

Factors that Impact the Evaluation of Teacher Evaluation of Teacher Effectiveness for Reappointment, Tenure and Promotion

Nelson Ngoh // University of Bridgeport

ABSTRACT

Institutions of learning hire faculty to teach and carry out certain contractual duties. The onus rests on the institution to retain those who are effectively carrying out the job for which they were appointed and to replace those who are not meeting up with the required assignments. The problem now becomes that of choosing the most appropriate instruments and the best qualified personnel to help evaluate these faculty in order to make the most informed decisions in this respect. This author chose some universities and closely examined the instruments used for this purpose as well as the personnel who are involved in determining the effectiveness of the teachers. The results revealed that some of the questions asked in the Faculty Evaluation Forms were ambiguous and therefore not valid for that purpose. Some of the personnel involved in this process of evaluation were not adequately qualified to determine the right faculty for reappointment, tenure and promotion. The article concludes with some recommendations to ameliorate and improve the validity and reliability of the evaluation.

INTRODUCTION

Evaluation is an important concept in education. Teaching, learning and evaluation are interwoven. Evaluation seems to be the foundation on which teaching and evaluation stand. Most of the time, appropriate decisions as to when or how to teach or learn are based on the results of evaluation. Evaluation carries the notion that the value or worth of someone or something is to be judged. In other words, evaluation is defined as the systematic process of determining the extent to which instructional objectives have been achieved by pupils. Evaluation is equally the means of determining which faculty member is to be hired, reappointed, tenured or promoted. Evaluation is inevitable in teaching and in all fields of activity including reappointment, tenure and promotion where judgments are needed.

It must be noted that the role of evaluation is intrinsic to the teaching-learning situation. Unfortunately, it would appear evaluation of students and faculty alike is all too often done as though it were extraneous to the main purpose of teaching (Ngoh, 2001). Proper care is not taken to select the most appropriate tools as well as the right and experienced individuals to carry out the evaluation. To be valid and reliable, evaluation is expected to exclude casual uncontrolled observation of the individuals concerned and emphasize on already identified and stated objectives.

For this paper, evaluation is the process by which the effects and the effectiveness of the teacher and teaching itself are being determined for the purpose of reappointment,

tenure and promotion to a higher academic rank. In this case it is the systematic process of determining the extent to which the faculty in question has discharged his/her duty as a teacher and as an academician or professional in his/her discipline. In this case, the paper argues that those doing the evaluation of faculty must be very qualified or knowledgeable in the meaning and purpose of evaluation or else the end results would be fraught with bias and subjectivity. Analogously, since you cannot give what you do not have, you cannot effectively evaluate when you do not know or do not have the appropriate experience or knowledge required. Furthermore, it is not only the status of the persons determining the effectiveness of the teacher that is of great importance. The type of instruments provided for this purpose could make the results of the evaluation inauthentic. For example, the type of questions asked about an individual could be misleading.

BACKGROUND

We may start by asking the main purpose of faculty evaluations. Are we right if we give the reason to be simply to promote faculty from Assistant through Associate to Full Professor? What might be the hidden intent of this? Could it be to increase the salaries as they move up the rungs? Could it be to motivate faculty to become more effective teachers? What does it mean to be an effective teacher? Are the authorities worried that the students are not learning well and the quality of graduates from their institutions are not high enough? According to Reeves (2004), Layne (2012) and Feldman & Lynch, (1988), effective teaching results in learning.

It is commonly noted that many students describe effective teachers as those who are interesting, knowledgeable, approachable, and helpful, who present materials well and motivate students by setting high standards. If learning is what is expected of teachers, will these descriptors necessarily lead to learning? What of other intervening variables that can offset these qualities? There are certain situations wherein students do not want to do science because of the stereotypes that science is difficult, or the students do not have the prerequisite background knowledge for that course. Such students fail the course and the faculty bear the blame and the brunt even when they qualify as effective teachers by all standards. It has been said (Weimer, 2013) that students are the ultimate “deciders” when it comes to whether or not they learn. Should faculty be evaluated based on their teaching or students’ learning? Which descriptors are thus most appropriate to be used for reappointment, tenure and promotion?

