.35	3.27
8 Interest in individuals	3.29	3.03	2.81
9 Sensitivity to trainee's understanding material	3.00	2.62	2.72
10 Availability for consultation	3.36	3.01	2.86
11 Trainees' freedom to question and express	3.29	2.86	2.94
12 Ability to answer questions	3.79	3.25	3.16
	3.40	3.02	2.96

C. Teaching Skill, Communication or Lecturing Ability

13 Knowledge of subject matter	3.43	3.13	3.24
14 Command of language	3.07	2.77	3.13
15 Enthusiasm in subject matter	3.50	3.16	3.27
16 Adaptation to trainees' level of comprehension	3.07	2.98	3.04
17 Usage of examples and illustrations	3.36	2.89	3.04
18 Relation to practical situations	3.07	2.92	2.97
19 Encouragement towards independent thinking	3.21	3.05	3.13
20 Development of new viewpoints and appreciations	3.00	2.86	2.90
21 Sense of humour	3.50	3.28	3.25
22 Summary and emphasis	3.64	3.41	3.20
23 Conducive atmosphere	3.07	2.77	2.92
	3.27	3.02	3.1

D. Course Content and Course Difficulty

24 Depth	2.79	2.72	2.95
25 Relevance	3.14	2.93	3.01
26 Challenging	3.14	3.19	3.08
27 Difficulty	2.43	2.34	2.60
	2.88	2.80	2.91

E. Assignments and Evaluation

28 Advice	3.29	3.15	2.83
29 Reflection of important aspects	3.43	3.25	3.05
30 Instructions	3.29	3.06	3.00
31 Feedback	3.36	2.98	2.87
32 Information of progress	3.00	2.86	2.58
	3.27	3.06	2.87

F. Impact on Trainees

33 Amount learnt	3.36	3.12	3.04
34 Interest stimulated	3.21	2.87	2.90
35 Motivation	3.50	3.13	2.95
	3.37	3.04	2.96

G. Global or Overall Ratings

36 Interesting	3.36	3.12	3.03
37 Worthwhile	3.71	3.48	3.34
38 Good model	3.79	3.13	3.17
	3.62	3.24	3.18

Number of trainees = 14

The two tables mentioned above provided the individual lecturer with information of his performance. It also give him an opportunity to compare his relative standing in his department with that of the sample of lecturers in MPSI.

Table 5 presents a list of positive comments and suggestions concerning the teaching performance of Lecturer 1. From this table, the lecturer was made aware of what the trainees thought of his teaching and their suggestions to improve his teaching of the course. The comments and suggestions reaffirmed the ratings on some of the items in Section 1 or they may be new items that needed to be given due consideration too.

Table 5
A Summary of the Trainees' Comments and Suggestions: Lecturer I

I. Comments on what the lecturer has done especially well.

Statement	Frequency (n=14)
Good teaching	5
Showed interest in trainees	3
Clear explanations	3
Interesting and active presentations	1
Hardworking/Diligent	5
Friendly	1

II. Suggestions to improve the lecturer's teaching of the course

Statement	Frequency (n=14)
Be more proficient in Bahasa Malaysia	3
Slow down in presentations	2
Slow down for topics that are difficult	2
Be more sensitive to our lack of understanding of the material	1
Be more creative	1
Be stricter in class control	1
Give extra classes	2
Give more examples	1
Have more discussions and interactions	2

