Using observation of teaching to improve quality

Author: McMahon, Barrett and O’Neill
Email: tim.mcmahon@ucd.ie
Date: 2007
Using Observation of Teaching to Improve Quality – finding your way through the muddle of competing conceptions, confusion of practice and mutually exclusive intentions.

Tim McMahon, Terry Barrett & Geraldine O’Neill
UCD, Dublin, Ireland.

Abstract:

This paper begins by reviewing some of the different models of third-party observation of university teaching that can be found in the literature. Having analysed these, it argues that - if ‘peer’ is taken to indicate equality of status - only one is genuinely a model of peer-observation. It proposes an alternative categorisation of third-party observations of teaching dependant on who controls the information generated by the process. A preferred six-dimensional model based on control by the person being observed of the data-flow, and other procedural aspects, is presented and explored. Evaluative comments, by university teachers who have undertaken the process, are presented to illustrate the benefits of adopting this model.

Introduction

Teaching, by definition, is an activity that is observed. It is only comparatively recently, however, that it has become a widespread practice for teachers in Higher Education to be observed by people other than those being taught - what we might call ‘third-party observation’ (Fullerton, 2003, p. 226). [This term is preferred to the more commonly used ‘peer-observation’ on the grounds that the latter is often a misnomer used to describe, among other things, observations of new staff by line managers, observations by external quality audit teams and observations by educational developers acting as gatekeepers to professional qualifications. Indeed, Shortland (2004, p. 222) quoting the UK lecturer’s
union, NATFHE, actually defines ‘peer-observation’ as a process of ‘third-party observation’ without stipulating any limitations on who the third-party might be.]

Third-party observations of university teaching, at least in the English-speaking world, first became common in the United States in the 1960s, where student-generated reviews of teaching staff eventually led to professional development programmes of which peer-observation was an integral part (D’Andrea, 2002). It took two to three decades for this practice to spread to Australia and the United Kingdom, becoming widespread in the latter only in the 1990s (Shortland, 2004, p. 220), and only in the first decade of the twenty-first century did it begin to appear in any significant quantity in the Republic of Ireland. Prior to the appearance of such observations, there seems, in all jurisdictions, to have been a commonly shared assumption that in order to judge the quality of teaching, it was enough to consider the subsequent performance of the students. Good teaching would lead to good results, bad results were a sign of poor teaching. Such a simplistic analysis, however, ignores the multiplicity of variables present in almost any programme of study in third level institutions. Students are, for example, rarely taught solely by one individual; thus, even if there was a direct relationship between quality of teaching and student performance, it would be difficult, if not impossible, to identify differences in the respective quality of the performances of individual teachers. Further, there will be those students who succeed despite bad teaching, and those who fail to perform to their full potential despite consistently excellent teaching. In the complexity of the interaction between the attributes and capabilities of individual learners and the totality of the influences of their social contexts, the quality of teaching is one factor out of many. It is, however, one that - if we are able to identify what constitutes ‘good’ teaching - is relatively easy to influence; at least, compared to the task of compensating for socio-economic inequalities.

For this reason, more and more of the many and varied quality improvement strategies attempted in Higher Education have come to include some form of third-party observation of teaching, with a view to using the results to improve the standard of teaching and the learning experience of students. More often than not, these are labelled ‘peer-observations’ even where there is a manifest lack of equality between observer and observed.

Methodology

This paper adopts a situated, practical, theorising approach (Usher & Bryant, 1989). An interrogation of the literature is used to illuminate reflection on the practice situation of using third-party observation of teaching as part of a credentialed teaching development programme (UCD, Dublin’s Graduate
Diploma in University Teaching and Learning). The question asked of the literature is ‘What are the different types of third-party observation of teaching?’ In reviewing the (portfolio-based) written reflections of lecturers on the programme, the question asked is, ‘What type of third-party observation of teaching are they experiencing?’ Following this identification of type, two final questions are asked, namely, ‘What benefits have accrued from the process?’ and ‘Are any of these benefits enhanced by the particular features of the type of observation being used?’

Categorisation of different models of third-party observation of teaching: a review of the literature

Gosling (2002), for example, identifies three general models of third-party observation differentiated on the basis of the relationship between the observer and the observed, and the use to which information relating to the observation is put. Unfortunately, he uses the generic term ‘peer-observation’ to describe all three, despite his explicit recognition of the difficulties and contradictions inherent in using this term, where there is a disparity of power or a lack of trust between observer and observed.