The main purpose of classroom instruction or teaching is to help pupils achieve a set of intended learning objectives. These educational objectives or outcomes include intellectual, emotional and physical changes. To see if these desired changes (learning) have taken place in the pupils, evaluation must be done. That is, the pupils' learning progress is periodically evaluated by tests and other evaluation devices. Similarly, the academic growth and performance of the faculty are periodically evaluated at many levels by students, colleagues and administrators. Therefore, the interdependent nature of teaching, learning and evaluation should often be recognized, particularly as Hoffman (1962) says that a pupil does not really know what he has learned until he has organized it and explained it to someone else. This should be equally true with faculty. In the process of preparing a file for evaluation, the faculty revisits her/his performance or achievements and realizes the level of his progress in the area of teaching, collaboration with colleagues,

scholarly activities, and their contributions to the institution and the society as a whole. The purpose of this paper therefore is to examine closely the instruments used and the qualification of those involved in the process of evaluating faculty for reappointment, tenure and promotion issues.

The evaluation of the faculty is based on the extent and quality of his/her function within the intellectual life of the University, in addition to fulfilling contractual obligations. The concern or premise here is that we are usually found assessing things that have very little bearing with what the faculty are expected to be or what and how they should be teaching. In many institutions of learning, those involved in the review usually consider faculty performance with respect to four functional areas. These include the faculty duties with students, colleagues, discipline and with the university in general (UB Faculty Handbook, 2002; Queens College – CUNY, 2010).

The faculty's classroom teaching, course innovation and design and availability for and effectiveness in advising is being evaluated by colleagues and students. Performance with students is followed by the faculty's performance with the discipline he/she teaches. Acceptable criteria in this area of discipline is depicted by "evidence of scholarly, professional and artistic growth, such as publications, research, exhibits, and performances; possession of and progress toward advanced degrees: lectures and presentations at other institutions: grants and fellowships; participation in scholarly or professional societies and activities; professional participation in the affairs of surrounding communities and the region" (UB Handbook, 2002).

As mentioned earlier, the faculty is also evaluated based on how well he/she works with other colleagues in the university. This means he/she needs to participate in interdisciplinary activities, where feasible in the university and elsewhere; participate in academic committee work and planning and in team-based instructional experimentation, and the ability to stimulate thought, research or professional activity among colleagues (CUNY, 2018). As regards the university in general, the faculty in question is expected to participate in or encourage clubs, forums, lectures, exhibitions, and performances as well as show effective participation in faculty meetings. This faculty must demonstrate a responsible attitude toward the development and maintenance of library, technical and other resources (F&M College, 2018; Hilderbrand, Wilson, & Dienst, 1971; UB Faculty Handbook, 2002).

In addition to the four areas addressed above, the dean or his designate visits the faculty's class to assess the actual classroom activities and how the faculty handles them. The reports of these areas (teaching, scholarship and performance in the activities of the institution and society) are reviewed by the School/College Personnel Committee and the entire University Personnel Committee before a recommendation for reappointment, tenure, or promotion is made to the provost and/or the president of the university for a final decision. The key question now is how qualified are the individuals involved in this process of evaluation? What experience do they have or how knowledgeable are they in the specific items listed above on which the faculty would be evaluated? How valid are the instruments like the Faculty Evaluation Forms used? (University of Hawaii, 2009; CUNY, 2018).

METHOD

It was noted that most of the universities and colleges across the nation used

basically the same procedure of evaluating their faculty for reappointment, tenure and promotion. For this paper, five institutions were selected including Queens College of the City University of New York (CUNY), Franklin and Marshall College (F&M) in Pennsylvania, University of Hawaii (UH), Fairfield University (FU) and the University of Bridgeport (UB) in Connecticut. This could be considered as convenient sampling. The dossier expectations that were reviewed for these institutions were similar. The faculty member's probationary period is seven years and the documents submitted usually include curriculum vitae, record of courses taught, syllabi and course assignments, and then an account of departmental, institutional, and professional contributions. The instruments like the Faculty Evaluation Forms used by the institutions are different in structure with the questions worded differently too. However, the overall contents of the questions on the evaluation forms were virtually the same. The levels of evaluation—which include the school/college committee, the dean, the university personnel committee and the administration taking part in the evaluation—were practically the same with only a change in terminology. The evaluation forms were used essentially by the students to answer questions about the performance of the faculty in teaching and in carrying out other professional functions. Comments and discussions were then made about the validity or relevance of the contents of the evaluation tools and the appropriateness of the qualification of the personnel involved in the evaluation of these faculty for reappointment, tenure and promotion.