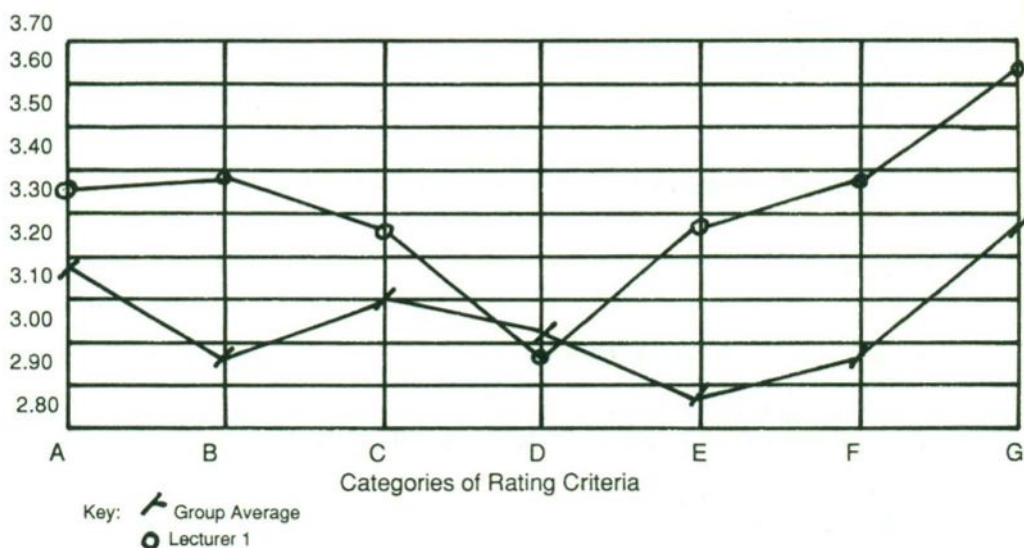
Figure 1 provides a graphic presentation of lecturer 1's mean score and the overall mean score of the sample of lecturers in MPSI. The categories of rating criteria which included organization, structure and clarity; lecturer-trainee interaction or rapport; teaching skill, communication or lecturing ability; course content and course difficulty; assignments and evaluation; impact on trainees; and global or overall ratings, were used for comparison. At one glance, lecturer 1 was able to see whether his performance based on the categories stated was above or below the overall mean score of the sample of lecturers. It was clear that he had scored very high in the categories on lecturer-trainee interaction or rapport, assignments and evaluation, impact on trainees, and global or overall ratings. His score for course content and course difficulty was weak. By referring to Table 3 and Table 4, he was able to identify that he needed to look into the depth of the course content covered and difficulty of course material.

After the results were shown to the sample of lecturers, the researcher interviewed them and they confirmed that the data feedback do provide useful information to them as individuals. They now have sufficient information on what they are good at so as to keep up their areas of strength and to seek ways to correct their weaknesses, thus striving continuously to master the art of effective teaching.

Figure 1

Comparison of Rating Profiles of Teaching Performance of Lecturer 1 and Sample of Lecturer

Likert Scale



Uses of the Results

All the respondents accepted the idea of using the results of trainee ratings for professional improvement and to support the lecturer's application for promotion, salary increment, confirmation of service, loans, transfers, further studies with full or half pay leave and scholarships.

Conclusions

The present system of evaluation for lecture is part of a policy requirement which is mandatory. Basically, the emphasis is on evaluation for the purpose of administrative decisions rather than for the purpose of development and improvement of instruction. As an exercise, it is initiated and implemented by the administrators once a year. The question raised about the present system of evaluation is that the information provided through the present system regarding the lecturer's ability to deliver ideas and instruction with clarity and orderliness, his knowledge of the job, and his interest in his job, are insufficient to confirm a lecturer's performance in the classroom with the range of "not satisfactory" to "excellent". Basically, the ratings by the administrators tend to generally reflect the overall reputation of the lecturer rather than judgment based on actual observations in the classroom. Therefore, trainee ratings constitute an important source of measuring the lecturer's teaching performance. It is an important facet in the lecturer's evaluation but is not the only facet. It is seen as a part of the other contributing facets which include administrators evaluation that is in existence, colleagues evaluation, alumni evaluation, professional evaluators evaluation, parents evaluation, public evaluation, and self evaluation. Since the ratings by the administrators reflect the general reputation of a lecturer rather than judgment based on actual classroom observations, the proposed system as a supplementary and complementary evaluation to the present system of evaluation will ensure broader based system of evaluation.

With the limitations of the present system, there is a need for the implementation of an additional system for the evaluation of lecturers by trainees as they are an integral source of information on effective teaching. There is implication of a need to maintain and extend such a system so that lecturers will benefit in terms of improving teaching and obtaining acknowledgement for excellence in performance. By and large, the administrators will also be able to obtain feedback from another relevant source about the classroom performance of the lecturers, with the intention of providing counseling or discussion sessions to guide the lecturers in improving their teaching, and at the same time to reward those who have shown excellent performance.

