

Power, policy and performance: learning lessons about lesson observation from England's Further Education colleges¹

Matt O'Leary, Centre for Research and Development in Lifelong Education (CRADLE), Faculty of Education, Health and Wellbeing, University of Wolverhampton, UK

Abstract

Lesson observation has been widely debated in education circles in recent times. From politicians to practitioners, everyone seems to have a view on it. Surprisingly, however, very little empirical research has been done on this important area of practice. With this in mind, this article explores some of the findings from a national research project investigating the use and impact of lesson observation on the professional lives of thousands of staff working in the Further Education (FE) sector. The project, sponsored by the University and College Union (UCU), adopted a mixed-methods approach and was carried out over a year (2012-13). This article argues that lesson observation has become a central crucible in which power is exercised over individuals and institutions alike under the guise of measuring and improving teacher performance. It considers what lessons can be learnt about its current use and concludes with recommendations as to what needs to change to enable the education sector to make better use of observation in the future.

Introduction

This article discusses the hot topic of lesson observation. Whether it is through the popular press, academic publications or social media, lesson observation has undoubtedly triggered intense discussion amongst the education community of late. From classroom teachers to senior managers, the Department for Education to Ofsted, there is no shortage of opinion about observation and its role in evaluating the quality of educational provision in England's schools and colleges. Despite its prominence in the educational arena as a tool for teacher evaluation and development, it may come as a surprise to some readers to discover that this important area of practice is relatively under-researched.

One of the key goals of this article is therefore to begin to address this shortfall in research evidence by exploring some of the findings from a national project (UCU 2013) investigating the use and impact of lesson observation on the professional lives

of thousands of staff working in a wide range of contexts and institutions in the Further Education (FE) sector in England. The project, sponsored by the University and College Union (UCU), adopted a mixed-methods approach and lasted a year (2012-13). It represents the largest and most extensive research study in the English education system to date and as such marks an important milestone in lesson observation research. In drawing on some of the project's key findings, along with previous cognate work, this article considers what lessons can be learnt about the current and continuing use of observation. It concludes by making recommendations as to how the education sector might make better use of observation in the future and what needs to change in order for this to happen.

The research context

The context in which this project was carried out was the FE sector in England. Although there are similarities between schools and FE providers, the contextual differences separating the two sectors are not insignificant and thus worth mentioning briefly. FE is often referred to as a 'complex' and 'diverse' sector (e.g. Huddleston and Unwin 2013). This complexity and diversity is typically reflected in the breadth of courses offered, the size of FE providers and the communities they serve, the wide-ranging profile of its staff and students, along with the fact that it is arguably the most market-tested area of public sector reform in England. It is the latter of these, in particular, that has impacted on the way in which observation has evolved and been implemented as an educational intervention over the last two decades, though this is an issue to which I will return below when discussing the status quo of observation. Before then it is important to understand the wider policy backdrop in which this project took place.

Lesson observation is a polemical issue that has attracted much discussion amongst the education community over the last year, particularly through social networking sites such as Twitter. In FE such discussion is not new, but has been simmering away

on the sector stove for up to a decade now. In the lead up to this project, the use of observation had already been identified as 'an increasingly common flash point in colleges, triggering local negotiations, and in some places industrial disputes' (UCU 2009: 1). The level of discontent experienced by union members in some colleges has even resulted in the boycotting of lesson observations altogether (UCU FE News, May 2012). Where industrial disputes have occurred, this has often been in response to what practitioners have perceived as draconian and punitive observation policies imposed on them by senior management. Such policies have largely sought to link the outcomes of formal, graded lesson observations to disciplinary or capability procedures. Thus the policy backdrop to the project was one in which observation had become a highly charged and contested area of practice for the FE workforce.

