

Experiential Learning Enters the Eighties, Volume 7, 1980

DESIGNING A COMPETENCY-BASED PEER ASSESSMENT SCALE FOR THE EVALUATION OF TEACHING IN HIGHER EDUCATION

Arthur Whatley, New Mexico State University
Thomas E. Cyrs, Jr., New Mexico State University
Pamela B. Eckard, New Mexico State University
Rafaela C. Acosta, New Mexico State University

ABSTRACT

A predominant theme in current literature on faculty evaluation is that the judging of the quality of instruction is much too complicated to be based solely on the filling out of rating forms by students, on administrator's personal assessment, or the consensus judgment of peers. Each of these is useful but no one is sufficient. Present literature yields little information in the area of peer assessment of faculty. This paper will detail the procedures used to develop a competency based faculty peer evaluation instrument for higher education as well as address several administrative/political issues likely to be raised by faculty and administrators.

INTRODUCTION

If this profession should prove itself unwilling to purge its ranks of the incompetent and the unworthy, or to prevent the freedom which it claims from being used as a shelter for inefficiency, for superficiality, or for uncritical and intemperate partisanship, it is certain that the task will be performed by others--by others who lack certain essential qualifications for performing it, and whose actions are sure to breed suspicions and recurrent controversies deeply injurious to the internal order and to the public standing of universities."

AAUP-1915

The bulk of current literature on faculty evaluation in higher education concerns student evaluation of instruction. A predominant theme in this literature is that the judging of the quality of instruction is a far too complicated process to be based solely on an administrator's personal assessment or the consensus judgment of peers or on the filling out of rating forms by students. Each of these is useful but no one is sufficient.

For sometime, the Center for Educational Development and Evaluation has focused, as one of its tasks, on peer assessment of faculty in higher education. The literature on faculty evaluation in higher education yields little information in the area of peer assessment of faculty. For the purpose of our study, peer is defined as a faculty member possessing qualities deemed suitable to act in the capacity of making judgments about another faculty member's teaching performance [4, pp. 165-6]. After conducting four cross-indexed computer searches, bibliographical reviews, telephone searches to known authors in the field of peer evaluation of faculty in higher education, and letters to centers of research and analysis in this area requesting information, articles, and additional sources of data on peer assessment, the Center compiled an initial file of over 500 references. This number was narrowed to just over two hundred of these.

Twenty-seven references were isolated which dealt with peer assessment in some detail. Of these twenty-seven articles, twenty were categorized as descriptive and seven as empirical studies. The major finding of this review of literature is that it appears that no systematic process of assessment of faculty in higher education has been developed that employs specific, peer-generated criteria for use by colleagues to rate their peers.

Why Peer Evaluation?

While studies show that students are qualified to rate classroom instruction on certain specific criteria such as:

1. Was teacher able to answer questions? To what degree?
2. Was student able to learn something new?
3. Was student given a syllabus for the class?

they are not in a position to make judgments in other areas, and should not be asked to do so [3]. Items such as:

1. Was material learned superficial or out of date?
2. Did the instructor demonstrate in-depth knowledge of his field?

and judgments about the accuracy, currentness, or sophistication of a teacher's knowledge can only be made by faculty peers conversant with the same field [3].

Centra [2] states that colleagues can play an important role in faculty evaluation:

1. Colleagues from the same or similar departments could contribute to a fair assessment of a faculty member's service and research activities, providing a quality as well as a quantity dimension to that aspect of faculty performance.

Additional aspects or dimensions of instruction which can best be rated by peers are given by Braskamp [1]:

1. Course organization
 - a. The syllabus adequately outlines the sequence of topics to be covered.
 - b. The outline and sequence of topics is logical.
2. Reading, Project, and Laboratory Assignments
 - a. Readings are appropriate for level of course.
 - b. The written assignments and projects reflect course goals.
3. Participation in University Community
 - a. The instructor is involved in student organized and sponsored activities.
 - b. The instructor participates in departmental seminars, activities, etc.

Experiential Learning Enters the Eighties, Volume 7, 1980

Finally, Shore [5] states that:

“Teacher improvement must be the primary focus of all assessment, though such assessment may serve secondarily to aid in promotion, tenure and RIF decisions. This represents a crucial shift away from the idea of evaluation as a judgmental tool which will set colleagues against one another. Instead, assessment should be a continuous, constructive process which will provide a teacher with feedback, guidance, and opportunities for improvement.”