Procedure for administering, analyzing, interpreting and reporting the results are important considerations involved in planning the evaluation of the teaching performance of lecturers by trainees. The implementation of this proposed system needs to be constantly monitored by the administrators and the authorities concerned. To ensure its success, issues like cooperation from the lecturers and trainees, allotment of time, availability of facilities and materials, professional development program and reward system need to be looked into. In addition, a meaningful evaluation can only be successfully carried out if the lecturers do not feel threatened by the trainee ratings.

The present system together with the proposed system of evaluation may have valuable implication for the development of a system of accountability. This is because the present system is used as a source to obtain feedback concerning the overall performance of the lecturers annually, whereas the proposed system is used as a source to assess the progress of the lecturer's teaching performance in the classroom and provide useful feedback concerning the individual lecturer.

Appendix A

A questionnaire was generated with thirty-eight items grouped under seven categories related to effective teaching in the teacher training colleges. They are:

A. Organization, Structure or Clarity of Subject Matter

1. Preparation for each class.
2. Presentation of material.
3. Pace or rate at which the material was covered.
4. Explanation of course material.
5. Interpretation of difficult or abstract ideas.
6. Use of class time.

B. Lecturer-Student Interaction or Rapport

7. Helpfulness and concern with whether students have learnt the material.
8. Interest in students as individuals.
9. Sensitivity to whether students have understood the material.
10. Availability for consultation with students.
11. Students felt free to ask questions or express opinions.
12. Ability to respond to questions raised by students.

C. Teaching Skill, Communication or Lecturing Ability

13. Knowledge of the subject matter.
14. Command of the language.
15. Enthusiasm about the subject matter.
16. Adaptation of the course to students' level of comprehension.
17. Expertise in selecting and using examples or illustrations.
18. Ability to relate course to practical situations.
19. Encouragement towards independent thinking.
20. Development of new viewpoints and appreciations.
21. Sense of humor.
22. Ability to summarize or emphasize major points in lectures or discussions.
23. Provision of a conducive atmosphere for work and learning.

D. Course Content and Course Difficulty

24. Depth of the course content covered.
25. Relevance of the material covered.
26. Challenging course.
27. Difficulty of course material.

E. Assignments and Evaluation

28. Advice on how to study for the course and examinations.
29. Important aspects of the course being reflected in the tests and examinations.
30. Instructions given for course work or assignments.
31. Feedback provided on course work or assignments.
32. Students informed of their progress.

F. Impact on Students

33. Amount learnt by students.
34. Interest stimulated in the subject area.
35. Motivation.

G. Global or Overall Ratings

36. Interesting course.
37. Value of the course.
38. Lecturer serving as a model.

References

- Abdul Aziz b. Mohd. Sultan. *Rasional, Objektif dan Komponen Kurikulum Perguruan Asas Tiga Tahun*. Kertas kerja dibentangkan di Program Orientasi Pensyarah-Pensyarah Maktab Perguruan Malaysia, Institusi Perguruan Pendidikan Negara, Sri Layang, Genting Highland, 1-3 Julai, 1985.
- Aleamoni, L.M. Development and Factorial Validation of the Arizona Course/Instructor Evaluation Questionnaire. *Educational and Psychological Measurement* 38(4), 1978: 1063-1067.
- _____. Student Ratings of Instruction. In Millman, J. (Ed.), *Handbook of Teacher Evaluation*. Beverly Hills, California: Sage Publications, 1981: 110-145.
- American Academy of Arts and Sciences. *Assembly on University Goals and Governance: A First Report*. Cambridge, Massachusetts: The Assembly, 1971.
- American Association of University Professors. *Project to Improve College Teaching: Newsletter*. Washington, D.C., 1970
- ASEAN Development Education Project. *Evaluation on the Effectiveness of Foundation /Professional Courses in Pre-Service Teacher Education Programmes in Malaysia*. Kuala Lumpur: Teacher Education Division, Ministry of Education Malaysia, October, 1983.
- Astin, A.W. and Lee, C.B.T. Current Practices in the Evaluation and Training of College Teachers. *Educational Record* 47, 1966: 361-375.
- Awang Had Salleh. *Malay Teacher Training in British Malaya (1878-1941)*. A General Survey. Unpublished B.Ed. (Post Graduate) Dissertation, University of Malaya, Kuala Lumpur, 1967.
- _____. *Malay Secular Education and Teacher Training in British Malaya (with special reference to the Sultan Idris Training College)*. Kuala Lumpur: Dewan Bahasa dan Pustaka, 1979.
- Bandura, A. *Social Learning Theory*. Englewood Cliffs, N.J.: Prentice Hall, 1977.
- Bejar, I.I. A Survey of Selected Administrative Practices Supporting Student Evaluation of Instructional Programs. *Research in Higher Education* 3, 1975: 77-86.
- Brickman, W.W. Student Evaluation of Professors. *School and Society* 94(3), 1966: 143-144.
- Broudy, H.S. Can We Define Good Teaching? *The Record Teachers College* 70(April), 1969: 583-584.
- Burton, C.E. *College Teaching*. New York: Harcourt Brace Jovanovich, 1956.
- Burton, T.H. and Prichard, K.W. *Excellence in University Teaching*. Columbia: University of South Carolina Press, 1975.
- _____. *Cenderamata* (SITC, Tanjung Malim), October, 1923.
- Centra, J.A. The Student as Godfather? The Impact of Student Ratings on Academia. *Educational Researcher* 2(8), 1973: 4-8.

