



Matthews, Richard and Noyes, Andrew (2016) To grade or not to grade: balancing formative and summative assessment in post-16 teacher trainee observations. *Journal of Further and Higher Education*, 40 (2). pp. 247-261. ISSN 1469-9486

**Access from the University of Nottingham repository:**

<http://eprints.nottingham.ac.uk/32027/1/Matthews%20%26%20Noyes%202015%20post-print.pdf>

**Copyright and reuse:**

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

- Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners.
- To the extent reasonable and practicable the material made available in Nottingham ePrints has been checked for eligibility before being made available.
- Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
- Quotations or similar reproductions must be sufficiently acknowledged.

Please see our full end user licence at:

[http://eprints.nottingham.ac.uk/end\\_user\\_agreement.pdf](http://eprints.nottingham.ac.uk/end_user_agreement.pdf)

**A note on versions:**

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact [eprints@nottingham.ac.uk](mailto:eprints@nottingham.ac.uk)

## To grade or not to grade: balancing formative and summative assessment in post-16 teacher trainee observations

Richard Matthews, University of Nottingham<sup>1</sup>

Andrew Noyes, University of Nottingham

### **Abstract**

The issue of whether trainee teachers in the post-16 sector should have their classroom practice graded has been debated for a number of years. The case for training courses retaining an emphasis on written and verbal 'developmental' feedback at the expense of 'judgements' appears to be lost. This paper is set within the context of an ever growing culture of performativity in English Further Education Colleges, where grading is regarded as an essential requirement to ensure high quality teaching. Tensions are explored between stakeholders who call for graded observations of trainees' classroom performance (e.g. Ofsted and FEC quality assurance managers), and classroom-based trainers and researchers who argue that grading is too judgemental and compromises the formative and developmental progress of trainees. The rationale for trainee teachers to have their classroom practice graded is contrasted with evidence that highlights the negative results of grading. The paper reports findings from the evaluation of an innovative, alternative strategy that addresses Ofsted's central requirement for trainees to know 'where they are' in their development by offering a middle way between grading and not grading trainees' classroom performance.

### **Introduction**

There is general agreement that the concept of teacher professionalism in the FE sector has changed considerably and for the worse over the last two decades (e.g. Avis, 2003, 2005; Gleeson et al, 2005). It has been argued that a culture of management-led performativity has had a negative impact on teacher identity and resulted in reduced autonomy and professional empowerment (Sachs, 2000, 2001; Ball, 2003, 2004; Hargreaves and Fullan, 2012).

A central thesis of these critiques is that the judgemental approach to assessing teachers' classroom performance is linked to audit trails, performance management, and other types of accountability and that these undermine the processes and development of reflective practice (Gosling, 2005). The observation of classroom practice has become "normalised as a performative tool of managerialist systems designed to ensure and

---

<sup>1</sup> Corresponding author: School of Education, Jubilee Campus, the University of Nottingham NG8 1BB. E-mail: [ttxrim@nottingham.ac.uk](mailto:ttxrim@nottingham.ac.uk)

improve standards, performance and accountability in teaching and learning" (O'Leary, 2012: 1). Research suggests that this process causes anxiety and loss of self-esteem amongst teachers, including trainees (Williams, 1989; Norrish, 1996; Lee, 2007). Trainee teachers, many of whom would be undertaking in-service training and therefore have considerable experience of teaching, feel defensive as a consequence of graded observations. They feel that they receive little subsequent developmental support in their classroom performance (Cosh, 1998).

The last two decades have seen considerable changes in what it means to be a teacher professional. A dichotomy has developed between the managerial demands for accurate quality assurance measurements and teachers' requirements for supportive quality improvement (Hargreaves and Fullan, 2012). Hargreaves and Fullan, (see also Sachs (2000) and Avis (2003)) argue that what teachers require to become professionals is the capacity to improve their classroom practice and to collaborate in developing teaching and learning strategies with fellow teachers. The current managerialist strategies, they argue, have failed to achieve an enhanced professionalism. If anything, the strategies have resulted in teaching becoming de-professionalised with teachers conforming to the observation requirements. What is needed for an active and dynamic teaching profession is for teachers to develop their confidence and expertise based on a culture that encourages empowerment and autonomy.

