

SET – the student evaluation of teaching staff - in secondary education

Lance King – 02/04/07

“Teachers are evaluated by one means or another in virtually every school district. And in most of those districts, teachers and administrators agree that the activity is ritualistic and largely a waste of time” (McLaughlin (1990) - as cited in Wilkerson, Manatt, Rogers & Maughan, 2000, p. 180)

How do any of us know for sure that we are doing a good quality job? Self appraisal will not suffice – it cannot be objective. Only by the evaluation of our performance by a third person (or persons) can we ever hope to receive objective feedback as to the quality of our output. Evaluation is the (hopefully) impartial assessment of performance in-situ, at the “coal face”, with the clients. Good quality evaluation is probably best performed by someone with:

- 1) an understanding of the parameters to be evaluated
- 2) the competence to judge between “superior”, “adequate” and sub-standard” performance in relationship to the parameters, and
- 3) the ability to mitigate any bias in their evaluation

and of course the time available to do the job well. In most organizations this evaluative role is generally handled either by a superior in the same organization, by a peer who is familiar with delivering the same or similar service or by the recipient of the service - the client.

The emphasis on the client as evaluator can be witnessed daily in hotels, in the insurance industry, in car sales – asking for client feedback on all aspects of service provided – but nowhere more so than in tertiary education. Over the last 30 years across the world, the tertiary education sector has mostly devolved the role of teacher evaluator (sometimes exclusively) into the hands of students. As of 1988, 80% of all liberal colleges in the US used systematic student ratings as all or part of the means for evaluating teachers (Langbein, 1994). Across tertiary education in the UK, data from SET (*the Student Evaluation of Teaching*) is considered as important evaluative information to judge academic staff on and as a guide for potential changes in course material and method of delivery (Shevlin, 2000). And in Australian universities Ramsden (1991) reports that

It is generally considered that student evaluations are more useful, accurate and valid than other measures of teaching performance and have the added benefit of being a direct measure of consumer satisfaction (p. 131).

The use of SET to support curriculum and personnel decisions in higher education in the USA increased from:

- 28% in 1973 to
- 68% in 1984 and then to
- 86% in 1993

Seldin, (1993).

until by 1994 45% of salary adjustment funds were allocated by teaching effectiveness criteria, as judged by students (Stratton, Myers and King, 1994).

In secondary education the picture is very different though. Traditionally teacher evaluation has been a process carried out by administration figures as part of an employee appraisal system the aim of which was to pick up any gross defects in performance and determine suitability for movement up through the salary grades. Good and Mulryan, (as cited in Wilkerson et al., 2000), in commenting on the history of teacher evaluation, stated,

Rating systems were developed primarily for reasons external to teachers, that is, to demonstrate to the public that students were receiving appropriate instruction or that teachers were competent, rather than to provide teachers with information that they might use to improve instruction (p. 180).

According to Wilkerson et al., the vast majority of personnel evaluation in secondary education is conducted for summative purposes – hiring, determining promotions, dismissal etc. And it may not still be true today but certainly in the early 90's these evaluations were typically conducted by an administrator untrained in personnel evaluation and the "whole evaluation often lasts 40 – 60 minutes – per teacher, per school year" (Haefele, 1993, p. 23).

In the US, "in 46 of 50 states, only a principal or other single individual assesses the performance of teachers" (Wilkerson et.al, 2000, p. 181). Performance evaluation for summative purposes of such a small sample of actual teaching practice yields little useful information for teachers on how to improve their teaching process and yet formative evaluation processes appears to be rare in secondary education (or rarely researched and published). Administrators themselves are not unaware of the limitations of the existing model – as David Brazer, Vice-Principal at Los Altos High School in California puts it – "the focus of evaluation efforts must be shifted from meeting the demands of the district office to fulfilling the needs of classroom teachers" (Brazer, 1991, p. 82). Classroom teachers need a system of evaluation that encourages them to try new methods of delivery, new strategies of teaching and gives them good quality feedback on the effects of such changes without threatening their job security or their pay packets but it seems (from one study anyway) that it is only when the regime of formal summative evaluation is removed that the teachers feel comfortable taking the risks of changing teaching practice (Grube, et al. 1988, pp 17-20).

Formative teacher evaluation is characterized by Barber (1990) as

...a set of procedures designed to assist teachers in improving their own teaching. Formative evaluation is a helping, caring process that provides data to teachers for making decisions about how they

can best improve their own teaching techniques, styles or strategies (cited in Mertler, 1997, p. 8).

Most importantly formative evaluation is not judgmental or punitive but rather is supportive of the professional development of the teacher and can be an effective mechanism for self improvement.