IDENTIFYING COMPETENCIES

Our approach to the development of a peer assessment scale is to identify conditions (areas of competencies) under which effective teaching takes place. An area of competence is a collection of skills and/or attitudes which summarizes a range or group of lesser skills and/or attitudes and is expressed in performance terms. These areas are frequently clustered in a variety of responses as determined by the variables in a given situation. Areas of competence are divided into professional dimensions such as “exhibits in-depth knowledge of subject” or “participates in curriculum development” and personal performance dimensions which includes “demonstrates personal availability to students” and “manages classroom time efficiently.”

Areas of competence should not be too global and all encompassing. They should be manageable yet not too specific. An area such as “gets along with students” is too global and inclusive. From this we could derive “respects students’ rights or ‘meets with students individually” as more specific yet not too specific for purposes of program development. How areas of competence are finally summarized is a matter of judgment to be decided upon by the expert panels

EXPERT PANELS

The process of deriving competencies through consensual judgment is a deductive technique utilizing expert panels. By consensus is meant, 100% agreement. This process is relatively easy to use and can be one of the least expensive. Essentially, panels of experts with common areas of interest are brought together and asked to identify their perceptions of the competencies necessary for good teaching. Through free association of ideas within the panel, these perceptions are further clarified, expanded and modified until the panel reaches consensus

Usually, two or three panels meet for one or two days in a common location, but debate separately. Each panel is composed of 6-10 members. Membership usually includes representative faculty, staff, administrators, and possibly consumers. Each panel is made up of the same number of representatives and is assigned a leader. The group leader is responsible for moving the panel towards its goal within the prescribed time framework. If more than one panel is working on the same problem, a process observer randomly visits each panel throughout the debate to insure consistency of approach with the panel leaders.

After the panels have derived the areas of competence, the leaders edit all data into one document and return this edited document to each panel member to review and approve. This is done to insure that the editorial process did not modify the content or intent of the individual panels.

Criteria for panel membership include: high level of

expertise; respected opinion either regionally, statewide, or nationally; prestigious in the field; well- published; nominated by peers.

CHARGE TO EXPERT PANEL MEMBERS

By the end of the session each panel will have identified as many areas of competence (personal and professional dimensions) as time will allow. Each area of competence will be accompanied by a series of behavioral descriptors which clarify and limit the meaning of each competency. Some of the criteria for an acceptable competency are:

1. There is consensus within each panel as to the relevancy of the need for effective and competent practice.
2. The area of competence is based on a contemporary, realistic framework as opposed to something in the near future.
3. The area of competence relates to training needs generally and not restricted exclusively to some area of specialty.
4. Each area of competency and accompanying behavioral descriptors should communicate the exact intent of the panel in as precise language as possible.

Below are examples of competencies in both the Personal and Professional Dimension:

Personal Dimensions

1. Demonstrates Respect For Students in Interpersonal Communications
 - does not talk down to students
 - respects personal rights of students
 - responds at students’ level of understanding
 - responds to students’ questions in a positive manner
2. Practices Professional Teaching Ethics
 - recognizes students’ rights
 - gives recognition to students for their original ideas
 - evaluates students only on class related objectives and activities
 - distinguishes between personal and professional judgment

Professional Dimensions

1. Manages Classroom Time Efficiently
 - prepares detailed plan for presentation
 - communicates expected readings and assignments
 - begins on time and stops on time
 - provides adequate time for discussions
2. Personalizes Instruction When Appropriate
 - builds on students’ prior academic experience has reasonable expectations
 - gives advice based on students’ background
 - recognizes students’ individual uniqueness

Experiential Learning Enters the Eighties, Volume 7, 1980

CRITICAL INCIDENTS

Once the initial list of competencies is developed, the next step is to generate a range (good to poor) of critical incidents for each competency. To do this we divide the initial list of 31 competencies into several shorter lists of five or six. These shorter lists are then distributed to the entire faculty. Each faculty member is asked to examine the list of 5-6 competencies and:

1. for each competency list two critical incidents of effective behavior, and
2. for each competency list two critical incidents of poor behavior.