DISCUSSION

When the questions in the evaluation forms were reviewed, some of them seemed to be out of place and therefore not valid; consequently the results thereof would not be very reliable. That is, by being students, they may not be knowledgeable in some of the concepts expected to be evaluated about their teachers. Some of the questions are simply convoluted and ambiguous, lacking validity. For example, at Franklin and Marshall College, there is a question like “To what extent was class discussion managed in such a way as to permit most members of the class to participate regularly?” Respondents are expected to check off from “very poorly managed” to “very effectively” (0%, 13%, 19%, 69%). Another question is “To what extent did the instructor employ fair and consistent grading standards?” (0%, 6%, 28%, and 67% which is “Never – Always”). One might ask if these students are themselves professionals to know what “fair and consistent grading standards” are and the appropriate ways of managing class discussions. Undergraduate students may find these questions difficult to answer. There is a chance that graduate students in a Teacher Preparation Program might have covered those topics of classroom management and grading styles and could give meaningful answers.

A question such as “did you gain ability to analyze and solve problems?” raises the question as to who is at fault if the student did not gain the expected ability. How effective is the said student in self-evaluation of his ability to learn (“analyze and solve problems”)? A teacher might teach well but the said student fails to gain the ability because the student was not paying attention. How does this student then effectively evaluate the teacher's effectiveness? How do students determine the “quality of exercises and assignments” if they are not professionals themselves? Do they know the good characteristics of exercises and assignments? The students of course can say how easy or difficult questions are. They

may not be able to even distinguish between difficult questions and the highly recommended challenging questions. Similar questions from the University of Hawaii such as whether “laboratory assignments seem carefully chosen” and “lab equipment was, on most occasions, effectively set up” are also very subjective. Decisions taken about the faculty based on these students’ answers will be a travesty of evaluation.

Underhill (1987) writes elaborately about testing techniques of spoken language. Queens College - CUNY asks students if the instructor’s command of spoken English is good. This question seems to be problematic considering the fact that diversity is highly encouraged in our institutions of learning now. How good are students themselves versed with the language to be able to judge whether the teacher’s spoken language is good? There are some courses filled with international students for whom English is not just the second language but may be the third and even the fourth. Even some of the home students do not display a high command of the English language and are therefore unable to effectively evaluate the language of their teachers. Students’ language of instruction in most cases is therefore questionable. Are these the people to assess the quality of the instructor’s command of spoken English? The most likely thing assessed by these students in some teachers might be the accents of teachers whose language of birth is not English. Accents therefore might contribute greatly to the refusal of reappointment and tenure of many international faculty if evaluators consider students grading of the spoken language.

How is whether “this course is required for a student’s major or minor” by the University of Bridgeport useful in determining the effectiveness of an instructor? I wonder how those who analyze the students’ responses interpret the ‘yes’ or ‘no’ answers in relationship to faculty effectiveness. Students in another case were asked to say whether the textbook chosen by the instructor was good. What is questionable in this is that said students do not yet know the qualities of a good textbook. How could people who do not know the characteristics of a good textbook be expected to rate an instructor’s ability to choose a valid textbook? In most cases, the students evaluate the book in terms of cost rather than quality. A good quality textbook that is expensive is usually rated as a bad book which means the instructor does not know how to choose a textbook. The professor in question is rated low.

These examples and many more indicate that evaluation of teacher effectiveness is quite often not effective because the students doing the evaluation themselves have no experience or are not knowledgeable in the concepts in the evaluation form. How many students actually know who a good teacher is? Orlando (2013) has described the nine characteristics of a great teacher. For many, especially students, a good teacher is one who does not give many assignments, who does not care or overlooks when students come to class, who stops the class well ahead of time, and one who gives the grades the students want not necessarily what they merit. Teachers who are described as “strict”, time conscious, and whose assignments are valid and there is reliability in their grading results are the “bad guys” and are rated very low.