The research project adopted a mixed-methods approach involving quantitative and qualitative methods of inquiry. An online survey, semi-structured interviews and focus groups were the main research tools used as part of a triangulated framework to address the project's research questions. Some of these questions were of a factual nature and thus lent themselves to a quantitative method of inquiry. Others sought to explore the experiences and perspectives of practitioners in the form of a narrative and so required a qualitative approach.

The sample for the first phase of the data collection (online survey) comprised UCU members, ranging from part-time tutors to senior managers. Approximately 4,000 respondents completed the survey (n = 3976). The second phase involved staff from a range of colleges across England, including UCU members and non-members. Purposive sampling was used to select the colleges to ensure a geographical spread, thus colleges were selected from the north, the midlands and the south of England. Naturally the scope of this article does not allow for detailed coverage and discussion of the data, but a full copy of the project report (UCU 2013) can be requested from UCU.

The status quo of lesson observation in FE

To get an overview of the contexts and purposes for which lesson observation was used in FE, a specific question was included in the online survey that required respondents to indicate the context that best described their most recent experience of observation. **Figure 1** below presents a summary of the responses to that question.

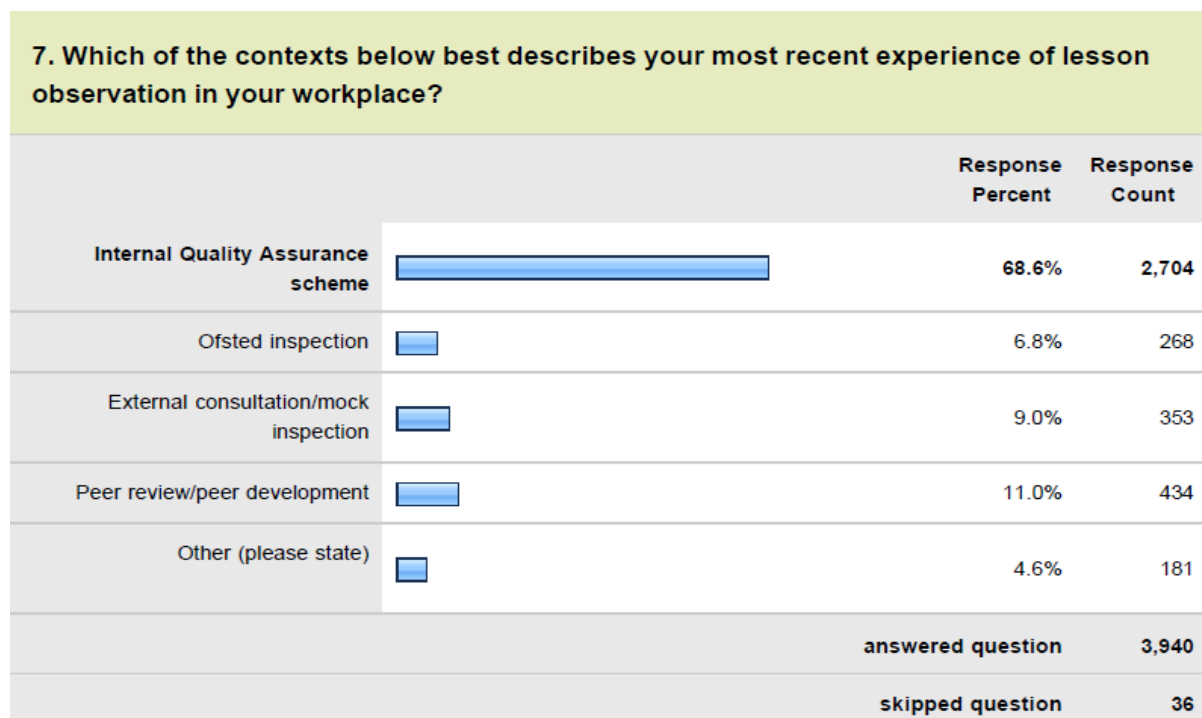


Figure 1 – Contexts of lesson observation

As **Figure 1** illustrates, only just over a tenth of respondents (11%) chose peer review/development, suggesting this particular use of observation was relatively marginal in FE, with much of it taking place in the context of Initial Teacher Education (ITE) courses. The most common response selected by over two thirds (68.6%) was the Internal Quality Assurance (QA) scheme, which typically mimics the approach adopted by Ofsted when carrying out lesson observations during inspections i.e. where the lesson is evaluated and graded against its 4-point scale (Ofsted 2012). Similarly, the context of ‘external consultation’ follows the Ofsted model and is used as a ‘Mocksted’, where external consultants are employed to carry out observations across the institution. So when combined, the first three contexts

listed in **Figure 1**, all of which adopted a similar performance management approach, amounted to over four fifths (84%) of responses.