The procedure will generate several hundred critical incidents for each competency. These incidents will then be edited and refined to a hierarchical list of 10-15 incidences for each competency. At this point the final instrument is complete--twenty-five or so competencies each with an accompanying hierarchical list of critical incidents.

PEER EVALUATION ISSUES

The basic issue of peer evaluation evolves around how evaluation information will be used. Of concern here is whether or not such data should be used in promotion, tenure, and merit pay considerations all part of the administrations appraisal system. When used in this way, several concerns must be addressed:

1. What happens to the faculty person (assuming the voluntary use of the instrument) who chooses not to be evaluated in this way? Is this person then guilty of poor teaching by implication?
2. Reliability and validity of the instrument must be provided when it is used to affect any employment condition if legal and ethical standards are to be met. Increasingly, institutions of higher education are being asked by government to substantiate the job relatedness of its faculty employment decisions. Furthermore, we owe it to our colleagues and to the teaching profession to establish credible evaluation procedures.

The second way in which this data can be used is in faculty development. Peer evaluation data using this instrument can be fed back to the faculty member and used for improving teaching. This is less controversial than the evaluation issue. Furthermore, the validity and reliability issues become less significant because employment conditions are not affected.

Yet, even if the data is used only for development purposes, a subtle implication remains because faculty are usually required to provide evidence for their teaching performance. If they do not use peer evaluation, what are they to use? Student evaluations as traditionally defined are of dubious validity; administrative evaluations are based on halltalk, hearsay, and unsubstantiated perceptions. In short, the pressures of the situation necessitate the use of the data for evaluation as well as developmental purposes. In the absence of competency-derived peer evaluation data, the entire process becomes mired in ambiguity.

An issue of secondary importance is: who is to do the evaluation and how often are the evaluations to occur? We feel that two or more peers chosen by the person being evaluated should do the evaluation. Two or more evaluators allow for checks on the internal consistency of the evaluation. The frequency of the evaluation can vary considerably. Surely several class periods should be used

staggered throughout the semester. Another approach might be to evaluate over a two-three week period.

Norming the instrument is also an important issue. Evaluation data in the abstract, without norms, is of questionable value. Norms provide a basis of comparison with other faculty. But developing norming data can be complex: can comparisons be made between graduate courses and undergraduate courses, courses with different subject matter, i.e., accounting vs. management, lecture vs. case courses, etc. Obviously considerable effort is required to develop useable norming data.

SUMMARY

As far as we know, ours is the first university-wide attempt to develop a competency-based peer review instrument. In light of EEO requirement to make conditions of employment job-related, societal pressure for accountability as well as increasing professional criticisms regarding traditional evaluation approaches, we feel we have taken an important first step to resolving many of these issues. Faculty acceptance is an all-important variable that is unknown as of this writing. In January, 1980, the entire faculty of New Mexico State University will be involved in the process of developing critical incidents and in general providing input into the project. Their support, assuming it is forthcoming, will allow us to implement an approach to the evaluation of teaching that has never before been tried. We are excited about the possibilities of our efforts for other organizations.

REFERENCES

- [1] Braskamp, L. A. "Colleague Evaluation of Instruction," Undated Monograph, Measurement Research Division, University of Illinois: Urbana, Illinois.
- [2] Centra, John A. "Colleagues as Raters of Classroom Instruction," Journal of Higher Education, 46(1), May/June, 1975, pp. 327-337.
- [3] French-Lazovik, Grace "Evaluation of College Teaching: Guidelines for Summative and Formative Procedures." Monograph prepared under the auspices of the Association's Commission on Liberal Learning, Washington, D.C.: Association of American Colleges, 1975.
- [4] Sauter, Robert C. and Walker, James K. "A Theoretical Model for Faculty Peer Evaluation," American Journal of Pharmaceutical Education, 40(2), May, 1976, pp. 165-6.
- [5] Shore, Jane E. "Faculty Evaluation in Higher Education," The Generator, 8(3), Fall, 1978, pp. 1-3.
- [6] AAUP 1915, Declaration of Principles quoted in Faculty Tenure. A report and recommendation by the Commission on Academic Tenure in Higher Education. Jossey Bass Publishers: San Francisco, 1973, pp. 43-44.