- _____. Effectiveness of Student Feedback in Modifying College Instruction. *Journal of Educational Psychology* 65(3), 1973: 395-401.
- _____. *How Universities Evaluate Faculty Performance: A Survey of Department Heads (GREB no. 75-Sbr)*. Princeton, J.J.: Educational Testing Service, 1977a.
- _____. The How and Why of Evaluating Teaching. In Centra, J.A.(Ed.), *Reviewing and Evaluating Teaching*. San Francisco: Jossey-Bass Inc., Publishers, 1977b.
- _____. *Determining Faculty Effectiveness*. San Francisco: Jossey Bass Inc., Publishers, 1979.
- _____. Using Student Assessments to Improve Performance and Vitality. *New Directions for Institutional Research* 20, 1978: 31-49.
- Cheah, B.K. Sultan Idris College, Cradle of Malay Resurgence. *The Straits Times*, 12 June, 1973.
- Cheeseman, H.R. *Education in Malaya 1930 to 1942 and 1946*. Kuala Lumpur: The Government Press, 1948.
- Clark, M.J. and Blackburn, R.T. *An Assessment of Faculty Performance: Some Correlates between Administrator, Colleague, Student and Self Ratings*. Ann Arbor: University of Michigan, 1971.
- Coats, W.D. and Swierenga, L. Student Perceptions of Teachers - A Factor Analytic Study. *The Journal of Educational Research* 65(8), 1972: 357-360.
- Cohen, P.A. Student Ratings of Instruction and Student Achievement: A Meta-Analysis of Multisection Validity Studies. *Review of Educational Research* 51(3), 1981: 283-284.
- Costin, F. *Survey of Opinions about Lectures*. Mimeographed, University of Illinois, Department of Psychology, 1968.
- Costin, F., Greenough, W.T. and Menges, R.J. Student Ratings of College Teaching: Reliability, Validity and Usefulness. *Review of Educational Research* 41, 1972: 511-535.
- Costin, F. and Grush, J.E. Personality Correlates of Teacher-Student Behavior in the College Classroom. *Journal of Educational Psychology* 65, 1973: 35-44.
- Crawford, P.L. and Bradshaw, H.L. Perception of Characteristics of Effective University Teachers: A Scaling Analysis. *Educational and Psychological Measurement* 28, 1968: 1079-1085.
- Crittenden, K.S. and Norr, J.L. Student Values and Teacher Evaluation: A Problem in Person Perception. *Sociometry* 36(2), 1973: 143-151.
- Deshpande, A.S., Webb, S.C. and Marks, E. Student Perceptions of Engineering Instructor Behaviors and Their Relationships to the Evaluation of Instructors and Courses. *American Educational Research Journal* 7, 1970: 289-305.
- Downie, N.W. Student Evaluation of Faculty. *Journal of Higher Education* 23, 1952: 495-496, 503.
- Dunkin, M.J. and Biddle, B.J. *The Study of Teaching*. New York: Holt, Rinehart and Winston, 1974.