Students’ comments are usually anonymous and are rarely challenged even at the level of the institutions, let alone in courts (due process). There have been cases where faculty have been questioned, fired or simply not promoted based on one-sided accusations by students and personnel involved. These questionable complaints have pushed some faculty to be “teaching to the test” or grading to please the students because they want students to evaluate them positively. These are poorly trained faculty who do not know that

grading of students' work should be objective and not based on confounding variables.

Colleagues evaluate their peers (peer review), for example, in a class visit. This is usually to assess whether the lesson goals/objectives are clearly stated and met and to describe the classroom environment and how the professor links the relevancy of the class activities to the field at large. This evaluator is expected to describe how the instructor uses instructional materials and other supportive learning/teaching aids, including the use of technological/A-V materials. Other things assessed include the nature of communication between the professor and the students and the professor's rapport with the students. The faculty's classroom management style or the type of disciplinary techniques used are of vital importance and are evaluated too.

The researcher had an opportunity to have access to class visit reports over several years. All these attributes being assessed are based on pedagogy and on assumption that the evaluator is himself/herself knowledgeable in pedagogical techniques. It was noted that some of these evaluators were given this duty just because they are older in service in the department. They themselves are not trained professional pedagogues. They have never done any courses in curriculum development or educational pedagogy and therefore lack the basic concepts and principles of education, teaching and classroom management. All they know is what they have been doing for a long time en guise of teaching. Perhaps they were hired to teach in the university just because they had Ph.D degrees and not necessarily based on any formal pedagogical qualifications. They simply use their long time experience as the standard expected of the colleague being evaluated. These experiences at times are totally not in line with currently acceptable pedagogical principles of teaching. The colleague being evaluated might be current in teaching methods but will be given a bad grade by this evaluator with archaic methods of teaching based only on longevity in service. Who suffers at this point is not the evaluator but the one being evaluated.

It is known that there are some faculty who do not do a good job of teaching. However, the students in that particular course might be a problem themselves. They are lazy and not willing to learn and all they do is to complain about one thing or the other and disrupt the class and thereby giving the impression that this instructor does not know how to teach. The evaluator who cannot distinguish between students' conduct and teacher's ability to create a conducive teaching-learning environment often marks down the instructor. The debate here then is who is qualified to make class visits to properly evaluate peers going up for reappointment and tenure.

There are levels of review that continue with this evaluation for reappointment, promotion and tenure. These levels are composed of faculty and administrators. There is the School or College Personnel Committee (S/CPC) consisting of at least two tenured faculty members from the given college/school or department. Cases arise when there are not enough qualified members to be in the committee in a given department/school and faculty from other departments are invited to sit in committees of different colleges/schools. Decisions or votes from such visiting committee members may not often be very reliable, lacking personal knowledge of the colleagues being evaluated. The recommendation of a S/CPC is based on the documents provided by the candidate and on personal knowledge and interaction with the faculty being evaluated. This committee therefore provides a College/School perspective of the faculty to the dean and to the University Personnel Committee (UPC). The decisions of the S/CPC might be influenced, positively or negatively, by personal knowledge of the colleague. Often too, S/CPC

members might be ignorant or not adequately educated in the area of specialization of the one being evaluated. The concern is whether the members of the committee are pretty versed with the demands of the different disciplines of their colleagues? For example, are language faculty familiar with the expectations of the science disciplines or history subjects? There are cases where the one evaluated has had a misunderstanding with the one evaluating and such strained relationships often lead to poor and biased judgments of the performance of the faculty.

The next level is that of the dean or director of the school or college. The advantage here is that the search for a dean is usually rigorous and ends up with one who is knowledgeable and has acquired experience to be able to say who is an effective teacher. Unfortunately, the shortcomings common with the colleagues who do the class visits and those in the S/CPC might be same with the dean. The dean's review and recommendation based on the direct interaction with the faculty in this case could be marred and biased too. However, the dean's recommendation is guided too by the recommendation from the S/CPC.