The problem that has emerged is how to develop reflective, autonomous professionals in a culture that emphasises 'measurement' of teaching performance rather than 'development'. Different solutions have been suggested. On the one hand O'Leary (2012) has called for a moratorium on grading classroom observations due to its negative impact on developing teacher professionalism, whilst Bousted and Hobby(2012) have acknowledged that although 'bad observation' exists, that is no excuse for not having an observation process that uses measurement (i.e. grading) in order to raise teacher 'performance'. The tension between managerialism and quality assurance on the one hand and professional development, reflection and personal development on the other can be seen in many sectors of education. We now turn to focus on one such context: teacher training in the post-16 sector.

During the past decade there has been a growing debate about whether trainee teachers (hereafter known as trainees) in the post-16 sector in England should have observations of their teaching practice graded using criteria<sup>2</sup> established by Ofsted. We will explore

---

<sup>2</sup>This paper refers to criteria used by Ofsted for inspections of teacher training in the Learning and Skills Sector from 2009 up until September 2012, at which point new criteria were implemented.

this debate in two parts. Firstly, we contextualise the pressure to grade trainee observations in the government's educational reform agenda. We examine the reasons why the majority of post-16 teacher training consortia in England have moved from a position of giving purely formative feedback and not grading their trainees' classroom practice, to a position of grading performance. The paper also discusses why some post-16 teacher training consortia have resisted the pressure to grade.

Secondly, this paper reports findings from an action research project that aimed to develop an alternative strategy to O'Leary's (2012) proposed moratorium on graded observations. This alternative approach addresses the tensions that have emerged in policy, practice and research. It accepts the political constraints of the educational world that require 'judgements' on performance, whilst at the same time emphasising teacher 'development' through enhanced autonomy and empowerment. The alternative strategy continues to allow managers to obtain data for monitoring quality assurance through graded observations, but affords trainees the opportunity to use observation feedback from their employers in a way that enhances their professional development. The research also supports the case for tutors and mentors to not grade trainees as part of their ITT course

### **The case for the grading of trainee teachers' classroom practice**

Ofsted became responsible for inspecting teacher training in the FE sector in 2001. In its first significant review of Initial Teacher Training (ITT) provision Ofsted (2003) voiced serious criticisms about the quality of training being offered. Courses lacked a "satisfactory foundation of professional development for FE teachers" (2). In particular there were criticisms made of the observation process. Tutors' judgements of trainees' teaching practice were judged to be weak, with a lack of clarity about what level trainees needed to achieve to meet the requirements of the teaching standards. Ofsted also highlighted the requirement for trainees to 'know where they were' in their professional development. Ofsted recommended that there should be more involvement of Human Resource (HR) managers in the management of ITT courses, with more integration of "the initial training of teachers with other aspects of the management of their staff" (Ofsted, 2003:2).

In its subsequent review of ITT in the FE sector between 2004 and 2008, Ofsted (2009) continued to criticise the poor judgements being made of trainees' progress and the limited contribution made by HR managers to forming those judgements. Observation of

classroom practice was described as being “one of the weaker aspects of provision provided” (5). Once again, Ofsted called for a stronger link between ITT and HR processes, suggesting that ITT tutors be part of the internal observation process for quality assurance. As a result, the guidelines for grading ITT courses in the 2008-2011 cycle of inspections focused specifically on the outcomes of the observation of trainees’ classroom practice. The Ofsted guidance did not specifically instruct HEI consortia to grade trainees. However, there was a clear assumption that tutors would make Ofsted-style judgements on their trainees. As part of the inspection process, inspectors would observe the complete process of tutors observing their trainees and then giving feedback. The inspectors would then determine the accuracy of the tutors’ judgements. A consequence of this increased focus on making Ofsted-style judgements on trainees’ classroom performance resulted in the majority of HEI providers implementing some form of grading policy.

### **The impact of grading observations on quality assurance processes in post-16 institutions**

FE colleges’ quality assurance processes typically include annual observation and grading of teachers’ practice. This process is normally linked to performance management or appraisal meetings where development issues can be raised with individual teachers to assist in the process of continuing professional development (Lee, 2007). With the data that is collected, institutions are able to self-evaluate the quality of teaching and learning on the basis of teachers’ grades. Whether having this data leads to improved practice is a moot point.