Formative evaluation can take place:

1) through feedback from peers as outlined by Munson (1998) through

Peer coaching where written and verbal feedback is provided to the observed teacher by the observer teacher who acts as coach, reinforcing correct instructional skills and identifying incorrect behaviours according to the standards of a prescribed model of instruction. Peer coaching requires teachers to judge one another's performance which ultimately can cause interpersonal problems.

or through

Peer observation, which eliminates the element of judgment.

The role of the observer teacher is to provide specific feedback on some aspect of teaching that has been chosen in advance by the observed teacher. This feedback is in the form of observed objective data that has been recorded by the observer and which the observed teacher critically reflects upon during the post-observation conference (Munson, 1998, p. 1).

or 2) from observations by supervisors or

3) from 'self-analysis'

but data from these last two sources can easily be biased towards maintaining the status quo as demonstrated by "a study of the effectiveness of middle and high school vocational education teachers in North Carolina found that most of the 243 responding teachers perceived their performance to be from "above standard" to "well above standard" and most of their principals perceived teacher performance to be "above standard"." (Jewell, 1990, p. 3).

Gaining evaluative feedback from students can demonstrate whether the teacher is meeting the needs of the students in the manner most appropriate or most effective for them but we need to ask in the case of secondary schools, whether secondary students are good judges of teaching effectiveness?

In a study involving 51 elementary, middle, junior high and high school principals who were asked to evaluate seven teachers, those that had access to student evaluation and summative assessment data on the teachers produced virtually the same ratings as those who only had access to the summative assessment data. This indicates that the student evaluation data supported rather than contradicted the summative data (Savage, 1986).

In Illinois in 1991 Linda Stroh compared university professors' and high school classroom teachers' evaluations of the performance of student teachers, with the student teachers' own self evaluations and with their high school students' evaluations and found that the high school student evaluations were "very comparable to the more "expert" evaluations" (Stroh, 1991, p. 81).

And the evaluations by 1,128 Anglo, Black, and Chicano students of 53 sixth grade teachers were found to "correlate very well with observers' teacher ratings." In this case the "process of teacher evaluation by students was also found to positively affect student outcomes" (Peck, 1978, abstract).

It would seem then that students as the recipients of teaching are well placed to evaluate the quality of that teaching provided for them and may even be able to offer insights into what constitutes effective teaching. In 1989 in Barbados 160 elementary and 60 high school students as well as 60 elementary and 30 high school teachers were asked to identify the characteristics of effective teachers. Both the student and the teachers groups identified the following common "most important" characteristics:

- 1) the giving of examples by the teacher
- 2) properly planned lessons
- 3) knowledge of subject matter
- 4) knowledge of student learning (Richardson et al. 1989, abstract).

These characteristics correlate well with evaluative data obtained in 2005 from 271 high school students in North Carolina who considered that good teachers will

- 1) keep the class moving in a positive direction, be patient, caring and understanding
- 2) be knowledgeable about the content and about the world
- 3) manage the learning environment to enhance the full participation of all students
- 4) explain things well and give good examples (Whitney, Leonard, M; Leonard, W; Camelio, M; & Camelio, V; 2005, p. 29).

But do teachers that students give the highest evaluative ratings to actually produce the goods? Do they help students learn and achieve academically or are they just those teachers who are the most entertaining?

A key study which seems to answer these questions was published in June 2000 and outlines a comparison of evaluations of teachers by principals, by students and self evaluations by the teachers themselves with the performance of four groups of students (K-2, Grade 3-5, 6-8 & 9-12) on criterion referenced reading, language arts and mathematics tests. The results were High, positive correlations between student feedback of teacher performance and student achievement in all three core subject areas. There was a highly significant,

positive relationship between the student feedback results and student achievement on the criterion referenced posttests in reading, language arts, and mathematics (Wilkerson, et al., 2000).

And interestingly enough in predicting student achievement on these same tests “the self-ratings, principal ratings, and the principal summative evaluations failed to meet the pre-determined rejection level for significance” (Wilkerson et al., 2000).

So it would seem from the evidence that

- 1) students are often good judges of what constitutes good teaching
- 2) student evaluation of teaching correlates well with other expert evaluations, and
- 3) students are sometimes better judges of the teaching necessary for high academic achievement than are principals or the teachers themselves.

So why is SET such a rare thing in secondary schools?

It would appear that one reason is resistance from teachers themselves.

Two studies from Nigeria illustrate the problem.

The first is from Joshua and Joshua who surveyed 480 secondary teachers from 20 schools and found “a significantly negative attitude to student evaluation of the teacher, irrespective of the use(s) to which the results of such evaluation will be put” (Joshua & Joshua, 2004, p. 12) And in subsequent discussion of the results they revealed this very attitude by stating, without any corroborating evidence that

Students’ rating, given the ages, social backgrounds and the orientations of Nigerian students, can hardly be valid, reliable, interpretable and generalisable. The tendencies are high that students will rate many other things of the teachers in addition to, or opposed to, quality of instruction and teaching performance (Joshua and Joshua, 2004, p. 13).