Gosling’s categorisation of third-party teaching observation regimes, although problematic (see below), is a useful response to the diversity of practice that exists in this area. There is, it seems, a new consensus that observations of individual teachers can lead to better quality teaching by those individuals and that, where this form of observation is widespread, there can be widespread improvement. Brown (1993), Beaty (1998), Martin & Double (1998), Bell (2001), Slade (2002), Hammersley-Fletcher & Orsmond (2004) and Shortland (2004) are among those who provide convincing evidence that third-party observations can lead, and have led, to improvements in teaching. This, however, is where consensus stops. Third-party observation of teaching may be widely practised in Higher Education, at least in most English-speaking countries, but it is not commonly practised. Practical and philosophical approaches vary greatly between countries, between institutions within countries and between units within institutions.

The first of Gosling’s models, ‘The Evaluation Model’, involves observation of a teacher by a superior within the management structure (or someone acting on behalf of such a person), and includes the formal assessment of an individual’s teaching skills as part of the management processes of a system or of an institution. Audits and inspections, whether internal or external, fall into this category, as do a whole range of managerial activities that include third-party observation. An observation may, for example, serve to inform an annual appraisal, confirm satisfactory performance during a probationary period or be part of the system for judging eligibility for promotion. Although often labelled ‘peer-observation’, these are all clearly processes that relate to evaluating and managing employees, and it can only be that aspect of the peculiar mythology of teacher professionalism that deems all academics to be peers within the academy that enables the label ‘peer’ to be applied to such inspections. It can be argued that such inspections are essential to quality assurance and accountability. Whether or not
this is true, they need to be distinguished from processes where the primary focus is continuing improvement through structured reflection.

Gosling identifies a second type of third-party observation which he locates in what he calls 'The Development Model'. This also involves formal summative assessment of an individual’s teaching skills. Rather than managing employees, however, the main purpose is to give expert feedback in order to encourage improvement in teaching competencies - often by way of an agreed action plan. This is the model used within programmes leading to a formal qualification in teaching, such as a Higher or Graduate Diploma in Ireland or a Post-Graduate Certificate in Education in the UK. It is also frequently used, particularly in England and Wales, as part of an institution’s formal professional development strategy, and can be found in those universities in Ireland where a teaching portfolio requiring some evidence of third-party evaluation of teaching forms part of the promotion system. Gosling suggests that this model can be clearly differentiated from the first but, in practice, the boundaries are anything but clear. It is, for example, difficult to find a ‘development model’ that does not contain elements of managerial appraisal. Observations carried out by educational developers tend to have the dual purposes of, on the one hand, giving feedback to enable improvement and, on the other, testing whether the teacher who has been observed meets certain standards - whether these are for the purpose of gaining a qualification, promotion or to satisfy internal quality assurance systems. It is hard to justify regarding this as qualitatively different from an observation that is overtly within the evaluation model. The key issue concerns what happens to the information generated by an observation. If reports on the quality of a teacher’s performance go back to that person’s line manager then, regardless of whether or not the observation was conducted by someone of equal status within the institution (literally a ‘peer’), the observation constitutes an evaluation and the ‘peer’ observer concerned must be seen as acting in place of the manager - however unwillingly. Similarly, where observations are overtly evaluational, it is often in-house educational developers who do the actual observation and who, as well as reporting to management the achievement (or non-achievement) of standards, will see it as their job to give constructive feedback.

Gosling’s third category is ‘The Peer Review Model’. This is intended to be a more formative process than the first two - one where teachers observe each other, discuss the experience and engage in ‘non-judgemental constructive feedback’ (Gosling, 2002, p. 5) - although the issue of how you give feedback without making judgements is not addressed. Mindful of the issues surrounding disparities of power and the need for mutual trust, he suggests that this model can only be fully successful where all the staff involved:

- are regarded as genuine peers, in which there is real mutuality and respect for each of the participants as equal, whatever their status in the department. (ibid, p. 2).