As many of our (in-service) trainees are members of the teaching staff, they too are observed under this quality assurance process and receive a formal grade each year. Trainees experience being ‘trainee as learner’ with their tutor as well as ‘trainee as teacher’ with their line manager. Ollin’s (2009a) research reported that HR managers regarded the observation of trainees to be part of the process of “driving up standards, driving up success rates [and] aspiring to excellence” (22). However, although the HR managers regarded the role of tutors as developmental when observing their trainees, “they also implied responsibility on the part of these tutors to assure standards of teaching” (22) against the Ofsted criteria, i.e. trainees were regarded as *teachers* rather than *learners* and tutors were expected to make summative rather than formative comments. While this process might be acceptable for trainees who have been teaching for several years, no allowance is made for new in-service trainees who have limited

experience in the classroom. There is a conflict here between the concerns of the HR manager seeking “good” grades and the ITT tutors privileging formative feedback with developmental goals and avoiding the use of summative grades in keeping with the formative assessment research (Black and Wiliam, 1998).

Ofsted has argued that the formative, developmental approach favoured by many ITT tutors (who regard trainees as primarily *learners* rather than *teachers*) hinders trainees knowing ‘where they are’. In contrast, Ofsted argues that the inspection-style process within post-16 institutions is a positive development allowing standards to “be applied in a quality assurance framework to the mutual benefit of all parties...” (Harkin et al, 2003: 41). Both Ofsted and HR managers’ primary concern is to make definitive judgements and use these to monitor, self-evaluate and drive up standards, although the mechanism by which measurement leads to improvement is unclear. In 2006 Ofsted reported that no trainees had demonstrated ‘very good’ or ‘outstanding’ practice, thereby concluding that the most capable trainees were not achieving their full potential. By 2009, Ofsted reported that approximately 10% of trainees were judged to be ‘outstanding’. Consequently, Ofsted and HR managers perceive that graded observations result in developing good teaching.

### **The case for not grading trainee teachers’ observations**

Tutors and their trainees are often resistant to grading because of the ways in which it undermines formative, developmental processes. Rather than being judged, trainees want help with their classroom performance; “they want to observe and be observed, with feedback” (Harkin et al, 2003:39). The growth of managerialism, inspection and self-evaluation in the FE sector during the last decade has been described as “surveillance...[and] ...another coercive and malign instrument designed to hold [teachers] to account” (Cockburn, 2005: 47, see also Hall and Noyes, 2009).

The perception that Ofsted-style inspection, and subsequent grading of performance, leads to improved classroom practice is questionable. As Thomson et al (2010) argue, the process of putting people into simple numbered categories leads to unhealthy modes of “calibration and forensic dissection” (652). Coffield and Williamson (2011) echo this concern in their criticism of how government policy on grading teachers’ classroom performance is seen as a means of addressing underperformance. Coffield and Williamson argue that rather than leading to improvement, graded observation has *caused* underperformance. As teachers, including trainees, become ‘mark hungry’, they

are pressurised to perform to a particular paradigm. Their learners subsequently react in ways to meet the requirements of the paradigm at the expense of real and meaningful learning. As a consequence, teaching and learning becomes distorted, contrived and artificial (Cockburn, 2005; O'Leary, 2012).

These are serious criticisms of the current emphasis on grading. Teachers can find themselves performing particular practices in a tick box culture which tends to trivialise the complex art-craft-science of teaching (Cockburn, 2005; Nasta, 2007; O'Leary, 2012). So it is that Ofsted-style observations can constrain trainees to prepare a particular type of officially sanctioned lesson (Lee, 2007; Thomson et al, 2010); teaching can become contrived and artificial and risk-taking is at best reduced, and at worst minimal (Cockburn, 2005; O'Leary, 2006; Orr and Simmons, 2010). Understandably, trainees adopt 'safe' approaches to teaching that meet a limited range of requirements, do not identify weaknesses for development and fail to experiment with new classroom approaches (Peake, 2006). At the outset of Ofsted's existence, Wragg (1994) argued that there should be less of an emphasis on "cosmetic changes...to please the observer" (96) and this holds true two decades on.