After the dean is the University Personnel Committee whose members, mostly tenured faculty and most of them already at the rank of associate professor, are from various schools/colleges of the institution. Often too, UPC members might be ignorant or not adequately familiar with the area of specialization of someone from another department. That is, the other members of this committee may not be very versed with the type of journals available there, the publications and the work-load of a faculty from a different college/school. However, the representative of that department very often helps in explaining the department's perspective when such a question comes up.

As with all the the levels of review, the UPC closely follows the directives of the handbook of the institution. The UPC bases its decision essentially on the information in the candidate's dossier, the peer review report, and recommendations of the S/CPC and that of the dean. Members of the UPC rarely personally know candidates from other colleges/schools of the university. As expected of all other levels of review, the UPC watches out that its recommendation is not influenced by personal knowledge of a particular colleague.

The Provost constitutes the next level of review. He bases his decision on the reports by all levels of review (the C/SPC, Dean and the UPC) below him. He uses his personal academic and professional knowledge to confirm or reject the recommendations made by the other committees. Again as with the other evaluators, there might be cases where the provost is academically deficient in a certain discipline and misinterprets the performance of a faculty under review and takes a questionable decision. The titles of books/articles published might not be considered acceptable as fitting the profile of the discipline of the faculty evaluated, even if they are. The problem here might be resulting from lack of knowledge, for example, by a pure linguist not familiar with science education or social studies concepts and expectations. The faculty evaluated is then rated low.

It is here noted that in many institutions, it is the provost, usually the Vice President of Academic Affairs, that makes the final recommendation about the reappointment, tenure or promotion of the faculty member in question. Such institutions believe that teacher effectiveness is an academic issue and should be appropriately decided by the provost who is in charge of academic affairs. Although the president is usually responsible for administrative affairs of the institution, some institutions include the president of the

university to dictate the final verdict in this process. Basically, what is the academic qualification of these individuals? Some may have only the first degree and others have at best graduate diplomas and have not had the necessary experience.

The role played by the provost or the president might not raise any eyebrow if the said individual were adequately qualified – has adequate credentials in academia and experience themselves in the process leading to tenure or promotion. It has been remarked, for example, that some provosts and presidents have not had the opportunity or responsibility of rising through the ranks of assistant, associate or full professor and they find themselves in the position to determine who should or should not be reappointed, tenured or promoted. The chances of taking appropriate decisions would be very slim.

It was noted that in most of the institutions of learning, the position of a president is purely administrative and that is why most of them are elected and not hired based on academic qualifications. Some presidents tend to make their decisions in these cases based on the financial stance of the institutions since promotion, for example, comes along with increase in financial remunerations. In this case, an individual might be refused promotion to a new rank because the institution has no money to pay for that rank and meanwhile that individual is academically ripe for that rank. It is not uncommon to see an efficient president of an institution with a first degree carrying out his administrative functions very well. Questions usually arise when such individuals challenge the recommendations made by academicians, members of the school or university committees, about the promotion or tenure of a faculty member.

CONCLUSION

As stated earlier, the purpose of this paper has been to examine closely the instruments used and the qualification or eligibility of those involved in the process of evaluation of faculty that merit reappointment, tenure and promotion. How do we know if the faculty is achieving the institution's specific teaching goals and other contractual duties and thus deserve to be retained? Selecting eligible evaluators is the challenge. Faculty evaluations done appropriately by the right students, colleagues and administration provide excellent feedback about student satisfaction of faculty teaching style, but they don't provide the important detail of how much the faculty is teaching.

If the qualification of those doing the evaluation and the instruments used are questionable, wrong data is obtained and, of course, wrong decisions are taken concerning the eligibility of those to be either reappointed, tenured or promoted. Therefore, the more accurately we judge our faculty using valid instruments, the more effective we will be in weeding out poorly performing teachers while maintaining and promoting qualified faculty and directing their teaching. This happens when sound teaching decisions are based on information from accurate, relevant and comprehensive evaluation. The contention here is that the effectiveness of the teacher and teaching depends, to a large extent, on the quality of the evaluation information and tools on which the decisions are based.