In this model, the primary purpose is that teachers, one acting as an observer and one being observed, engage in mutual reflection on a teaching session in order that both can learn from it in order to inform future practice. This, then, is the only Gosling model that can really be said to describe ‘peer’ observation, and it can be clearly differentiated from the other two by reference to what Gosling calls ‘confidentiality’ - by which he means who controls what is reported, to whom and in what format. This is the crucial difference that makes this model compatible with the idea of the teacher as reflective practitioner - since it is the only model in which the reflection does not take an inevitable second place
to concerns about managerial/inspectorial judgements. (It is, presumably, the absence of these particular forms of judgement that leads Gosling to describe the process as ‘non-judgemental’.)

Even though Gosling is explicit on ‘confidentiality’ being an essential part of the boundary conditions of this model, it has become an all-too-frequent practice for institutions to claim that all third-party observations conducted by academics on other academics are ‘peer-observations’, regardless of the relative status and power positions of the observer and observed, and regardless of what happens to the information gathered. The importance of this latter point should not be underestimated. Shortland (2004) documents a ‘typical’ (p. 219) scheme where trios of teachers working within subject disciplines were required to report on mutual observations via a ‘one-way upward communications channel’ (p. 226). What is most interesting, and optimistic, about this report is that despite receiving no feedback from management, some, at least, of the teachers involved used the observation process to devise and inform their own self-directed professional development operating in parallel with, but divorced from, the official reports. In effect, two observation schemes were operating— one formal, where evaluative data was passed upwards, a second informal, where data was used within the team ‘to promote self-knowledge and personal development’ (p. 227).

Thus, while being a useful categorisation for some purposes, the lack of firm boundaries, arguments over definitions and, frankly, occasional deliberate obfuscation by managers faced with teacher resistance to appraisal by observation, can make it difficult to identify actual practice as being in any particular one of Gosling’s three models. This makes them of limited utility in describing the reality of the disparity of practice, since aspects of all three modules are likely to be found in most actual systems. The permeable nature of the boundaries compounds the problems inherent in describing observations where there is a disparity of power as being ‘peer’. This allows, among other things, misleading descriptions to be adopted where, for example, staff-appraisal or assessment of candidates for teaching qualifications by observation is sold as peer-review - and such deceptions, in turn, exacerbate any pre-existing teacher suspicion of any form of third-party observation.

Bell (2002), who was responsible for one of the most interesting and informative reports of a (genuine) peer review regime in practice (see Bell, 2001), takes a slightly different approach to categorising models by postulating two variants, the ‘Performance Model’ and the ‘Development and Training Model’. These correspond to Gosling’s first and second models, but within these models she considers the added dimension of the comparative formality of process. In her formulation, Gosling’s ‘Peer Review Model’ becomes an informal version of the ‘Development and Training Model’. Differences in terminology notwithstanding, the correspondences between Bell and Gosling are such that our critical discussion of the latter can be taken to apply equally to the former. Bell, incidentally, seems even more aware of the inappropriateness of describing ‘performance’ or ‘evaluation’ observation as ‘peer’. Conversely, she feels that it is too narrow a descriptor for those regimes that involve
structured reflection, since it is the reflection, rather than the observation, that is crucial. (Incidentally, the team responsible for the Dublin diploma also felt reflection-on-observation to be more important than the actual observation. As we shall see, it was this focus on reflection that enabled them to design an assessment system that did not require data from the observation.)

The lack of clarity in both the theory (what is meant) and practice (what is being done) of ‘peer’ (third-party) observation of teaching permeates the literature, and makes things confusing for the unwary researcher and/or practitioner.

**Categorisation of third-party observation of teaching in terms of who controls the information flow: a new proposal.**

Close study, however, of both theory and practice allows a different categorisation; one based not on who does what but, rather, on who controls the data flow - who, in other words, decides what is reported, how it is reported and to whom. This is, fundamentally, an issue of power; but a very specific power: that of reporting professional competence. In this instance, information is, literally, power and, unless restricted by enforced confidentiality, the desire to look good on a report being sent up the hierarchy must be of greater concern than possible gains from exposing areas that could benefit from reflection and development. Concern about the content of third-party observation reports being seen by superiors or others is not confined to those whose teaching skills leave much to be desired. Indeed, it is much easier to give a slick, packaged presentation where students are expected to remain passive and quietly absorb the information being offered than it is to operate an active classroom or lecture room, where students are not only learning but are seen to be learning. Weak teachers, therefore, tend to adopt the active teacher/passive student mode when being observed, and so the process does not capture their weakness. More importantly, most teachers, like most professionals, have a natural desire to have their contributions to society recognised and validated, and third-party observations that are passed up the hierarchy or otherwise circulated provide a rare chance for this to happen - especially in societies where teachers are undervalued. What is more natural (or more reasonable) than to put on your best show when managers, superiors or ‘inspectors’ are present?