- Feldman, K.A. Consistency and Variability Among College Students in Rating Their Teachers and Courses: A Review and Analysis. *Research in Higher Education* 6, 1977: 233-274.
- _____. Course Characteristics and College Students' Ratings of Their Teachers; What We Know and What We Don't. *Research in Higher Education* 9, 1978: 199-242.
- _____. The Significance of Circumstances for College Students' Ratings of Their Teachers and Courses. *Research in Higher Education* 10, 1979: 149-172.
- French, G.M. College Students' Concept of Effective Teaching Determined by an Analysis of Teacher Ratings. *Dissertation Abstracts* 17, 1957: 1380-1381.
- Frey, P.W. A Two-Dimensional Analysis of Student Ratings of Instruction. *Research in Higher Education* 9, 1978: 69-91.
- Gadzella, B.M. College Student Views and Ratings of an Ideal Professor. *College and University* 44, 1968: 89-96.
- Gaff, J.G. and Wilson, R.C. Faculty Cultures and Inter-Disciplinary Studies. *Journal of Higher Education* 42, 1971: 186-201.
- Gage, N.L. What Do We Know About Teaching Effectiveness? *Phi Delta Kappan* October, 1984: 87-93.
- Greenwood, G.E., Bridges, C.M., Ware, W.B. and McLean, J.E. Student Evaluation College Teaching Behaviors Instrument: A Factor Analysis. *Journal of Higher Education* 44(8), 1973: 596-604.
- Gustad, J.W. Policies and Practices in Faculty Evaluation. *Educational Record* 42, 1961: 194-211.
- Hamzah bin Salas. *Program Pendidikan Guru*. Tanjung Malim: Maktab Perguruan Sultan Idris, 4 April, 1985.
- Herman, J.J. *Developing an Effective School Staff Evaluation Programme*. New York: Parker Publishing Company, Inc., 1973.
- Hoffman, R.W. Students Portray the Excellent Teacher. *Improving College and University Teaching* 11 (Winter), 1963: 21-24.
- Howard, G.S. and Maxwell, S.E. Do Grades Contaminate Student Evaluations of Instruction? *Research in Higher Education* 16, 1982: 175-188.
- Kaplan, R.M. Reflections on the Dr. Fox Paradigm. *Journal of Medical Education* 49, 1974: 310-312.
- Keller, W.D. On Teaching and Learning. In Buxton, T.H. and Prichard, K.W., *Excellence in University Teaching: New Essays*. Columbia, South Carolina: University of South Carolina Press, 1975.
- Kent, L. Students' Evaluation of Teaching. *Educational Record* 47 (Summer), 1966: 376-496.
- Kulik, J.A. and McKeachie, W.J. The Evaluation of Teachers in Higher Education. In Kerlinger, F.N. (Ed.), *Review of Research in Education*, Vol. 3, Itasca, I.L.: F.E. Peacock Publishers, 1975.