RECOMMENDATIONS

Changing the way faculty are assessed can dramatically improve teaching effectiveness. The evaluations done by the right people concerned can have a tremendous

influence on the lives of the faculty being evaluated and the institutions as a whole. Evaluation therefore should not be lightly or casually made by any type of person and using poorly formulated instruments. Carefully collected data by those who understand the intricacies of evaluation would definitely help retain professors who understand the learners, plan better learning objectives and good teaching strategies for them. They will further determine the extent to which the teaching objectives are being achieved and also plan to contribute effectively in the growth of the discipline, profession and the institution. It is then a good idea that such decisions are being taken by a number of capable people who are scrupulously scrutinized to ensure fitness-for-purpose and not by one person. The concept of shared governance should be enforced such that the chances that an individual should take a final decision about a faculty should not exist. The questions in the Faculty Evaluation Forms should be revised to reduce ambiguity and increase validity. These questions should be sensitive to cultural diversity and the level of the students. Qualified evaluators will not include confounding variables in their evaluation. To facilitate the work of the evaluators, the university handbooks ought to be revised to be much more specific in describing requirements for reappointment, tenure and promotion. Vagueness surely leads to poor interpretation of stated criteria thereby resulting in subjectivity which leads to lack of validity and reliability in the evaluation. The qualification of those doing the evaluation at each level must be reviewed so that valid decisions should be made such that reappointment, promotion and tenure should truly reflect academic prowess of a faculty.

REFERENCES

- Bambrick-Santoyo, P. (2010). *Driven by Data: A practical Guide to Improve Instruction*. San Francisco: Jossey-Bass.
- CUNY. (2018). *Procedures for Tenure and Promotion*. Retrieved June 8, 2018 from <http://www.qc.cuny.edu/about/administration/provost/facultystaff/pages/procedures-for-tenure-and-promotion.aspx>
- Dougherty, E. (2012). *Assignments Matter. Making the Connections that help Students Meet Standards*. Alexandria, VA: ASCD
- Feldman, J. M., & Lynch, J. G. (1988). Self-Generated Validity and other Effects of Measurement on Belief, Attitude, Intention and Behavior. *Journal of Applied Psychology*, 73(3), 421-435.
- Franklin and Marshall College. (n.d.). *Tenure Information*. Retrieved June 8, 2018 from <https://www.fandm.edu/provost/tenure>
- Hildebrand, M., Wilson, R. C., & Dienst, E. R. (1971). *Evaluating University Teaching*. Berkeley: Center for Research and Development in Higher Education.
- Hoffman, B. F. (1962). Electrophysiology of the conducting system of the heart. *Transactions of the New York Academy of Sciences*, 24(8), 886-890.
- Layne, L. (2012). Defining Effective Teaching. *Journal on Excellence in College Teaching*, 23(1), 43-68
- Ngoh, N. N. (2001). *Understanding examinations: A practical guide to measurements and evaluation for teachers*. Yaounde, Cameroon: Presse Universitaire de Yaounde.
- Orlando, M. (2013). *Nine Characteristics of a Great Teacher*. Retrieved from <http://www.facultyfocus.com/articles/philosophy-of-teaching/nine-characteristics->

- of-a-great-teacher/
Queens College – CUNY. (2010). *Queens College Course Information System*. Retrieved October 6, 2010 from <http://courses.qccuny.edu/?list=Sections&id=5043&display=survey>
- Reeves, B. D. (2004). *101 more Questions & Answers about Standards, Assessment, and Accountability*. Denver Colorado: Advanced Learning Press.
- Stanford Teaching Commons. (n.d.). *Assessing Student Learning*. Retrieved from <https://teachingcommons.stanford.edu/resources/teaching/evaluating-students/assessing-student-learning>
- Underhill, N. (1987). *Testing Spoken Language: A Handbook of oral testing techniques*. Cambridge: Cambridge University Press.
- Weimer, M. (2013). *Learner-Centered Teaching: Five Key Changes to Practice*. San Francisco: Jossey-Bass.