The second study (also from 2004) is of 2,310 students’ in 60 secondary school in 12 states (out of the 36 states of the Federal Republic of Nigeria) where the researchers were making an honest attempt to find out about the quality of secondary teaching in Nigeria from the perspective of the students themselves. The researchers Akpotu and Oghuvbu (2004) revealed that the common opinion in the public of Nigerian secondary teachers was that they were “deprived, dejected, forsaken, ...poorly trained, poorly paid and poorly motivated” (p. 52) and they were seeking to find out if this opinion was borne out in the experience of the students. They were also seeking to test out the effectiveness of appraising teachers by using student evaluations. The findings of the study were claimed to be that (in general):

Nigerian secondary school teachers are effective in class attendance, competent in content and pedagogy, have a positive relationship with students and disciplinary qualities. The results

indicate that students perceive their teachers as efficient in their job performance (Akpotu & Oghuvbu, 2004, p. 52)

And also they found that the method of student evaluation was highly effective and they recommended that:

“The Nigerian educational administrators, planners and policy makers should as a matter of urgency incorporate students’ assessment in their teacher evaluation policy” (Akpotu & Oghuvbu, 2004, p. 53).

They also recommended that such evaluations

“.....should form part of the overall annual evaluation of the teachers and should be used as the basis for regular in-service training programmes for teachers” (Akpotu & Oghuvbu, 2004, p. 53).

So from within the same educational structure we get a complete endorsement of SET from outside the system and complete rejection from inside, by the teachers themselves, which while both being gross generalizations and appear to be very political in nature certainly highlight the difference of opinions about SET in secondary circles.

Why is this?

Having worked in over 150 secondary schools in New Zealand, the UK and Thailand in the last 15 years the author has much anecdotal evidence which suggests that many teachers in secondary schools are afraid of the implications of SET. On the one hand many feel that they maintain only a tenuous authority over students at the best of times and to let those very students have any influence at all over the performance of their job, any influence over the processes of their teaching, would destroy that authority. And also on the other hand they are simply afraid that if they opened themselves up with SET that they would receive negative feedback on their teaching.

But are these opinions borne out by the facts?

Individual teachers have written papers outlining their own personal use of SET – for example Marvin Hoffman (1992) reported he uses student evaluation very effectively

...when I need help with a problem that has me stymied; when the activity we’re engaged in has no visible product or outcome and I need an indirect measure of its usefulness; when the end of a project calls for closure; and when it’s time to take a sweeping assessment at semester-end (p. 79).

And he suggests more broadly that “Properly done, evaluations can help middle and high school teachers improve the quality of teaching and learning (Hoffman, 1992, p. 79)

In New Zealand, "*Assessment for Learning* has been an integral part of the national assessment strategy since 1999" (Olsen, 2005, p. 9). Assessment for learning rests on 5 key principles:

- 1) the provision of effective feedback to students
- 2) the active involvement of students in their own learning
- 3) adjusting teaching on account of the results of assessment
- 4) recognition of the profound influence assessment has on the motivation and self esteem of pupils, both of which are critical influences on learning
- 5) the need for pupils to be able to assess themselves and understand how to improve (Chamberlain, 2001, p1)

It seems to me that if we have in place strategies to achieve the above with students it is incumbent upon us to also apply the same criteria to teachers and enable every teacher to use an effective system of self-assessment or evaluation to improve their own performance along exactly the same lines.

In order to be congruent with the assessment for learning programme as applied to students, *Evaluation for Learning* should operate on the same principles, which could be:

- 1) the provision of effective feedback to teachers
- 2) the active involvement of teachers in their own learning
- 3) adjusting teaching on account of the results of evaluation
- 4) recognition of the profound influence evaluation has on the motivation and self esteem of teachers, both of which are critical influences on learning
- 5) the need for teachers to be able to assess themselves and understand how to improve

From the evidence it would seem that one way to achieve this would be through the use of a well designed system of SET.

The last paper I wish to reference was written in 1997 by Craig A Mertler and it reports on the use of a student evaluation instrument called SE3T, administered to volunteer students across 7 high schools in two school districts in North Florida on two occasions during the 1995-1996 school year. The experimental design appears very thorough and carefully constructed - all data were captured anonymously on questionnaire forms, were collected by a person other than the teacher being evaluated, the data were collated, analysed and reported back to the teacher within one week. For each item on the questionnaire the teacher only received the mean response and the range of responses. Any written responses to the open ended questions were given to the teacher verbatim but were typed up so no handwriting identification was possible. Each teacher only received their own data with no reference to any other teacher and no other person was privy to their data. After the second administration of the SE3T the teachers were all surveyed as to their opinions of the process

as to its usefulness, efficiency, and whether the feedback they received had impacted at all on their teaching practice.

