

Case article Quality enhancement in Higher and Further Education; grooming the growers in the groves of Academe.

Date August 2010

Published By Dr. Sandy Ochojna

Dr Sandy Ochojna is an independent survey research advisor. Between 1986 and 2008 he was the Manchester-based director of several well-known international market research companies; for the ten years prior to that he was Passenger Manager at Strathclyde Passenger Transport Executive in Glasgow. From time to time he feels compelled to set down his thoughts on topical survey issues.

For some time now I have been an adviser to the Electric Paper Company; its EvaSys software package offers colleges and universities an effective means of collecting paper and on-line students' assessments of their teachers. I have talked at various EvaSys seminars with the people who manage such assessment procedures and it is clear to me that there are some instances where they feel that the process can generate misunderstanding between themselves and academic staff....or, to put in another, and perhaps more understandable, way, between those carrying out the evaluations and those who are being evaluated.

Clearly, deep down, nobody really likes to have their performance assessed, either at work or at home or at play. However, I am a bit surprised to come across this in the field of education where the concepts of evidence-based deduction, testing, examining, comparing, marking and correcting are all fundamental parts of the teaching process. Surely Kierkegaard was right;

"Instruction begins when you, the teacher, learn from the learner; put yourself in his place so that you may understand....what he learns and the way he understands it."

So why the problem?

A lemon for the institution?

The challenge for all performance evaluation is to have a system which is robust enough to withstand the scrutiny it is subjected to by those in the frame who have been 'found' wanting; the problem with any process whose results can be used, rightly or more often than not wrongly, to rank 'performance' is that for every high flyer there is one at the bottom....and of course that person's situation is always atypical.....and indeed that may so be. And to meet the challenge any institution must set out very clearly what it is trying to achieve with its evaluation process.....and this is where the big issue lies, not with academics or non-academics, but with the institution itself. The institution has to have a clear understanding of exactly why such exercises are being commissioned, what they can be expected to deliver, and what good they can bring to the institution in the long run. Many organisations are not that clear on this.

So, what should institutions be seeking from student perceptions of their teachers? Well, the first thing to acknowledge is that any such appraisal does not collect facts, it merely is noting the perceptions of students, and there is no way of stating categorically that such perceptions are indeed 'correct' or even relevant. But such views are accurate in respect of how students, for whatever reasons, think a teacher performs in relating to them

That is, student assessment is valid in measuring how well the teacher communicates, explains, helps them, supports them, and deals with the lecturing environment. What this approach cannot do meaningfully is assess the quality of the course itself or the quality of the information being imparted....this must rely upon peer review and good old externally adjudicated examination results. All that a performance regime can seek to assess is the quality of the teaching, not the structure or content of any course, nor the quality of student outcomes and competencies.

It's all about raising teaching standards to the highest level attainable within the institution. It's not about who is seen to be best and who is worst....but of course letting the worst see the best in action can only help raise standards overall.

The lesson? It must be made very clear that student assessment of teachers is about **how** a course is presented, nothing more. But this can be a very useful and productive exercise.

An apple for the teacher?

Clearly, such an assessment of teacher performance depends upon many factors, but the key ones are;

- the actual proficiency of the teacher
- the type of students (their proficiency?) in the class
- the nature of the course subject.

It must be accepted that some students will be more motivated than others, that some subjects will be seen to be 'easier' than others, and some students will never be pleased, others will be determinedly indifferent, that some courses are optional and others might be mandatory, that some sessions might involve practical work and others might be entirely 'book-based'. And of course, students' assessments will be coloured by how well they are coping with the course work; some will find it too demanding, some will find it satisfyingly challenging, and others will find it to be just too easy. All in all then, can we really compare a history teacher with a metalwork instructor; can we learn anything by looking at the perceived lecturing proficiency of a Latin scholar and an astronomer? Of course we can, as long as we remember that we are not interested in content, only delivery.

And of course, we must remember to take into account the nature of the students making these assessments. Yes, just because teacher A scores very highly in their student assessment doesn't mean automatically that they are a very good teacher....it might just mean teacher A has a class of disinterested students who are happy to be left to do as little as possible.

The upshot of this is that for an assessment programme to be fair to teachers, and so be accepted by them, it must ensure that the system is capable of discriminating not just amongst teachers, but amongst students, and perhaps courses, as well.

Planting and pruning; survey phase 1 - courses

The whole exercise is about teacher delivery, and so we are looking at such aspects as clarity of presentation, quality of handouts, perceived degree of teacher preparedness, willingness to field questions, ability to get everyone involved, and so on. And to place such a student

assessment in context we must also ask about the course itself - how 'easy' it is, how interesting, how well structured and so on. And finally we must consider the respondent by looking at such aspects as student demographics, degree of choice in taking the course, and overall satisfaction with it.

In an ideal world then we would simply ask all students about every class or module they attend.....and with current technology this is not a problem, though it certainly would place an enormous burden on student respondents. But we can reduce the burden significantly if we split the survey programme into two distinct phases, one looking at the various elements of **the course**, and the other concentrating on individual **teachers**.