Further, it is not unusual for it to be the most innovative, admittedly sometimes idiosyncratic, but highly effective teachers who are most concerned about the reports - and this is because there is all too often a perception that good teachers’ conceptions of what constitutes good teaching are not shared by senior management. (That they are often wrong and that senior managers are usually well aware that ‘good’ teaching means challenging, motivating and activating students does not change the effect of this perception.) This situation is compounded by the effects (which go way beyond the boundaries of the jurisdictions involved) of some of the ‘tests’ of ‘good teaching’ imposed by some of the States of the USA, and the experiences relayed on the grapevine in the UK (and, occasionally, openly stated in conferences) of external peer reviewers who were less than current in their
understanding of teaching theory. These have led to the even more widespread perception that external reviewers acting on behalf of funding agencies are frequently not competent to judge teaching. (In this case there is less chance of the perception not reflecting the reality.)

The result is that where third-party observation reports are seen by managers, inspectors or other agents of the public or funding bodies, there is a strong imperative to concentrate on existing strengths rather than on the development and improvement of weaker areas; yet, if we want to improve the overall quality of teaching, it is these areas that we need to subject to the rigours of (genuine) peer-observation in order for teachers to be able to engage in effective self-managed professional development. Thus, there is a need to switch the guiding imperative to one where teachers actively seek out areas that need developing, subject them to scrutiny by peers, and use the professional dialogue and shared reflection that follows to improve professional practice. For this to work, however, the teachers concerned must be freed from the fear that exposing weakness within this process could adversely affect their careers. Consequently, it is essential that it is the person being observed who has total control over what happens to the information generated by the observation and subsequent discussions.

In terms of theory, this leads, inevitably, to the conclusion that what really matters in differentiating systems of third-party observation is whether or not the person being observed has full control over what happens to information about the observation. Thus, despite many variations in detail, regimes of third-party teacher observation can be most usefully categorised as falling into one or other of the following two types.

**Type-A: control by observe**

In this model, access to the information generated by the observation is exclusively in the hands of the person being observed: the observer being bound by a strict code of confidentiality. This is the only form that can, legitimately, be called peer-observation, since in any other scenario the observer/reporter has power over the observed and the relationship is not one of equals.

**Type-B: control by others**

In this model, access to the information generated by the observation is not exclusively in the hands of the person being observed; rather a report can be made about them without their consent and without their having control over content or distribution.
Recent studies have identified that aversion to being monitored leads to a reluctance to participate in peer-observation (Peel, 2005; Adshead et al., 2006). This suggests that type-A schemes are likely to be more successful in prompting improvements in teaching than type-B. In order to explore how strongly university teachers tend to agree with this proposition, a study at UCD, Dublin, looked at the written reflective statements of 22 teachers who had experienced a type-A process as part of a Graduate Diploma in University Teaching and Learning.

At this point, it is appropriate to ask whether peer-observation that takes place inside a credentialed programme can be anything other than type-B.

It could be argued that if one of the reasons for undergoing peer observation of teaching is to produce evidence in order to achieve a qualification, then the need to satisfy a third-party with superior power (that of being able to award the qualification) has all the hallmarks and implications of the type-B model as described above.

In the case of the Dublin diploma, however, this was not the case. The relevant diploma module did not require any evidence from the observation beyond that it had, in fact, taken place. What was assessed in the module was evidence - in the form of a reflective statement* of how a teacher had done something different following feedback from an observation. The assessment for the diploma did not require data about the actual observation. Indeed, although many chose to reveal such data because they felt it helped explain the reasons for, and form of, their subsequent actions, such revelations were actively discouraged precisely because the course team wished to establish the principle that the data from the actual observations were irrelevant to the passing of the associated module.