- Labyrinth* No. 5, November 1979: 10-30. Cited in Murray, H.G., Evaluation of University Teaching: A Selective Bibliography. In *Newsletter* (Ontario University Program for Instructional Development), 1980.
- Langen, T.D.F. Student Assessment of Teaching Effectiveness. *Improving College and University Teaching* 14 (Winter), 1966: 22-25.
- Marsh, H.W. Students' Evaluations of University Teaching: Dimensionality, Reliability, Validity, Potential Biases and Utility. *Journal of Educational Psychology* 76, 1984: 707-754.
- Marsh, H.W., Fleiner, H. and Thomas, C.S. Validity and Usefulness of Student Evaluations of Instructional Quality. *Journal of Educational Psychology* 67(6), 1975: 833-839.
- McKeachie, W.J. Teaching Tips, *A Guidebook for the Beginning College Teacher* (6th Ed.). Massachusetts: D.C. Heath and Co., 1969: 215.
- _____. Student Ratings of Faculty: A Reprise. *Academe* 65, 1979: 384-397.
- Menges, R.J. The New Reporters: Students Rate Instruction. In Pace, C.R. (Ed.), *Evaluating Learning and Teaching: New Directions for Higher Education*. San Francisco: Jossey-Bass, No. 4, 1973.
- Miller, M.T. Instructor Attitudes Toward, and Their Use Of, Student Ratings of Teachers. *Journal of Educational Psychology* 62(3), 1971: 235-239.
- Miller, R.I. *Developing Programmes for Faculty Evaluation*. San Francisco: Jossey-Bass, Inc., Publishers, 1974.
- _____. *Evaluating Faculty Performance*. San Francisco: Jossey-Bass, Inc., Publishers, 1974.
- Moomaw, W.E. Faculty Evaluation: The State of the Art in the South. *Faculty Development and Evaluation in Higher Education* 3, 1977: 6-14.
- Mowrer, O.H. *Learning Theory and Behavior*. New York: John Wiley and Sons, 1960.
- Murray, H.G. Predicting Student Ratings of College Teaching from Peer Ratings of Personality Traits. *Teaching of Psychology* 2, 1975: 66-69.
- _____. Teaching Ratings, Student Achievement, and Teacher Personality Traits. Paper presented at the Annual Meeting of the Canadian Psychological Association, 1978.
- _____. Low-Inference Classroom Teaching Behaviors and Student Ratings of College Teaching Effectiveness. *Journal of Educational Psychology* 75, 1983: 138-149.
- Naftudin, D.H., Ware, J.E. Jr. and Donnelly, F.A. The Dr. Fox Lecture: A Paradigm of Educational Seduction. *Journal of Medical Education* 48, 1973: 630-635.
- Newman, A.D. and Rowbottom, R.W. *Organizational Analysis*. Illinois: Southern Illinois University Press, 1968.
- _____. *New Straits Times*, 7 October, 1980.

- Oviatt, A.B. Reflections on Effective Teaching. In Buxton, T.H. and Prichard, K.W. *Teaching: New Essays*. Columbia, South Carolina: University of South Carolina Press, 1975.
- Pendidikan Guru* (Terbitan Khas, Sempena Perayaan Jubli Intan, MPSI) 29 November, 1982.
- Prichard, K.W. Teaching Education Courses: The Other Side of the Tracks. In Buxton, T.H. and Prichard, K.W., *Excellence in University Teaching: New Essays*. Columbia, South Carolina: University of South Carolina Press, 1975.
- Principal's Dairy* (Sultan Idris Training College, Tanjung Malim), 1941.
- Proceedings of the Federal Council*. Federated Malay States: Government Press, 1917.
- Rushton, J.P., Murray, H.G. and Paunonen, S.V. Personality, Research Creativity, and Teaching Effectiveness in University Professors. *Scientometrics* 5, 1983: 93-116.
- Scheck, D.C. and Bizio, S. Students' Perceptions of the Ideal Professor. *College Student Journal* 11(4), 1977: 335-343.
- Scott, C.S. Some Remarks on Student Ratings: Validation. *American Educational Research Journal* 12(4), 1975: 444-445.
- Seldin, P. New Emphasis in the Evaluation of Professor. *PhiDelta Kappan* 56, 1975: 496-497.
- Sherman, B.R. and Blackburn, R.T. Personal Characteristics and Teaching Effectiveness of College Faculty. *Journal of Educational Psychology* 67, 1975: 124-131.
- Smith, A.A. College Teachers Evaluated by Students. *Sociology and Social Research* 28, 1944: 471-478.
- The Times of Malaya* (Weekly). Ipoh, 30 November, 1922.
- Tomasco, A.T. Student Perceptions of Instructional and Personality Characteristics of Faculty: A Canonical Analysis. *Teaching of Psychology* 7, 1980: 79-82.
- Tuckman, B.W. and Oliver, W.F. Effectiveness of Feedback to Teachers as a Function of Source. *Journal of Educational Psychology* 59(4), 1968: 297-301. Winstedt, R.O. *Education in Malaya*. Singapore: Fraser and Neave Ltd., Printers, 1923.
- Wong, F.H.K. and Chang, P.M.P. *The Changing Pattern of Teacher Education in Malaysia*. Kuala Lumpur: Heinemann Educational Books (Asia) Ltd., 1975: 12-15.
- Wotruba, T.R. and Wright, P.L. How to Develop a Teacher-Rating Instrument: A Research Approach. *Journal of Higher Education* 46(6), 1975: 653-663.
- Zainuddin bin Abu. Sedikit Kenangan-Kenangan untuk Albam. *Albam Kenangan*, 1955-1956: 36-37.