The conclusions of the paper were that:

...it can be concluded that the process of collecting student feedback regarding teachers' classroom performance is a very useful one to teachers who voluntarily participate in the process. In contrast to the deficit model of teacher evaluation (ie. infrequent classroom visits and evaluations by administrators), this process is seen as something done for teachers rather than against them. The participating teachers attitudes, desires, and actions in this study toward improving their teaching behaviours serve as strong support for this conclusion (Mertler, 1997, p.17).

And they also found that once teachers were familiar with the evaluation process they were able to make it self-sustaining – “Analyses of the data revealed that overall reaction of the teachers to this process of collecting student feedback was highly positive and that it was feasible for teachers to conduct this process on their own” (Mertler, 1997, p.1).

One is left wondering why more secondary teachers don't use SET???

REFERENCES

Akpotu, N. E., & Oghuvbu, E. P. (2004). Performance appraisal of the Nigerian secondary school teachers: the student perspectives. *ISEA*, 32(3), 44-57.

Brazer, S. D. (1991). The search for meaning in teacher evaluation. *Educational Leadership*, 48(6), p. 82.

Chamberlain, M. (2001). Developing assessment for learning. Retrieved May 16, 2006, from www.nzpf.ac.nz/resources/magazine/2001/nov/new_page_12.htm

Grube, G., Cram, H. G. & Melchior, T. M. (1988). *Educational Leadership*, 46(2), 17-21.

Haefele, D. L. (1993). Evaluating teachers: a call for change. *Journal of Personnel Evaluation in Education*, 7(1), 21-31

Hoffman, M. (1992). “How am I doing’?: the importance of evaluation in the classroom. *English Journal (High School Edition)*, 81(4), 79-82.

Jewell, L. R. (1990). Evaluating Teaching Effectiveness of secondary vocational educators. *Journal of Vocational and Technical Education*, 6(2), 3-17.

Joshua, M. T., & Joshua, A. M. (2004). Attitude of Nigerian secondary school teachers to student evaluation of teachers. *Teacher Development*, 8(1), 67-80.

Langbein, L. I. (1994). The validity of student evaluations of teaching. *Political Science & Politics*, 27(3), 545-558.

Mertler, C. A. (1997). *Students as stakeholders in Teacher evaluation; teacher perceptions of a formative feedback model. Paper presented at the annual meeting of the Mid-Western Educational Research Association.* Chicago, IL.

Munson, B. R. (1998). Peers observing peers: the better way to observe teachers. *Contemporary Education*, 69(2), 108-111.

Olsen, L. (2005). Classroom assessments stir growing global interest. *Education Week*, 25(6), 8-13.

Peck, R. F., (1978). Student evaluation of teaching: a multivariate validation study. *Research and Development Center for Teacher Education publication*, Texas Univ., Austin.

Ramsden, P. (1992). *Learning to teach in higher education.* London: Routledge.

Richardson, A. G., Thomas, A. A. (1989). *Characteristics of the effective teacher as perceived by pupils and teachers: a Caribbean case study. Paper presented at the Annual Meeting of the American Educational Research Association.* San Francisco, CA.

Savage, T. V. & McCord, M. (1986). *The use of student evaluation in the assessment of teacher competence. Paper presented at the Annual Meeting of the American Educational Research Association.* San Francisco, CA.

Seldin, P. (1993). The use and abuse of student ratings of professors. *Chronicle of Higher Education*, June 12, 40.

Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The validity of student evaluation of teaching in higher education: love me, love my lectures? *Assessment and Evaluation in Higher Education*, 25(4), 397-405.

Stratton, R. W., Myers, S. C., & King, R. H. (1994). Staff behavior, grades, and student evaluations. *Journal of Economic Education*, 25 (1), 5-15.

Stroh, L. (1991). High school student evaluation of student teachers: how do they compare with professionals? *Illinois School Research and Development*, 27(2), 81-92.

Wilkerson, D. J., Manatt, R. P., Rogers, M. A., & Maughan, R. (2000). Validation of student, principal, and self-ratings in 360° feedback for teacher evaluation. *Journal of Personnel Evaluation in Education*, 14(2), 179-192.

Whitney, J., Leonard, M., Leonard, W., Camelio, M., & Camelio, V. (2005/2006). Seek balance, connect with others, and reach all students: high school students describe a moral imperative for teachers. *The High School Journal*, 89(2), 29-40.