**Reflections on being observed teaching within a system where the observer retained full control of the information flow**

The issue of control over the peer-observation process was the dominant theme in the written reflections on peer-observations by lecturers who had been observed. (Each of the 22 teachers was assigned a number at random. Where quotes are given, the numbers in parentheses indicate the individuals assigned this number.) From their writing it is possible to identify six dimensions of control:
1. Control over whether or not to participate in the first place (control over entry to the POT scheme).

2. Control of choice of observer (note: in the system under review, peer-observers were formally known as ‘peer-mentors’).

3. Control of focus of observation.

4. Control over how feedback shall be given (including control over the structure and form of any documentation used to record feedback).

5. Control over all information generated by the observation.

6. Control over what is done as a result of the observation (including planned improvements and whether or not to engage in further observations).

The first four dimensions can be said to relate to the pre-observation stage of the process, while the final two encompass the post-observation stage.

**Observee Control over the pre-observation stage of the process**

In the case of the 22 lecturers surveyed, they cannot be said to have had a free choice over whether or not to engage in the process in the first place because the process was required by the credentialed programme on which they were enrolled. Entering the credentialed programme was, however, entirely voluntary, and they did have control over all the other dimensions.

Being able to choose an observer rather than having them assigned was seen as very important by all 22 of the lecturers:

The most important part of the process was the ability to select a peer-mentor I could feel comfortable with. I wanted to select someone who was an experienced teacher who could help me identify areas that needed further development, someone who would not be judgemental, but rather constructive and open. Finally, I wanted someone with whom I could continue to discuss my teaching practice over time. [9]
Two participants specifically mentioned that it was important to be able to ‘trust’ and ‘respect’ the ‘peer observer’, and that this affected their choice of observer [2,8].

Another key dimension of control in the pre-observation phase was that concerning the focus of the observation that was discussed at a pre-meeting. A clear majority of the observees felt that this was an essential part of the process.

At the pre-observation meeting I highlighted to the observer particular areas of concern I had regarding the session and recommended the focus of the observation be the extent to which I had been successful in promoting active learning . . . [3]

In addition to control over the focus of the observation, those being observed were also given a choice over how the observation was recorded. Many participants commented on how this choice over the format of the feedback documentation actually helped them to decide on the focus of the observation.

Overall, an analysis of the written reflections demonstrates that the participants perceived that they were in a type-A system*/where they had total control of the data flow. It also confirms that most considered it important that they had control over all of the following dimensions of the pre-observation phase: choice of mentor, focus of the observation and structure of the feedback documentation.

**Observee control of the post-observation flow of data**

Participants highlighted two issues of post-observation control, namely: control of feedback data and choice over whether or not to engage in peer-observation in the future.

The confidentiality of the data generated was considered to be a very positive aspect by all participants.

. . . peer-observation as practised in UCD, has several very positive features. It is non-threatening. It is confidential. As I have developed a trust relationship with the observer it will likely become progressively more useful to me and my students as time goes by. [18]
Following feedback on the peer-observation session, participants drew up action plans to develop and improve their teaching. The content and form of this action plan was controlled by the participants but was, most often, in the form of a learning contract with their peer mentor.

By producing a learning contract I have begun to move to a new level of practice with an ability to initiate change to improve my work. [10]

Having used peer-observation of teaching once, participants had a choice whether to use it in the future or not. One participant reflected that:

I consider that the focus on the fine detail requested by me in this peer-review, essentially the specific ‘fine tuning’ component of POT (peer-observation of teaching), is not the major benefit of the process. As in the story of Jack and the Beanstalk, it is not the individual golden egg that will benefit Jack and his mother, but rather the goose that can produce the egg. In the same way I feel I have almost been “empowered” by going through the learning process of POT, both being the teacher being observed and the observer, my intention is to make it an integral part of my continuing professional development. [8]

Participants on the Graduate Diploma in University Teaching and Learning had, however, no control or choice about whether or not to engage in the process of peer-observation. Peer-observation was a mandatory part of a core module:

My peer-observations took place in the context of my undertaking UCD’s . . . Diploma in University Teaching and Learning. It was a mandatory component of this course. I have little doubt that I would not have partaken in a peer observation process under other circumstances. I initially approached this component of the course with some scepticism and also a certain amount of trepidation . . . However, as the (diploma) progressed and the rationale and the process of peer-observation was explained, my scepticism largely vanished and my fears were greatly reduced although I did still feel nervous about the process. . . . By the time it came for me to be peer observed, I was certainly open to the experience and I approached it as an opportunity to explore approaches to enhance student learning in my lectures. [2]