The neo-liberal reform agenda has given greater power to FE quality assurance processes as they seek to produce high performing, productive workers. O'Leary (2012) explains the implications that this has for the relationships between the observer and teacher, particularly when Ofsted's criteria now determine what can be considered 'normal': "Those that are able to manifest such normalised behaviour form a homogenous community; those that fail to do so are identified through 'gaps' in their assessed performance" (O'Leary, 2012: 14). This process does little to enhance teacher autonomy but rather tends to disempower teachers. Instead of adopting a deficit model, which pre-supposes that there is something missing in the trainees' performances, there needs to be an approach that motivates and develops that self-confidence that can lead to professional growth (Shortland, 2004; Cockburn, 2005; O'Leary, 2006). Observation should be a positive, developmental experience with trainees being freed from the fear of harsh criticism and the excessive anxiety that can arise from being graded. Cockburn (2007) argues that trainees need to be supported to become reflective and develop their own action research as part of their continuing professional development. Trainees who collaborate with their tutor to plan for a developmental classroom activity can reflect on the event and make subsequent improvements.

The deficit, grading approach has not achieved what it was intended to do and the potential for classroom observation to be used as a tool for professional development

has got lost. For example, O’Leary (2012) reported that many observers found it very difficult to provide constructive advice following lesson observations as they were driven by the demands of the quality assurance system. Observers found themselves assuming a judgemental role and marginalising the developmental potential of the observation process. O’Leary explains that this change is not only having an impact upon the trainees: “alongside these policy developments there has been a simultaneous reduction in the autonomy of tutors, with limited opportunity for them to shape and influence their professional development and identity” (25).

There is general agreement amongst scholars that a model of observation which does not involve grading is preferable to the approach used by Ofsted and HR managers (Cosh, 1998; Cockburn, 2005; Peake, 2006). It is the importance of formative feedback to trainees that is central to any discussion about whether grading should take place or not. Such formative feedback plays a crucial role in the development of trainees (Hardman, 2007). Peake (2006) has shown that trainees who have experienced both quality assurance graded and non-graded ITT observations have found the ITT experience more progressive and developmental to their professional practice. They also felt that regular non-graded observations which enabled action points from one observation to be implemented for a subsequent observation had more impact on practice than did an annual, graded quality assurance observation, where development points were not revisited (Fawbett, 2003).

### **Alternative strategies for the observation process of post-16 trainees**

In spite of the above discussion of the demerits of grading, there is a clear requirement for tutors to be able to make judgements against the Ofsted criteria regarding trainee progress and for the trainees to know ‘where they are’ in relation to the official standards. Indeed, inspectors have repeatedly asked HEI institutions that do not explicitly grade the classroom practice of trainees how they know if trainees have achieved the required standards. In particular, do tutors and trainees know if they are ‘outstanding’ or have the ‘potential to be outstanding’?

A decade ago a handful of the forty HEI consortia that deliver post-16 ITT in England graded the teaching practice of their trainees. At the time of writing over 50% now have some form of grading using Ofsted criteria. The changes appear to have been made as a direct result of Ofsted inspectors challenging HEIs during their inspections. The following



contrasting approaches illustrate how and why ITT partnerships have reacted in different ways to the pressure to grade trainees' classroom performance.

'University of the North' operates in partnership with FE colleges delivering post-16 ITT in the north of England. Up until 2009 the HEI and its partnership colleges were committed to a policy of not grading their trainees' classroom practice. They followed a developmental approach through formative feedback. They believed, for the reasons given earlier in this paper, that this strategy reduced trainee stress and supported trainees in taking risks and in developing self-reflection. It was believed that this process led to improved teaching and learning. Tutors argued that grading, which is essentially a judgemental process, would undermine the advantages of a developmental approach. Alongside this process trainees would experience graded observations organised by their institutions' HR managers to satisfy quality assurance requirements.

However, it was recognised by tutors that there were different perceptions of the process and content of feedback between Ofsted graded lessons and those focused on developmental feedback. Tutors realised that their trainees were not confident in understanding the coded language used in feedback and often could not understand the difference between a graded quality assurance observation and a formative ITT observation. It was felt that trainees did not know where they stood in regard to their classroom performance particularly at the end of a two year training course (Ollin, 2009a and 2009b). This HEI consortium took the view that some form of grading was required as part of the teacher training process. The argument put forward to justify the consortium making this decision was that

both tutors and trainees [need to be] clear about expected standards of teaching, including the possibility that a trainee could fail to meet these standards...[and]... that tutors and trainees have a shared understanding of these standards the trainee has achieved, whilst keeping the main forms of observation as developmental and formative (Ollin, 2009a: 58).

Consequently this consortium has now implemented an observation process whereby most tutor observations have remained as developmental, with no grading. But two formal grading points have been established, one at the end of Year 1 which indicates a 'potential' grade against the Ofsted criteria, and one at the end of Year 2 which indicates a summative grade. Tutors can, therefore, continue to emphasise the developmental aspect of the observation process whilst trainees gain some idea of 'where they are' in Ofsted terms. Trainees will also be able to compare the grade given by their tutor with grades given as part of the quality assurance process in their institution.

Compare this with the 'University of the Midlands', which operates in partnership with FE colleges delivering post-16 ITT in the Midlands of England. Like the University of the North it has traditionally been committed to conducting non-graded observations. Although trainees were graded by the institutions where they were employed as part of the quality assurance processes, tutors have continued to feel that graded observations for ITT undermine the developmental nature of the training process. This partnership has recently been involved in two Ofsted inspections within 18 months of one other. On both occasions inspectors wanted to know why trainees' classroom practice were not graded. The inspectors accepted the position of the partnership that although the observations were developmental, tutors were able to make accurate judgements against the Ofsted criteria.

The tutors in this partnership have repeatedly debated the issue of whether or not to grade their trainees and have recently confirmed their commitment to continue the policy of not grading. They felt that they could make necessary and accurate judgements whilst giving developmental feedback. However, three problems still exist with this approach. Firstly, trainees are not necessarily aware of the judgements that are made by their tutors and consequently cannot self-evaluate their progress against the Ofsted criteria. Secondly, the trainees still receive a grade as part of their institutions' quality assurance process; some students have commented on the incompatibility with their grades when compared to their tutors' feedback. Thirdly, employers and future employers have no indication of their trainees' professional competence as identified by an HEI at the end of the course.

### **A middle way: empowering trainee teachers in their professional development**

So far we have discussed the pros and cons of grading classroom practice. Whatever position one takes there is a *realpolitik* of Ofsted inspections and quality assurance processes. Indeed, as part of the professionalization process ITT tutors need to help trainees understand how to deal with such realities. O'Leary's (2012) call for a moratorium on graded observations is unrealistic. The world in which we now live is such that managers and future employers need common measures of professional competence. It is, therefore, manifestly unlikely that the production of quantitative measurement of trainees' classroom performance is going to cease particularly with the majority of ITT consortia having now implemented some form of grading. Elsewhere, Boustead and Hobby's belief (2012) that the observation process would improve once 'bad' observations had been weeded out fails to address the fundamental tension

between a summative, judgemental assessment process and a formative, developmental one. What is needed is a middle way that recognises the requirements of managers to obtain quantitative data on classroom practice whilst enabling developmental opportunities for trainees and avoiding the need for tutors and mentors to grade trainee observations.

Such a middle way has been trialled in a small action research study at the FE College in which one of the authors lectures and which we call Local College. This research resulted from a desire to seek an alternative to the opposing pressures of grading trainees' professional practice or not. As proponents of action research explain (for example Hopkins, 1993) following the specification of a particular practice-based problem (here discussed in the previous sections), the action researcher develops an intervention, implements and then evaluates this, thereby leading into a further research cycle and so on.

This action research study investigated the trainees' experiences of the employers' observation process (where a grade was given), compared this with their tutors' and mentors' observations (where no grade was given) and then trialled a new strategy. A PGCE class of seventeen trainees was involved in the study at Local College. The study was undertaken towards the end of term 1 of the trainees' second year of their PGCE course. The course is part-time, with trainees attending Local College for one day a week. Most trainees have teaching hours in Local College but some work in other local post-16 education institutions. During the first year of the course, their classroom practice had been observed twice by a tutor and twice by their mentor using the observation protocol of the University of the Midlands. Observations were ungraded and formative, with action points developed that had been negotiated between observer and trainee. During their first year of the PGCE course the trainees had also been observed, at least once, against Ofsted's criteria as part of their employers' quality assurance procedures; the trainees' received a grade on their performance in these 'quality' observations and received feedback that was of a summative nature.

The research was carried out with all seventeen trainees in the class acting as a focus group. In the initial discussion meeting, the group reflected upon their experiences of the two approaches to observations. The feedback from the Ofsted-style observations typically consisted of a grade and a list of strengths and weaknesses. The trainees were not averse to being graded; they understood the requirement to be assessed and welcomed the judgement of 'where they were' in their classroom performance. However, the trainees were critical of the feedback they received. The observers had not discussed

any strategies to help them address the identified weaknesses. No link was made between the observation process and the institutions' appraisal systems where a developmental action plan could have been developed. The trainees reported that they felt that the process was too judgemental and used by managers to primarily gather quantitative data on performance, rather than provide opportunities for professional development. They also expressed a sense of powerlessness in this judgemental process feeling that they were unable to take part in a discussion about the feedback. They felt that these types of observations were something that was 'done' to them rather than being supportive and collaborative.

The trainees gave a contrasting view of the 4-6 observations undertaken as part of their PGCE course. These had not been graded and no summative judgements had been made. The trainees reported that following each of these observations developmental points had been discussed with the observer, targets were agreed and became a focus for subsequent observations. Trainees commented that they found this process very helpful in their professional development and that the discussion with the observer enabled them to feel part of the process. In contrast to the other observations, they did not feel that observations were being 'done' to them and that they felt a sense of personal and professional autonomy.

Following the discussion, a new approach was developed and trialled with the aim of making the Ofsted-style observations as empowering and as useful for the trainees' professional development as their PGCE observations had been. During a PGCE session the trainees each collated the positive points which were made by their tutors and mentors, together with the actions points that had been negotiated for future development. They compared and contrasted these outcomes against the Ofsted criteria and against feedback from the Ofsted-style observation they had received in Year 1. The trainees then awarded themselves self-assessment grades for their PGCE observations and identified their own action plan for developing their classroom practice. The self-assessment grades were not for public use but to inform personal professional development.

Following this process each trainee met with their personal tutor where the focus was on classroom practice with reference to the Ofsted criteria. Each trainee was asked to report their self-assessment grade and explain it using the Ofsted criteria. For all of the trainees, there was close agreement between the self-assessment of progress and their tutor's judgement. This indicated that trainees were able to understand the Ofsted criteria and accurately identify their strengths and areas for development.

At the end of this process the trainees reported that:

1. they felt that they had a better understanding of the Ofsted criteria than when they had been previously observed as part of their employers' quality assurance procedures
2. their self-confidence had increased as they felt able to identify a summative Ofsted grade, thus fulfilling Ofsted's requirement of 'knowing where they are'
3. they were less apprehensive of Ofsted-style grading and had a sense of empowerment. Being observed using Ofsted criteria would not deter them from experimenting with teaching strategies. They believed that post-observation discussions should be more dialogic and that they would be able to confidently justify their decisions and actions.

The first action research cycle suggests that it is possible to overcome some of the negative aspects of grading whilst enabling trainees to have some indication of their position when measured against the Ofsted criteria. The strategy seeks to preserve the developmental role of the ITT tutor whilst allowing the trainees to use their tutors' and mentors' feedback to reflect on 'where they are' against the Ofsted criteria and compare this to feedback received from quality assurance graded observations.

An important element in this strategy is to empower the trainees and to develop a sense of professional autonomy. Trainees are able to develop an understanding that Ofsted grading criteria is not the only way to assess how far a lesson is 'outstanding' or 'good'. By encouraging trainees to be critically aware of the Ofsted criteria they can trial more high-risk, developmental lessons with the confidence to justify their actions.

The trainees' comments offer a challenge to the mainstream orthodoxy outlined earlier in this paper that is deeply critical of the impact of grading trainees' professional practice. The evidence from this small study suggests that trainees do see some merit in having their observed teaching practice graded, as part of their institution's quality procedures, and are not therefore completely averse to complying with Ofsted. What they want is to be involved as equal partners – as professionals – in the observation and assessment process, ensuring that graded observations are supportive and developmental. What particularly emerged was the trainees' requirement for more than just receiving developmental feedback, emphasised by several writers above, but the opportunity to discuss both the strengths and weaknesses of their practice and to have some autonomy in being able to task risks.

## **Concluding Comments**

In this paper we have discussed the tensions for ITT trainees between receiving summative/graded feedback and formative/developmental feedback from formal observations of practice. Building on this we have presented some evidence from action research conducted in an FE college that shows that both of these demands can be satisfied via a middle way involving trainee self-assessment. This middle way has five advantages over both the call for formal grading and those resisting change.

1. It fulfils Ofsted's desire that trainees should 'know where they are', enabling a more accurate, consistent and transparent approach. The evidence from this small study suggests that trainees' self-assessments are highly likely to be the same as that of an official judgement. More importantly, the trainees' self-evaluation tends to be more holistic, drawing upon a range of lessons.
2. This approach supports many of the aspects of educational 'good practice' highlighted in this paper. It avoids the excesses of the surveillance aspect of grading whilst emphasising the trainees' self-reflection and development. This is suitable for trainees new to teaching and those who have some experience, as well as those teaching in different contexts.
3. This approach helps trainees to compare their own perceived Ofsted grade with grades received for quality assurance purposes. Having reflected on where they think they are in relation to the Ofsted criteria, they should be able to undertake a professional discussion with their observer (either their ITT tutor, or whoever has undertaken the quality assurance observation) should there be a marked discrepancy between the two grades. This is a professional competence that we consider is very important.
4. Negotiations between tutors and trainees of the outcomes of their ungraded observations could lead to a summative grade being determined towards the end of the PGCE course. These could be used to inform current and future employers.
5. Finally, and perhaps most importantly, this small scale study demonstrated the extent to which empowering the trainees by giving them the autonomy and confidence to develop challenging lessons enhanced rather than limited their professional development.

This approach balances the strengths/demands of grading and not grading observations by occupying the middle-ground and in doing so makes a strong contribution to the professional development of trainees in Further Education Colleges. It allows HR

managers to continue obtaining quantitative data they deem necessary to judge the quality of teaching and learning and which feeds into the statistical reporting that is undertaken across an institution. Importantly, this approach also retains the developmental observations by tutors and mentors, thereby giving trainees strong formative feedback that can become a tool that enhances their capabilities to a high standard. Whilst recognising that the managerial processes are not going to go away any time soon, professional learning and the development of ITT processes can be developed enabling trainees to develop reflection, self-determination and strong professional identities.

Working in this way enables trainees to understand the demands of external agencies such as Ofsted without reducing their practice to the unidimensionality of that particular grading system. This approach suggests an alternative strategy for enhancing professional learning. It enables trainees to develop an understanding the rules of accountability frameworks whilst at the same time enabling them to develop a measure of independence.

There would be some benefit of evaluating these models in a more systematic way in future. But for now we have presented this issue as an example of how current performative and managerial cultures in teacher training in Further Education Colleges can be embraced enabling the trainee 'voice' to be heard and for the trainees' development to be enhanced through informed discussion.

## References

Avis, J. (2003) Re-thinking trust in a performative culture: the case of education. *Journal of Education Policy*, 18 (3), 315-332.

Avis, J. (2005) Beyond Performativity: reflections on activist professionalism and the labour process in further education. *Journal of Education Policy*, 20 (2), 209-222.

Ball, S. (2003) The teacher's soul and the terrors of performativity. *The Journal of Education Policy*, 18 (2), 215-228.

Ball, S. (2004) Education reform as social Barberism: economism and the end of authenticity. *Scottish Educational Review*, 37 (1), 4-16.

Bousted, M., and Hobby, R. (2012) This is your profession – so take charge of it. *The Times Educational Supplement*. 12<sup>th</sup> October. London.

Cockburn, J.(2005) Perspectives and Politics of Classroom Observation. *The Research and Development Bulletin*, 3 (2), 45-56.

Coffield, F. and Edward, S. (2009) Rolling out 'good', 'best', and 'excellent' practice. What next? Perfect practice? *British Educational Research Journal*, 35 (3), 371-390.

Coffield, F. and Williamson, B. (2011) *From Exam Factories to Communities of Discovery. The democratic route.* (London, Institute of Education, University of London).

Cosh, J. (1998) Peer Observation in Higher Education: a reflective approach. *Innovations in Education and Training International*, 35 (2), 171-176.

Fawbert, F. (Ed) (2003) *Teaching in post-compulsory education: learning skills and standards.* (London, Continuum).

Gleeson, D., Davies, J., & Wheeler, E. (2005) On the making and taking of professionalism in the further education workplace. *British Journal of Sociology of Education*, 26 (4), 445-460.

Gosling, D. (2005) Peer observation of teaching. *Seda paper 118.* London: Staff and Educational Development Association Ltd.

Hall, C. and Noyes, A. (2009) New regimes of truth: the impact of performative school self evaluation systems on teachers' professional identities. *Teaching and Teacher Education*, 25 (6), 850-856

Hardman, J. (2007) *The Use of Teaching Observations in Higher Education.* ESCalate: Education Subject Centre Advancing Learning and Teaching.

Hargreaves, A & Fullan, M. (2012) *Professional Capital. Transforming Teaching in Every School.* (Abingdon, Routledge).

Harkin, J., Clow, R. & Hillier, Y. (2003) *Recollected in Tranquillity – FE teachers' perceptions of their Initial Teacher Training.* (London, Learning and Skills Development Agency).

Hopkins, D. (1993) *A Teacher's Guide to Classroom Research.* 2<sup>nd</sup> Edition. (Buckingham, Open University Press).

Lee, J. (2007) Are you being observed? In *The Times Educational Supplement, Further Education*: page 1, 15<sup>th</sup> June. London.

Nasta, T. (2007) Translating national standards into practice for the initial training of Further Education (FE) teachers in England. *Research in Post-Compulsory Education*, 12, (1), 1-17.

Norrish, J. (1996) *Mind the gap: thoughts on self-help and non-judgmental observation in the classroom.* Available at: <http://www.zait.uni-bremen.de/wwwgast/tesl/ej/ej05/a3.html> [Accessed 5<sup>th</sup> July 2006].

Ofsted (2003) *The Initial Training of Further Education Teachers.* (London, Ofsted).

Ofsted (2006) *The Initial Training of Further Education Teachers: 1<sup>st</sup> Annual Report.* (London, Ofsted).



Ofsted (2009) *The Initial Training of Further Education Teachers: an overview 2004-2008*. (London, Ofsted).

O'Leary, M. (2006) Can inspectors really improve the quality of teaching in the PCE sector? Classroom observations under the microscope. *Research in Post-Compulsory Education*, 1 (2), 191-198.

O'Leary, M. (2012) *Surveillance, performativity and normalised practice: the use and impact of graded lesson observations in Further Education colleges*. Accessed from Acaademia.edu (1<sup>st</sup> November 2012).

Ollin, R. (2009a) *The grading of teaching observations: implications for teacher educators in Higher Education partnership*. Final Report. (Huddersfield, University of Huddersfield).

Ollin, R. (2009b) What factors do tutors take into account when deciding whether a trainee's teaching is outstanding? *At the Journal of Vocational Education and Training Eighth International Conference: Researching Vocational Education and Training*, 3<sup>rd</sup>-5<sup>th</sup> July 2009, Worcester College, Oxford. [WWW] <http://eprints.hud.ac.uk/55343/>

Orr, K., and Simmons, R. (2010) Dual Identities: the in-service teacher training experience in the English further education sector. *Journal of Vocational Education and Training*, 62 (1), 75-88.

Peake, G. (2006) *Observation of the practice of teaching*. The consortium for post-16 compulsory education and training – Final Report. (Huddersfield, University of Huddersfield).

Sachs, J. (2000) The activist professional. *Journal of Educational Change*, 1, 77-95.

Sachs, J. (2001) Teacher professional identity: competing discourses, competing outcomes. *Journal of Education Policy*, 16 (2), 149-161.

Shortland, S. (2004) Peer Observation: a tool for staff development or compliance? *Journal of Further and Higher Education*, 28 (2), 219-228.

Thomson, P., Hall, C., Jones, K. (2010) Maggie's Day: a small scale analysis of English education policy. *Journal of Education Policy*, 25 (5), 639-656.

Williams, M. (1989) Developmental Classroom Observation. *English Language Observation Journal*, 43 (2), 85-91.

Wragg, E. (1994) *An Introduction to Classroom Observation*. (London, Routledge).