Surveys amongst all students into their satisfaction with their courses and how they are run in general could be undertaken usefully say mid-way through term one (so that respondents have some meaningful experience of the course) and at the end of the year (when their expectations will be more realistic and their awareness of what other courses have to offer in terms of delivery will be broader). With this vast dataset we can undertake some fairly simple but very powerful statistical analysis based on simple satisfaction ratings which can...

...by using Key Driver Analysis, identify which of the delivery aspects measured have the most impact (positive or negative) on overall student satisfaction

...by using CHAID, (**CHI**-squared **A**utomatic **I**nteraction **D**etector) the data itself can identify which of all the student/course combinations are the most discriminatory in segmenting the student body into satisfied and dissatisfied students. Being a purely statistical and mechanistic assessment of the data means that it is a far more powerful approach than having management try and identify problem areas or put forward performance hypotheses based on their 'experience', 'knowledge of the college', or intuition. The CHAID process encourages the institution to ask such questions as, for example, why are the most satisfied students more likely to be women, vocational, under 20 while the least satisfied are male 25-40 year olds in a science course? Is it the course, the type of student, or the teaching? If younger students are more likely to be satisfied than older ones, then teachers with older students in their classes could be expected to have scores lower than those with younger students, regardless of their teaching competencies.

However, the power of such internal benchmarking depends to a great degree on the **whole institution using the same set of questions** and the same fieldwork method(s), or at least adopting a series of key questions, so that meaningful comparison can be made across all parts of the institution. If this is not done, and faculties or schools go off on their own tack, then a great deal of student input, in the form of survey participation, is being abused and wasted. Furthermore, if this is not done then the institution loses the ability to track its overall perceived performance year on year....and that's another big loss.

Picking; survey phase 2 – teachers

The phase 1 course review means that when students are asked about teacher performance they are asked only what is relevant...to them, and to you... while staff, academic and non-academic, can begin to appreciate that overall teacher performance may, in many cases, be driven as much by who or what is being taught as it is by how it is being taught. As a consequence, any formal assessment based on the results of student surveys can be modified to take account of such factors.

The secret of a successful, on-going continuous student survey programme lies in sample selection, that is the selection of which students and which courses/lessons to sample. And here we have two options. We could concentrate survey activity on those courses which phase 1 has highlighted to be performing below par; this will limit student involvement in this phase and also yield the highest chance of identifying failings in delivery. Or we can continue to give all students a voice in the assessment programme but limit their input via implementing an appropriate sampling approach.

The unit of sample depends on what has to be measured, based perhaps on the initial comprehensive survey as suggested above. We could simply assess the **performance of the teacher** by taking a sample of all his or her students; or we could try to be a bit more sophisticated and develop a **sample which reflects each teacher's various courses as well**. The latter approach means we can 'control' for different types of students in the different classes as well as accepting that a teacher may perform better in one discipline area than in another.

At its simplest we should consider developing a sample frame which lists every opportunity for teacher-student interface and comprises every

sampling unit of teacher-course combination: for every teacher there is a list of his/her courses on one side (the full spectrum of teaching delivery), and the list of participating students on the other (the full spectrum of eligible assessors). Ideally, each teacher-course sampling unit should achieve c100 student assessments over its period of delivery, which could be a term, a short course of a fortnight, or a year; in practice there are probably few opportunities for such an intense review, and many courses will not have that many students, in which case the simple thing is just to try and get as many as possible to take part.

While this does make it simpler to carry out the survey in class, and so benefit from the higher response rate that the herd instinct can generate, it does mean that the questionnaire must be very clear in stating that the assessment relates to several weeks' of teaching, otherwise many respondents might refer solely to the latest lecture in which the teacher may well have invested more effort than normal given that he or she was aware that forms were to be distributed on that day. In contrast, an e-survey can be very selective in regards to which student is asked about which teacher-course combination, the memories of the most recent lecture are not 'ringing in respondents' ears', and teachers do not have the 'luxury' of pitching up a prize-winning performance for the 'day of judgment'.

The complexity or otherwise of this regime depends on the size of the institution and the number of courses it offers. It also depends upon whether the institution wants, or indeed needs, to assess performance at the course or particular lecture level. But if a sampling strategy is adopted as set out above then the frequency with which any student is asked to complete a survey is minimised, and with the electronic sampling databases and returns monitors built into systems such as EvaSys it is easy to control which students are asked to assess which teachers.

Cultivating the Tree of Knowledge

Student surveys of teacher classroom delivery is a perfectly valid way of helping institutions assess the performance of teachers. If used sensibly and effectively it helps institutions understand the dynamics of their teaching processes, appreciate the factors which lie outwith the command of the teacher, hence enables institutions to develop realistic measures of teacher performance. This one process can indeed compare the performance of the Latin master and the star-gazer.

Identifying those teachers who teach well and those who could be expected to teach better offers the institution the opportunity for internal benchmarking and the development of 'best practice'. If managed in conjunction with all parties the results can only lead to a better teaching experience for the student.

"The authority of those who teach is often an obstacle to those who want to learn."

Cicero may have been correct two thousand years ago, but today perhaps technology like the EvaSys package offers the means to prove him wrong.

Sandy Ochojna

www.thesurveydoctor.co.uk

August 2010

thesurveydoctor.co.uk