Although there was evidence of 'alternative models' in practice, they were relatively marginalised and tended to operate on the peripheries of most formal systems of accountability. For example, ungraded models of observation were in use in some institutions, though only accounted for a tenth of current practice. Similarly, peer observation, whilst not uncommon, occurred mainly as an informal, unaccredited activity that staff undertook on a voluntary basis. These alternative models were rarely viewed by senior management teams (SMTs) with the same level of importance as their performative counterparts and tended to be valued more highly by practitioners. Furthermore, there was evidence of apprehension among some providers in implementing alternative and/or ungraded models of observation on a formal basis for fear of going against normalised practice and leaving themselves open to increased scrutiny from Ofsted, as discussed below and elsewhere (e.g. O'Leary and Gewessler 2014).

The fact that lesson observation has become associated with performance management in education over the last two decades is no accident but a consequence of the wider political and economic 'neo-liberal reform agenda' that has sought to 'transform the working cultures of public sector institutions' (O'Leary 2014: 11). During this period, the medium of classroom observation has evolved into a catch-all crucible in which performance data and evidence of improvements in teaching and learning simultaneously bubble away. Driving this reform agenda are three interwoven policy 'technologies' that have shaped teachers' work, notably 'markets', 'managerialism' and 'performativity' (Ball 2003). The way teachers experience these 'technologies' in their working lives is concretised through systems of quality assurance, target setting, continuous improvement etc, or what Green (2011) neatly refers to as the language of 'managerialese'.

In discussing the impact of high-stakes testing on teachers' work, Stevenson and Wood (2013) argue that 'it is the pincer movement of markets and managerialism that have combined to effectively and radically re-shape teachers' experience of work' (p. 43), underpinned by the aim of making the complex processes of teaching and learning measurable. Smith and O'Leary (2013) refer to this as 'managerialist positivism', where the complexity of the teaching and learning process is reduced to the presentation of quantitative performance data, which is valued more highly than the qualitative kind on the premise that its outward-facing measurability imbues it with an increased authenticity and credibility:

Managerialist positivism functions to provide and impose quantitative wholeness on the unfinished totality of the present for the (political) purposes of policy intervention and the allocation of financial resource. It can be seen as the ideological veil that normalises the representation of complex sociological and qualitative phenomena in reductive and numerical forms (Smith and O'Leary 2013: 246).

The FE workforce perspective on current approaches to lesson observation

Figures 2 and **3** below contained a list of questions relating to graded and ungraded approaches to lesson observation respectively. Overall, levels of agreement regarding the benefits and effectiveness of graded observation as a method of teacher assessment and improvement were consistently much lower than disagreement as shown in **Figure 2** below and markedly lower compared to ungraded approaches in **Figure 3**.

10. I believe that graded lesson observations ...						
	Strongly agree	Agree	Disagree	Strongly disagree	N/A	Rating Count
are essential for monitoring the quality of teaching and learning	7.4% (274)	32.3% (1,194)	38.2% (1,412)	21.3% (788)	0.7% (24)	3,692
are essential for improving the quality of teaching and learning	6.6% (244)	27.1% (998)	42.2% (1,556)	23.5% (866)	0.7% (24)	3,688
are essential for the continuing professional development (CPD) of staff	5.9% (218)	23.6% (871)	44.5% (1,644)	25.4% (938)	0.6% (21)	3,692
are the most effective method of assessing staff competence and performance	1.8% (67)	12.1% (446)	46.8% (1,724)	38.4% (1,414)	0.9% (34)	3,685
are a reliable indicator of staff performance	1.1% (40)	11.1% (412)	43.9% (1,626)	43.0% (1,591)	0.8% (31)	3,700
have helped to raise the standards of teaching and learning in my workplace	2.1% (79)	17.9% (658)	44.2% (1,625)	32.5% (1,195)	3.3% (121)	3,678
have helped me to improve as a classroom practitioner	3.3% (120)	19.3% (711)	42.4% (1,564)	32.4% (1,195)	2.6% (97)	3,687
are the fairest way of assessing the competence and performance of staff	0.9% (32)	9.7% (357)	46.7% (1,716)	41.4% (1,522)	1.4% (51)	3,678
are a necessary part of staff appraisal	2.5% (92)	23.4% (864)	42.4% (1,562)	30.6% (1,129)	1.1% (40)	3,687
should no longer be used as a form of assessment	37.0% (1,365)	30.4% (1,120)	23.1% (852)	7.0% (257)	2.5% (94)	3,688
					answered question	3,735
					skipped question	241

Figure 2 – Graded lesson observations

Approximately two fifths (39.7%) of all participants agreed that graded observations were essential for monitoring the quality of teaching and learning compared to just under three fifths (59.5%) who disagreed. This was by far the highest level of agreement recorded for the questions on graded lesson observations across all groups. One plausible explanation as to why there were higher levels of agreement in response to this first question across all groups compared to other questions in this section of the survey might be that as graded observations have become normalised in FE in recent years (O’Leary 2013), so staff have become

increasingly conditioned to expect to be graded on their classroom performance whenever they are observed. As Brian, a curriculum coordinator and interviewee in one of the focus groups aptly commented:

We're so used to getting a grade now when we're observed that even if a colleague does something like a peer observation of you, there's a part of you that still wants to know how they'd grade it even though that's not the point.

Two thirds (65.7%) of all respondents disagreed that graded observations were essential for improving the quality of teaching and learning and three quarters (74.8%) disagreed that graded observations had helped them to improve as classroom practitioners. This level of (dis)agreement was similarly reflected in responses to the question relating to whether graded observations had helped to raise the standards of teaching and learning in their workplace. However, levels of agreement amongst senior manager and observer respondents contrasted starkly with practitioners with over four fifths (86.4%) of senior managers and over half of observers (59%) agreeing with the statement that graded observations had helped to raise the standards of teaching and learning in their workplace. This is an interesting contrast in perspectives given that annual graded observations tend to be resource-intensive activities in FE with many providers investing a lot of time and money in them. Such schemes are expected to meet the dual purpose of monitoring and improving the quality of teaching and learning. Yet if there is such a significant level of disagreement across groups as to the benefit of graded observations, one has to question the extent to which they satisfy either or indeed both of these purposes.

The highest and most striking levels of disagreement recorded in the responses to the use of graded lesson observations converged around the topic of teacher assessment. Over four fifths (85.2%) disagreed that graded observations were the most effective method of assessing staff competence and performance. A similarly

high level of disagreement was recorded in response to whether they were regarded as a reliable indicator of staff performance. However, the highest level of disagreement (over 88%) of all the questions in this section was the response to whether graded observations were considered the fairest way of assessing the competence and performance of staff. In contrast, only a tenth (10.6%) of all respondents agreed with this statement. When comparing cross tabulations, similarly high levels of disagreement featured in the responses of all practitioners regardless of the length of teaching experience. Over two thirds of observers also disagreed with these two statements, with senior managers being the only group to register greater levels of agreement than disagreement. This was completely out of sync with all other groups and was indicative of a significant divide in perceptions between senior managers and the rest of the FE workforce.

It was interesting to note that in response to the final question, two thirds (67.4%) agreed that graded lesson observations should no longer be used as a form of teacher assessment, though once again there were differences in opinion between practitioners, senior managers and observers with the latter two strongly in favour of retaining the use of graded observations. Such differences in opinion are not surprising given the importance attached to the collection of this quantitative performance data in many FE providers.

The lesson observation 'grade profile' (i.e. statistical data sets of how many lessons are graded as a 1, 2, 3 or 4 in any given year), as it is commonly known, has become custom and practice in FE in recent years and is relied on heavily by SMTs as a key tool with which to attempt to measure and compare levels of staff performance internally and against national benchmarks. Thus the compilation and scrutiny of statistical data from annual graded observations is seen as an essential component of the performance management cycle for SMTs in monitoring and assessing the quality of teaching and learning across the institution, despite the scepticism

expressed by some senior managers in other studies as to the value of this practice (e.g. O’Leary 2013). Others have commented on such practice as an example of ‘scientific management’ or ‘neo-Fordism’ (e.g. Boocock 2013), the basis of which is to attempt to categorise teachers’ work into quantifiable units that it claims can be measured ‘scientifically’.

Despite the positive responses in support of the use of graded observations expressed by SMTs in discussion of the survey data, there was a surprising dearth of comments to reinforce this level of support in the qualitative data. Even in those instances where comments were broadly supportive of the use of graded observations, they were consistently accompanied by a conditional statement, as the quote below concisely exemplifies:

Everybody wants to do a good job; lesson observations, be them graded or ungraded, announced or otherwise, are a key part of the Quality Assurance and a positive experience if the emphasis is on the development of the practitioner and are not used as an excuse for disciplinary or 'capability' procedures (Rose, senior manager)

In contrast to **Figure 2** above, the responses to questions about ungraded lesson observations in **Figure 3** below reflected a very different picture. Overall, ungraded observations were viewed more favourably than their graded counterparts, with responses to similar questions registering a much higher level of agreement among practitioners and observers, with senior managers being the only outlier once again.

Over four fifths of respondents (81.2% and 81.3%) agreed that ungraded lesson observations were a more effective means of improving the quality of teaching and learning and played a more important role in the CPD of staff than graded models.

Levels of agreement were slightly lower among observers, though responses still registered at least two thirds. Such high levels of agreement may well be due to the removal of the ‘fear factor’ often associated with the high-stakes nature of grading teaching performance, particularly when so much is dependent on the outcome of these annual assessments that capture such a minute part of teachers’ work.

11. I believe that ungraded lesson observations ...						
	Strongly agree	Agree	Disagree	Strongly disagree	N/A	Rating Count
are more effective in improving the quality of teaching and learning than graded lesson observations	31.0% (1,103)	50.2% (1,786)	12.5% (445)	2.8% (100)	3.5% (126)	3,560
play a more important role in the CPD of staff than graded lesson observations	30.0% (1,064)	51.3% (1,821)	12.3% (436)	2.8% (98)	3.6% (129)	3,548
are more effective in assessing staff competence and performance than graded observations	27.2% (963)	49.1% (1,738)	16.3% (578)	3.3% (117)	4.1% (147)	3,543
are a reliable indicator of staff performance	16.0% (566)	43.0% (1,520)	30.5% (1,079)	6.5% (229)	4.0% (143)	3,537
have helped to raise the standards of teaching and learning in my workplace	11.4% (401)	33.3% (1,174)	21.4% (756)	5.3% (187)	28.7% (1,012)	3,530
have helped me to improve as a classroom practitioner	14.9% (527)	33.8% (1,195)	20.3% (717)	5.4% (192)	25.6% (908)	3,537
should replace the use of graded lesson observations	40.9% (1,450)	35.7% (1,266)	13.4% (475)	4.0% (143)	6.0% (211)	3,545
				answered question		3,588
				skipped question		388

Figure 3 – Ungraded lesson observations

Only senior managers disagreed (63.6%) that ungraded lesson observations were more effective than graded ones. Given previous discussion about the importance of collecting quantitative performance data and the reliance of SMTs on the observation grade profile as a measurement tool, such a response is hardly surprising. However,

it might also be considered symptomatic of a risk-averse culture amongst SMTs to explore alternative models of observation or assessment for fear of failing to comply with the hegemony of normalised practice in FE. In the following quote, Paul, a senior manager, illustrates this by describing how his college contemplated moving to an ungraded model but with an Ofsted inspection imminent decided against it:

We toyed with a number of models and we had links with another outstanding college and knew they'd decided to scrap grading altogether, although they did this shortly after a successful Ofsted inspection and it's quite interesting when you look at the colleges that do abandon grading, they're almost exclusively colleges who have just been through a successful Ofsted inspection so they're not expecting an inspection team to return for a number of years. We weren't in that position because we were inspected in 2009 so we are expecting to be inspected this year. So we didn't really feel the time was right or it might be advisable to lose grading altogether just before an inspection.

Paul's comments not only reinforced arguments about the normalisation of graded observations in FE (O'Leary 2013), but also revealed the apprehension of some providers in choosing to implement alternative and/or ungraded models. Added to this was the conditionality of professional autonomy and how it was linked to inspection performance. For Paul, those recently judged successful were more likely to experiment with new approaches, as 'they're not expecting an inspection team to return for a number of years' and as such were afforded more freedom to do so.

Where opinion seemed more divided was in the responses to whether ungraded lesson observations had helped to raise the standards of teaching and learning in the workplace and to improve the classroom practice of staff. Just below half of respondents (44.7% and 48.7%) agreed with both of these statements, compared to a quarter who disagreed. The percentage of 'N/A' responses was noticeable,

approximately a quarter of responses for both statements. The main explanation for such a high level of 'N/A' responses is likely to be that ungraded models of observation were not in use in the workplaces of these respondents, thus they were unable to express an opinion on their impact on practice. Finally, it was interesting to note that over three quarters of respondents (76.6%) agreed that ungraded lesson observations should replace graded observations, with SMTs again proving the exception. Furthermore, the high levels of agreement recorded regarding the use of ungraded models of observation were echoed in the qualitative data, as the two excerpts below serve to illustrate.

I believe in observations between respected and supportive peers, who know the teaching area, who know the issues amongst the students, and who would provide constructive and supportive feedback to encourage improvement. If we have to have formal observations by managers, then they also need to be properly conducted, with knowledge as above and most definitely ungraded, with constructive feedback (Sarah, curriculum coordinator)

The more frequently ungraded observations occur (within reason!), and the more constructively the ensuing feedback is given, the more a culture of embracing peer observation as a tool for professional development without the element of suspicion and fear develops. This can only be a positive thing for students, teachers and colleges (Respondent 905)

Much of the project's qualitative data converged around the counterproductive effects of observation, highlighting the predominant perception among many participants that the use of observation in FE was deemed problematic rather than productive. With regards to the matter of improving standards of teaching via formal observation schemes, for example, there was a groundswell of opinion

among the study's participants that the impact of such practice was at best negligible and at worst detrimental to the professionalism of practitioners.

One of the main findings to emerge from the project was the widespread discontent felt amongst participants towards the use of lesson observation as a form of teacher assessment. Graded models of observation were repeatedly criticised by a significant majority of participants for being little more than a 'box-ticking' exercise and, in some instances, a 'disciplinary stick' with which 'to beat staff'. They were also identified as a major cause of increased levels of stress and anxiety amongst teaching staff. Whilst there were significant differences in interpretations as to the value attached to graded observations by practitioners and SMTs, this disparity was illustrative of how, for many observees, the outcomes they experienced did not equate with the purported aims of those who were responsible for overseeing and implementing them i.e. senior managers and observers. As far as practitioners were concerned, performance management models of observation were of little relevance to their professional needs and failed to improve their practice. For them, such models existed purely to furnish SMTs and external agencies like Ofsted with quantifiable data that could be used to exercise managerial control over judgements about the effectiveness of teaching and learning. In contrast, there was a consensus amongst participants that low-stakes, peer-based models of observation were most conducive to sustainable change and professional learning and thus should be at the forefront of most providers' use of observation and wider CPD strategy. However, the importance of these models seemed undermined by the on-going external demand for statistical performance data.

Implications for the future use of observation in education

The project's key findings revealed an appetite for change among the majority of participants as to how observation is currently used in FE, as well as the ways in which teacher performance is assessed and managed. The first recommendation to

emerge was the need for observation to be exploited more as a formative, supportive intervention with the emphasis on helping teachers to improve their practice rather than the current deficit model that focused mainly on attempting to measure and judge it in an isolated, episodic manner. Thus the professional needs of practitioners should be placed at the forefront of any observation scheme rather than the production of performance management data. Providing both observees and observers with opportunities to engage in substantive professional dialogue as part of the observation process is vital, though if this is going to work in practice then sufficient time needs to be allocated and embedded into teachers' workloads at the start of each academic year. This time allocation needs to allow for: 1) a pre-observation meeting; 2) feedback and 3) feed-forward meetings as well as the observation of the lesson itself as a minimum commitment.

Following on from this, the second recommendation identified the need for a shift towards more peer-based models of observation. While only a small minority of participants expressed a desire to see an end to the use of lesson observation per se as a form of teacher appraisal, the majority acknowledged that it had an important role to play in teacher assessment and development. They did so, however, on the proviso that certain models were deemed to be more beneficial than others, particularly peer-based models with a focus on enhancing professional learning. These have the potential to redress some of the power imbalances associated with top-down, deficit models and encourage a greater sharing of practice and dialogue that could be mutually beneficial for observer and observee. This is perhaps best encapsulated in the comments of one of the project's participants:

The best model has to be one based on a genuine spirit of enquiry and research. To explore what's happening in that messy business of learning, and to be a starting point for professional discussion and debate. This means that both people involved in the discussion - teacher and observer need to be equal

partners in the process, both working to improve things for teachers and learners. This means sharing a common purpose - why are we doing this and what do we both need to get out of it? (Vera, observer)

Thirdly, there was an overwhelming desire to move away from a system that relies heavily on a narrow evidence base to a more fully-inclusive, multi-dimensional model of assessing teaching competence and performance. The current reliance on annual graded observations as a means of measuring a practitioner's professional competence was considered an inequitable and reductive practice. Given the misgivings surrounding the validity and reliability of observation as a method of assessment discussed at length in the project's report, there can be no justification in linking the outcome of formal observations directly to an institution's capability or disciplinary policy. The project's findings highlighted an increasing demand to make the process of teacher evaluation more inclusive by extending it beyond the lens of lesson observation and drawing on other sources of evidence (e.g. student feedback, student achievement data, peer review, self-assessment, external verification) so as to ensure a more triangulated and reliable evidence base for assessment. How such different data sources might be harnessed into a coherent framework of assessment undoubtedly represents a challenge, but it is a challenge worth tackling for all those committed to the ongoing improvement of teaching and learning.

Fourthly, there is a need to explore alternative approaches to the current, dominant model of graded lesson observation. These alternatives should seek to draw on elements of existing practices but also make greater use of recent advances in the research of observation and cognate areas that lend themselves to exploiting the potential of observation as a medium for professional enquiry and learning. Such alternative approaches might therefore wish to draw on models like 'differentiated observation' or 'lesson study' as a starting point (see, for example, O'Leary 2014, Chapter 8, for an account of alternative models of observation).

Concluding comments

For any of the recommendations discussed above to lead to meaningful and sustained change, this requires a root and branch reform of the way in which observation is conceptualised and engaged with as a form of educational intervention. As a starting point, educational leaders, inspectors, policy makers and practitioners alike need to break free from the assessment straitjacket that currently constrains the perception and implementation of observation. In my view, this is arguably the biggest obstacle that prevents the English education system from harnessing the potential of observation as a tool for meaningful and sustained professional learning and development.

However, in order for this to happen, there needs to be a ‘thinking outside the box’ when it comes to how observation may be used as a source of evidence in the educational arena. Tinkering with prevailing normalised models of observation is, at best, only likely to have minimal impact and offer short-term solutions to longstanding issues. Although removing the graded element would certainly represent a step in the right direction, for example, it cannot be considered a panacea in itself. In a similar vein, recent calls for the abolition of lesson observation from the inspection process are a classic example of ‘throwing the baby out with the bath water’ and as such represent a knee-jerk reaction to a much more complex problem than the one they claim to solve. Ultimately, what both of these strategies fail to address are the deep-rooted political and epistemological issues surrounding the use of observation as a method of assessment. At the heart of any such discussion is the acceptance that the use of observation is not purely an act of pedagogy but one that is underpinned by issues of hierarchical power and professional trust. Until these issues are acknowledged and discussed by education professionals in an open forum, then any attempts at reforming the way in which the sector makes use of observation are unlikely to progress.

References

- Ball, S. J.** (2003) The teacher's soul and the terrors of performativity. *Journal of Education Policy*, Vol. 18(2) pp. 215-228.
- Boocock, A.** (2013) Observation of teaching and learning: teacher development or micropolitics and neo-Fordism. *Journal of Further and Higher Education*, Vol. 37(4), pp. 482-503.
- Green, J.** (2011) *Education, Professionalism and the Quest for Accountability*. New York: Routledge.
- Huddleston, P. & Unwin, L.** (2013) *Teaching and Learning in Further Education – Diversity and change*, 4th Edition. London: Routledge.
- Ofsted** (2012) *A good education for all*. London: Ofsted Publications Centre.
- O'Leary, M.** (2013) Surveillance, performativity and normalised practice: the use and impact of graded lesson observations in Further Education Colleges, *Journal of Further and Higher Education*, Vol. 37(5), pp. 694-714.
- O'Leary, M.** (2014) *Classroom observation: a guide to the effective observation of teaching and learning*. London: Routledge.
- O'Leary, M. & Gewessler, A.** (2014) 'Changing the culture: beyond graded lesson observations'. *Adults Learning – Spring 2014*, Vol. 25, pp. 38-41. Available online at: <http://www.niace.org.uk/sites/default/files/documents/adults-learning/2014-spring/AL-Spring-2014-Vol25-pg38-41.pdf>
- Smith, R. & O'Leary, M.** (2013) New Public Management in an Age of Austerity: Knowledge and Experience in Further Education, *Journal of Educational Administration and History*, Vol. 45(3), pp. 244-266.
- Stevenson, H. & Wood, P.** (2013) Markets, managerialism and teachers' work: the invisible hand of high stakes testing in England. *The International Education Journal: Comparative Perspectives*, Vol. 12(1), pp.42–61.
- University and College Union (UCU)** (2009) *Lesson Observation: UCU Guidelines*. [Accessed 24/06/2013] Available at: <http://www.ucu.org.uk/index.cfm?articleid=2969>.

University and College Union (UCU) (2012) *UCU FE News – May 2012*. [Accessed 19/06/2012] Available via: <http://www.ucu.org.uk/>.

University and College Union (UCU) (2013) *Developing a National Framework for the Effective Use of Lesson Observation in Further Education*, project report completed November 2013.

¹ This paper was first presented in a modified form at the ResearchED Midlands Conference, Birmingham, UK on 5th April 2014.