Where peer-observation takes place in another context (where it is not a

mandatory part of the course) then there is the possibility of academic staff having control over whether they personally decide to participate in a peer-observation initiative that is operating in their department/special interest group/school/faculty or college. Having said that, the Graduate Diploma was not compulsory, and academic staff made a personal choice as to whether or not to participate.
A model of peer-observation with control by the observee

From the (unprompted) reflections of the teachers undergoing the UCD peer-observation process, it appears that type-A models of peer-observation (control by the observee) can be said to have six key dimensions as shown in Figure 1.

Figure 1.

In the case of the UCD process, participants had total control over five of the six dimensions.

All 22 of the UCD participants noted (in one way or another) that the peer-observation process, where they had control, choice and confidentiality, encouraged them to try teaching strategies new to them, as well as getting feedback on established methods. One participant, who asked a peer-observer to
observe her in sessions she was having difficulties with, reflected that being able to choose the observer was key to using peer-observation in this developmental way:

I think if the ability to select one’s peer observer is denied, then I would be far less inclined to select a teaching session that I had some difficulty with, and do what most of my trainee teachers did at second level and actually rehearse the class in advance so as to put on a good performance for fear of not being applauded. [9]

Another participant chose to be observed once using a well-established lecture method and, subsequently, in trying out a new teaching approach. Having control of the foci of the two observations was central to this developmental approach:

I asked my peer to observe:

1. a “traditional” style first year lecture
2. a new approach (for me) with my second year course where I divided the class into small groups and asked them to solve some problems during the lectures. This was my first attempt at using such a technique.

I chose to have the first year lecture observation carried out first because I was confident about my ability to do this well. I was more nervous about the second observation as it was a first time experience for me in trying out a new teaching method . . . The process has to some extent re-energised my teaching. [2]

It was evident from the reflections that the real sense of control and choice involved encouraged real development of and improvements in teaching:

The affirmation and encouragement was for me perhaps the most important aspect of the process. This was all the more the case in that in both of my peer-reviewed lectures I was experimenting with new techniques in teaching. The need for affirmation should not be underestimated in lecturers who rarely ever have the privilege of having qualified colleagues sit in on lectures. [1]

The other major benefit of peer-review was that it gave me more confidence to experiment with new techniques. [8]
Conclusion

This paper has fulfilled two functions. First, from a review of the literature it has proposed a theoretical model of peer-observation where control of the data-flow, and other key elements of the process, rests with the person being observed. Secondly, it has explored the dimensions of this model in practice. Evidence from the use of the model in UCD, Dublin, suggests that observee control over five of the six identified dimensions of the process was fundamental to encouraging a focus on improvement. Thus, in answer to the question, ‘What benefits have accrued from the process?’ each of the 22 UCD teachers could point to tangible and documented improvements in teaching that resulted from the observations. In answer to the question, ‘Are any of these benefits enhanced by the particular features of the type of observation being used?’ the participants were in no doubt that having the control over the five key dimensions of:

1. choice of observer
2. focus of observation
3. form and method of feedback
4. resultant data-flow
5. next steps

encouraged them to focus on improvement of practice rather than demonstration of existing good practice.

This is not to say that type-B observations (those where control is with someone other than the person being observed) cannot lead to effective professional development. Even what appear to be meaningless paper-chasing exercises, such as that reported by Shortland, have been shown to lead to improvement, but there has to be a serious question as to reliability when the imperative is to demonstrate competence to a third-party.

There is no guarantee that type-A models will inevitably lead to improved teaching, but it seems more likely theoretically, and the Dublin experience seems to bear this out in practice*/in as much as the teachers involved strongly expressed the view that without control over the five dimensions (and especially of the data flow), they would have made different choices over what they allowed to be observed.
Type-A models produce genuine peer-observation by equalising the power relationship between the observer and the person observed. As control of all derived data remains with the latter, the threat to career prospects of situations where the results of observations might be communicated to managers disappears; and so does the temptation to simply showcase existing strengths. The imperative is switched to one where teachers actively seek the help of colleagues in improving the learning experience of students.

References:


