

# **Exploring the role of lesson observation in the English education system: a review of methods, models and meanings**

Matt O'Leary

*School of Education, University of Wolverhampton, Walsall, UK,*

University of Wolverhampton, School of Education, Walsall Campus, Gorway Road, Walsall, WS1 3BD, UK. Email: [moleary@wlv.ac.uk](mailto:moleary@wlv.ac.uk)

## **Abstract**

Lesson observation has a longstanding tradition in initial teacher education programmes in England. Over the last two decades it has also established itself as an important tool for measuring, assuring and improving the professional skills and knowledge base of in-service teachers and lecturers in schools and colleges. This article reviews a range of related literature across the three education sectors in England (i.e. schools, Further Education (FE) and Higher Education (HE)) in order to compare and contrast the role of observation. In doing so it discusses the key themes and issues surrounding its use in each sector and identifies both common and contrasting patterns. In schools and FE observation has become entrenched as a mechanism of performance management systems, from which a dominant yet contested model has emerged that relies on a simplified rating scale to grade professional competence and performance. In contrast, in HE there is limited evidence of it being linked to the summative assessment of staff, with preferred models being peer-directed and less prescribed, allowing lecturers greater autonomy and control over the use of observation and the opportunity to explore its potential as a means of stimulating critical reflection and professional dialogue about practice among peers.

**Key words: lesson/classroom observation, professional development, teacher assessment, performance management.**

## **Introduction**

This article examines the role of lesson observation in the English education system.

It is divided into three main sections. The first section reviews studies of observation across the Further Education (FE), Higher Education (HE) and schools' sectors in England and discusses key themes and issues. This is followed by an exploration of typologies of observation by examining the interrelated questions of what models of observation exist, in what contexts do they occur and what are their purported aims?

The final section draws together the key themes and issues discussed, concluding

with a synopsis of what extant research reveals about the use of observation in the English education system.

Much of the existing literature on observation is located in the schools' sector with a particular focus on the *practice* of observation and takes the form of textbooks, rather than research-based texts as the first section below explores. Besides there being fewer studies in FE and HE, those that do exist have occurred mainly in the last decade. This in itself is significant as it highlights how observation has a longer history in the schools' sector (Grubb 2000; Wilcox & Gray 1996).

## **A review of literature on lesson observation across educational sectors**

### *The schools' sector experience*

It was during the 1980s that observation materialised as well-established practice in schools. Although it had long been and continues to be a pivotal method of assessment in Initial Teacher Training (ITT) courses, its rise to prominence for qualified teachers was closely linked to wider political reforms at the time, which demanded increased public accountability and an educational reform agenda determined to impose greater control over what teachers did in the classroom (Lowe 2007). Amid concerns about standards and the quality of teaching, observation emerged as a key method of collecting evidence on which to base subsequent systems of teacher evaluation. Wilcox and Gray (1996) referred to the 'dominance' of observation as the main method of collecting data about what went on in

classrooms. This dominance was crystallised by the introduction of a cycle of inspection by the Office for Standards in Education (Ofsted) and formal appraisal for teachers from 1992-94 (Wragg *et al.* 1996).

Wragg's seminal work, *An Introduction to Classroom Observation* (1999), remains the most comprehensive and widely cited textbook on the subject to date. It is located in the context of English primary schools and covers a wide range of themes related to observation pedagogy and theory, although the emphasis is largely on the practical application of observation as a pedagogic tool. Given the breadth of coverage in Wragg's work it is helpful to use it as a starting point from which to explore some of the key issues and themes in the field and to relate these to cognate studies.

At the beginning of his book Wragg raises the issue of the reliability of observation as a form of assessment. He remarks that 'we often "observe" what we want to see' (1999, p. vii). His comment draws attention to the subjectivity of observation and how events are 'inevitably filtered through the interpretive lens of the observer' (Foster 1996, p. 14). The subjectivity of observers' interpretations is a common theme in the literature, particularly when discussing notions of *good practice* (e.g. Fawcett 1996; Montgomery 2002; Tilstone 1998). Wragg maintains that 'mostly when we talk about a "good" teacher, an "effective" strategy or a "bad" lesson, we are referring to our own subjective perception (*op. cit.*, p. 60). He exemplifies his argument by recalling a session in which thirty five highly experienced teacher educators were

shown a videotape of a student teacher's lesson and were asked to grade it on a scale of A-E (A at the top end of the scale and E at the bottom). Their grades varied from a D at the lowest end to a B+ at the top end of the scale. Such differing judgements illustrate the issue of observer subjectivity and reinforce the unreliability of observation as a sole method of assessment, especially when a grading scale is used to measure performance.

Wragg is critical of hierarchical grading systems as he claims that 'the nature of the levels can still be vague and diffuse, using words like "adequate" or "considerable" that are open to widely differing interpretations (*op. cit.*, p. 103). It cannot be assumed that there is a shared understanding among observers or observees as to the meaning and interpretation of value-laden terms such as 'good' and 'outstanding', as used by Ofsted (See Brooks 2009; Wolf 1995). Wragg (1999) argues that these terms, together with the assessment criteria that underpin them, need to be carefully defined when used and attempts made to establish a collective understanding. But even when such attempts are made e.g. standardisation exercises for observers in observation systems and associated assessment criteria, the limitations of what they can achieve need to be acknowledged (Montgomery 2002). In other words, whilst they might be useful in raising collective awareness among observers, it is unrealistic to expect the assessment criteria to be uniformly and consistently applied. Though it has to be said that this is not a phenomenon specific to observation as a method of assessment, but reinforces more widely held beliefs

among key researchers in the field that 'assessment is not an exact science and we must stop presenting it as such' (Gipps 1994, p. 167).

One of the underpinning issues traversing the different contexts and purposes of observation in schools is the notion of *teacher effectiveness*. Campbell *et al.* (2004, p. 3) define teacher effectiveness as 'the impact that classroom factors, such as teaching methods, teacher expectations, classroom organisation, and use of classroom resources, have on students' performance'. However, they are critical of what they see as its narrow conceptualisation in previous studies, where a 'goal-oriented model for measuring teacher effectiveness' (p. 61) has been the preferred approach. For them, too much importance has been attached to the notion of student achievement when attempting to measure teacher effectiveness. They argue that although learning outcomes are the most common form of measurement, there is no clear evidence to suggest that there is a cause and effect relationship between the two. Such simplistic models fail to consider many variables often beyond the control of the teacher that can affect students' performance in any given lesson.

According to Montgomery (2002, p. 18), it was not until the 1980s that discussion surrounding notions of teacher effectiveness in schools came to the fore. Much of the ensuing discussion was characterised by what she described as the 'lack of a coherent and shared theory and practice of teaching'. She emphasises the importance of defining the characteristics of teacher effectiveness if the assessment of teachers

via observation is to become a legitimate and meaningful activity. Yet she is critical of competency-based models used for teacher appraisal that adopt a 'checklist' approach to observation, as they reduce the teacher's role to that of a technician. As Kincheloe (2004, p. 60) argues, 'teachers in the technician paradigm are conceived as a unit of production of an assembly line' or 'operatives' who perform a set function. For Montgomery (2002), competency-based systems exist only to serve 'auditing and accounting purposes' (p. 19). They do little to respond to teacher behaviour that is deemed unsatisfactory or ineffective, as the 'how to do it' is very often left unanswered (p. 39).

In their two-year monitoring study of teacher appraisal in English primary and secondary schools, Wragg *et al.* (1996) highlighted what they described as a 'snapshot' approach to observation (i.e. one-off observations) as one of the main obstacles to identifying incompetent teachers. The reason for this was because such teachers could deliver the 'rehearsed' lesson as a one-off performance, hence avoiding detection. As a means of avoiding this scenario Wragg *et al.* advocated the need for a series of observations to be carried out as part of a longitudinal approach to construct a more realistic picture of a teacher's classroom competence. However, Campbell *et al.* (2004, p. 133) maintain that 'even successive observations of a teacher will only ever supply a collection of snapshots rather than a full picture of teacher behaviour over the year'. While they acknowledge the importance of observation as a source of evidence for systems of teacher appraisal, they also remark that as a

method of data collection 'it is often used with little regard for, or knowledge of, its characteristics' (p.133). What they mean by this comment is that despite its widespread use as a means of gathering data, there is a lack of rigour in its application and insufficient awareness on the part of those carrying out observation of its limitations as a method.

Wragg (1999, p. 3) succinctly summarises some of the paradoxes involved with observation when comparing the ways in which teachers respond to the different contexts in which it occurs and its application as a multipurpose tool in the following comment:

Skilfully handled, classroom observation can benefit both the observer and the person being observed, serving to inform and enhance the professional skills of both people. Badly handled, however, it becomes counter-productive, at its worst arousing hostility, resistance and suspicion.

The rules of observer-observee engagement are likely to differ according to who is observing whom, in what context and for what purpose. Underpinning Wragg's comment and the observer-observee relationship are the notions of power and authority. As Wragg comments, 'the actual or perceived power relationship between observer and observed is not just a sociological concept, but rather a reality that needs to be recognised' (*op. cit.*, p.62). For example, if a head of department is observing a newly qualified teacher, to what extent does s/he feel able to challenge their assessment? Are efforts made to ensure that the observee's voice is heard?

What also links these questions is the degree of ownership and autonomy afforded to the observee.

Ownership and autonomy are identified as key features of successful observation schemes, which are characterised by a move away from authoritarian models where observation is something that is 'done to teachers' to a more egalitarian approach in which ownership of the process is devolved to teachers (e.g. Metcalfe 1999; Tilstone 1998). Tilstone (1998) advocates 'partnership observation', a term which she uses to express a more collaborative, egalitarian relationship between observer and observee. She argues that:

Such partnerships will only work if the [observer] is not regarded as an authoritarian figure and is able to take on the role of facilitator with the teacher in control of direction of the observation and consequent actions (p. 60).

Metcalfe (1999, p. 454) reflects on his experience of the use of observation from the perspective of both a researcher and an Ofsted inspector:

What is becoming clear in schools is that classroom observation, as an aspect of monitoring and evaluation, is felt to be most acceptable when it is part of a broader approach ... in which teachers work collaboratively as opposed to a 'bolted on' approach, which is felt to be connected with 'checking up', accountability and control.

A similar approach to observation is supported by two of the main professional associations for school teachers, the National Union of Teachers (NUT) and the Association of Teachers and Lecturers (ATL) (ATL 2008; NUT 2006). It is their shared belief that observation constitutes an important element of a teacher's professional development and as such should be 'neither a burden for the teacher concerned nor



an opportunity to “police” a teacher’s performance’ (NUT 2006, p. 5), but ‘should be conducted in a manner that equates to a professional dialogue’ (ATL 2008, p. 1).

The advantages of observation programmes that prioritise development over surveillance are well documented. As Wragg argues (1999, p. 17), ‘good classroom observation can lie at the heart of both understanding professional practice and improving its quality’. When it is used insightfully observation can have a profound impact, which ‘can lead to a more open climate, greater trust between colleagues, and the development of strong professional relationships’ (Marriott 2001, p. 3). One of the biggest obstacles to the creation of such a climate would appear to be the issue of grading.

Graded observation has been a contentious issue in schools and provoked a resolute response on the part of the two largest professional associations for schoolteachers, the NUT and the National Association of Schoolmasters Union of Women Teachers (NASUWT) to campaign against its use (e.g. NUT 2007). Both unions believe that grading encourages school management to view observation as a surveillance mechanism with which to monitor the quality of teachers’ work, instead of seeing it as a valuable means of stimulating professional dialogue. Marriott (2001) maintains that the grade can take on such importance that it threatens to undermine the value of the dialogue and feedback between the observer and the observee. Both parties can ‘become hung up on what the grade means rather than how to improve the

teaching' (2001, p. 46). Such is the anxiety surrounding grading that the 'teacher may become over-concerned about whether he or she has "passed"' (*Ibid.*).

One of the most recent developments in the field of observation in schools has been the use of 'lesson study' as a model for improving teaching and learning. One of the unique characteristics of lesson study is how it seeks to involve the learners in the discussion and analysis of the observed lesson(s). According to Lieberman (2009), lesson study puts student and teacher learning at the centre rather than teacher evaluation. In drawing on Stigler & Hiebert (1999), Lieberman (2009, p.83) traces the origins of lesson study to Japan, where it has been used as 'a form of teacher professional development in the improvement of mathematics and science education'.

As a 'model' or 'process' of professional development, lesson study typically involves a group of teachers collaboratively planning a lesson, observing someone teaching it, meeting to discuss their observations and reflections on the lesson and publicly cascading what has been learnt about the process of teaching and learning (Lewis *et al.* 2006). In short, lesson study is designed to promote a collegial approach to observation, enabling teachers to take control of their own development and conceptualise improvement as a shared experience (Lieberman, 2009). (See the section on 'Typologies of observation' below for further discussion of how lesson study operates as a model of observation).

As part of a pilot project in England co-funded by the Teaching and Learning Research Programme (TLRP), the National College for School Leadership and the Centre for British Teachers (CfBT), teachers from twenty schools (both primary and secondary) across eight local authorities were involved in the trialling of lesson study with the aim of answering the project's key question: 'Would Lesson Study work in the UK and if so would it do so in a way which would add value to the range of professional development approaches already in use?' (Dudley 2007). The research was conducted in two phases: the first phase from 2003-06 and the second from 2007-2010.

One of the findings to emerge from the first phase of the project was that lesson study was found to be a 'popular, powerful and replicable process for innovating, developing and transferring pedagogic practices' (*Ibid.*). According to Dudley 'it was popular with both experienced and less experienced teachers alike' and had a 'demonstrable impact on the quality of teaching and on pupil progress and attainment' (*Ibid.*). It remains difficult to analyse the validity and reliability of these claims given that these were emergent findings from an ongoing research project, the methodology of which was not made transparent.

The second phase of the research aimed to explore the 'critical features of teacher learning in lesson study' and what distinguished it as a form of classroom inquiry.

Some of the key findings indicated that lesson study fostered joint risk taking among

teachers, enabled teachers to develop evidence-based practice to inform their professional learning and empowered them to take control of their professional development through their own classrooms (Dudley 2008).

Dudley (2007, 2008) claims that the use of lesson study has subsequently become more widespread across schools in England in recent years, with particular emphasis on mathematics and literacy teaching, though this claim is not supported by any quantitative evidence. Notwithstanding the lack of data to support his claim, there is evidence of a growth in the use of lesson study as a form of intervention in teacher development in other countries. Lieberman (2009), for example, has reported on the popularity of lesson study increasing in the USA over the last decade. In her research Lieberman (2009) found that lesson study encouraged greater openness among staff, which helped to expose vulnerability as an issue that affects both experienced and novice teachers. Reinforcing some of Dudley's (2007, 2008) findings, Lieberman argues that lesson study has helped to foster a collegial approach to teacher development through peer observation and thus prompted teachers to take more risks in their teaching.

### *The FE sector experience*

There has been comparatively little research regarding the role of observation in FE. One of the first studies was Cockburn's (2005). His qualitative research consisted of interviews and focus groups with observers and observees as well as documentary

analysis of the College observation policy and feedback reports. The research was based in one college and although the number of participants is not specified, one gets the impression that it was a relatively small cohort. The study's aim was to report on the perceptions and attitudes of staff to the use of observation.

Cockburn (2005, p. 376) provides what he refers to as a 'typology of resistance' of those who expressed negative views about observation. Some of the issues that he lists resonate with other studies (e.g. O'Leary 2006, 2011; Wragg *et al.* 1996). For example, he suggests that there is evidence of 'artificiality' in lessons as a result of being observed, which leads some tutors to adopt an 'orthodox style of delivering lessons' (p. 380) on the basis that there is a 'formula' for *effective teaching*.

The credibility of the observer emerges as a contested issue amongst observees in Cockburn's study, specifically relating to their experience and suitability to perform the role. Like Wragg (1999) above, he makes reference to the power relationship between observer and observee, arguing that 'the observer is commonly perceived as possessing greater power', which 'is legitimised by organisational arrangements' (2005, p. 384). This is a phenomenon that has been commented on by other writers in the field particularly regarding how it threatens to undermine the developmental potential of observation. This is accentuated when it is used to fulfil the dual purpose of performance management requirements and the developmental needs/goals of teachers (e.g. Ewens and Orr 2002; O'Leary 2006, 2011).

Hardman's research (2007) adopted a case study approach in which the practice of observation was explored in three FE colleges and three HEIs. In seeking to identify differences between the two sectors, she reported that observation was much more heavily associated with Quality Assurance (QA) and performance management systems in FE colleges where there was an emphasis on grading staff. This contrasted with HE where there was a tendency for it to occur in developmental contexts, reflected in the prevalence of informal and ungraded peer observation as the most common model of practice and discussed in the following section. This viewpoint is echoed by Armitage *et al.* (2003, p. 50) who refer to observation as the 'mainstay of the [FE] institution's quality assurance process', although they are also keen to stress its value as 'the basis of some of the most useful professional reflection you can undertake in order to improve performance' (p. 47).

The three FE colleges in Hardman's study all had observation policies and procedures that sought to combine the purposes of QA requirements for Ofsted, together with internal staff development agendas. Observation schemes are undoubtedly time intensive and expensive for colleges. With limited budgets it is hardly surprising that colleges should attempt to dovetail two different purposes into one scheme. However, as Hardman suggests, the effectiveness of such a strategy is questionable. QA requirements appear to take precedence over the developmental needs of teachers. Furthermore, the use of observation for QA purposes is not without its controversies **as has been discussed elsewhere (O'Leary 2011).**

Postlethwaite (2007) states that 'observing classes as part of quality assurance procedures has become a contentious matter in many FE colleges' (cited in James & Biesta 2007, p. 168). This is a viewpoint shared by the main professional association for the sector, the University and College Union (UCU) who have 'call[ed] for a code of practice over how such work is carried out' (Lee 2007, p. 1) as, unlike the schools' sector, there is 'currently no national agreement on lesson observations with the Association of Colleges' (UCU 2009, p. 2). This is in response to complaints from members regarding the draconian, 'intimidatory and not supportive' way in which observation is being used by managers in some colleges (Lee 2007). At the centre of the debate is the issue of grading.

From 2006-09, researchers based at the University of Huddersfield published three separate reports (Burrows 2008; Ollin 2009; Peake 2006) into the use of observation in the context of ITT programmes as part of the 'Huddersfield Post-Compulsory Education and Training (PCET) Consortium', the largest provider of courses leading to professional awards for tutors in the sector.

The first of these reports (Peake 2006) explored the perceptions of teacher educators and trainee teachers concerning the purposes of observation and sought to identify examples of good practice. The research methods consisted of two survey questionnaires, one for trainee tutors and the other for subject specialist mentors, as well as interviews with eleven teacher educators working at five different centres. In

total there were one hundred and thirty four responses to the trainees' questionnaire and only four responses to the mentors' questionnaire, which had been sent to twelve mentors in total. No explanation was provided for why there was such an imbalanced ratio between these two groups, which was surprising given how the report emphasised the importance of subject mentors in the observation process.

The report is largely descriptive and although links are made to previous studies in the field, these are not integrated into a coherent conceptual framework, but appear as a fragmented and arbitrary collection of comments. Although the research lacked a suitable theoretical framework with which to analyse the data, its key findings revealed some interesting areas of discussion that resonated with issues covered in related studies. Below is an adapted summary of these based on the original report:

- Importance of the observer being a subject specialist
- Conflicting purposes of ITT and QA observations
- Concerns regarding the lack of consistency and standardisation of practice between observers across the consortium
- Trainees avoid taking risks in observed sessions
- Value of peer observation in professional development
- Observation is resource intensive – it is time-consuming and expensive



The second report (Burrows 2008) focused on exploring trainees' perceptions of observation as part of the 'new curriculum' i.e. the increase in observation to a minimum of eight hours under the LLUK qualifications (2007), together with the undertaking of observation by mentors to support the development of subject specialist pedagogy. The project was underpinned by two aims:

To identify the perceptions of trainees of observations within the new curriculum [and] to formulate an action plan to improve observations based upon the analysis of the research (p. 5).

Although the research aimed to explore the main issues confronting both pre and in-service trainees and their observers during observation, only the former were included as part of the sample. Eighty participants completed a questionnaire similar in focus and design to that used in Peake's (2006) study, followed by four focus groups in which sixty five trainees participated. Like Peake's study, there is a lot of descriptive detail but no theoretical underpinning. Notwithstanding these shortcomings, the key findings summarised below add to the undeveloped field of observation in FE:

- More structured training of observers needed, particularly subject mentors who are unfamiliar with the observation process
- Over fifty per cent of respondents thought that ITE observation should not be graded compared to a third who thought that it should.
- Observation as 'formative' assessment is a key means of supporting professional development and more value should be attached to it than course assignments. Peer observation is highly valued.

The third report (Ollin 2009, p. 5) was carried out:

in response to the introduction of Ofsted's new grading criteria for inspection of ITE in the [sector] (2009), which state that over fifty per cent of trainees need to be judged 'outstanding' for an ITE provider to achieve the highest inspection grade.

The research explored the implications of introducing the Ofsted scale for teacher educators across the consortium. Ollin remarked that the grading criteria on Huddersfield's programmes 'previously operated on a pass/fail basis' and that the transition to the Ofsted scale 'will influence the way that Certificate in Education/PGCE programmes are developed and delivered' (p. 7). This is a very significant point highlighting the challenges faced by providers as a result of having to adapt to changing policy, while also protecting the values and beliefs that underpin many ITE programmes i.e. the emphasis being on encouraging teacher development through a combination of formative and summative assessment where the former is prioritised over the latter. Nevertheless, Ollin fails to point out that providers are not obliged to adopt Ofsted's scale as a replacement for their own assessment criteria. It is the choice of an institution to decide to do so or not.

The research was 'qualitative and interpretive in nature' and used interviews and observations. These included a small sample size of nine case studies of 'outstanding' and 'weak' trainees during which the ITE tutors were observed carrying out observation and giving feedback, followed by a semi-structured

interview 'focussing initially on understandings and judgements of "outstanding" related to the specific observation' (p. 15). Some of the key findings revealed:

- Mixed interpretations as to what constitutes 'outstanding'. The conclusion reached is that 'outstanding teaching is more than the sum of its parts' (p. 30).
- The need to consider the effect of context on the notion of 'outstanding'.
- Tensions between the 'dual identities' of those in-service trainees observed on ITE programmes and as part of internal college QA schemes. In the first instance they are seen as 'students' with 'developmental needs' and in the second as employees with an obligation to prove their professional competence. Similar tensions were revealed by observers who were involved in observation as ITE tutors but also in college observation for QA purposes.
- Resistance on the part of observers to grading based on the premise that it undermined the developmental nature of ITE observation.

Like the two previous reports from the Huddersfield Consortium, this report is descriptive rather than critical. There is no discernible theoretical framework through which the data are analysed. The research concludes that the Ofsted grading criteria are likely to have a significant impact on future ITE programmes across the Huddersfield consortium and present a challenge in 'balancing the underpinning values related to the learning and development of trainees with increasing demands for standards of teaching to be monitored and assured' (p. 6). It is surprising though that there is no reference to previous studies carried out in this specific field (e.g.

Cope *et al.* 2003; Sharp 2006), which have argued that the graded assessment of observation for trainee teachers is unsustainable on the basis that:

There is no published research which confirms that meaningful grading is possible. Attempts to implement grading schemes ignore the lack of support from research and imply that the assessment of teaching is based on measurement rather than professional judgement (Cope *et al.*, 2003, p. 683).

### *The HE sector experience*

It is only in recent years that observation has begun to emerge in HEIs. This has been partly fuelled by QA demands for greater accountability but more increasingly as a result of its potential for supporting the CPD of lecturers (Hammersley-Fletcher & Orsmond 2004, 2005; McMahon, Barrett & O'Neill 2007; Shortland 2004). Unlike FE and the schools' sector, its use is much less commonplace or prescribed. There is less evidence of links to formal, centralised QA systems and it appears to operate mostly on an informal, voluntary and departmental basis. Hardman's research (2007) reveals that it also occurs as part of academic programmes of professional study for staff, such as the recently created postgraduate award in Learning and Teaching in HE, a compulsory qualification for new staff in many post-1992 universities.

The dominant model used in HE would appear to be 'peer observation' (Hammersley-Fletcher & Orsmond 2004, 2005; Peel 2005; Shortland 2004). Shortland (2004, p. 220) defines peer observation as 'peers observing each other's teaching to enhance teaching quality through reflective practice, thereby aiding professional development'. There are some researchers, however, who contest the generic

application of peer observation in HE as an all-encompassing term for observation and instead prefer the label 'third-party observation' (Fullerton 2003; McMahon, Barrett & O'Neill 2007). For them the term peer observation refers to a specific model of observation based on a collaborative partnership between peers, which is underpinned by 'equality between observer and observed' (McMahon, Barrett & O'Neill 2007, p. 500). This is a legitimate and helpful terminological distinction to make, especially if we are to avoid a blurring of the boundaries between the different models, contexts and purposes of observation as examined in the next section.

Extant research reveals a commonality in the key issues, most of which centre on the perceived opportunities and threats associated with the use of peer observation in HE. Peel (2005) is mindful of its potential danger as a surveillance tool on an institutional level. Research carried out among GP teachers revealed opposition to schemes that used peer observation to address the twin aims of teacher development and QA. Such schemes were considered 'unlikely to succeed if seen to be conveying quality assurance in the guise of tutor support' (Adshead, White & Stephenson 2006, p. 72). The transparency of the aims and objectives of any peer observation scheme in HE is regarded as fundamental to avoid it being viewed with suspicion by lecturers (Hammersley-Fletcher & Orsmond 2005; Peel 2005; Shortland 2004).

In a small-scale qualitative study involving eighteen interviews with lecturers from two academic schools of a post-1992 English university, Hammersley-Fletcher and

Orsmond (2005) explored their experiences as participants in a peer observation scheme. Their findings revealed uncertainty regarding the expectations of their roles as both observer and observee. Some lecturers felt uncomfortable about providing critical feedback for their peers, which the writers (2005, p. 218) recommend 'must be presented in ways that are constructive and will lead to new understandings and improved practice'. The uncertainty and unease expressed by lecturers showed how a shared understanding of what was meant by the term 'critical feedback' was missing. It also exposed their lack of experience in providing constructive feedback.

In Hammersley-Fletcher and Orsmond's study the success of the peer relationship between observer and observee was seen to be dependent on the notions of trust and confidentiality. These were considered fundamental to facilitating honest reflection. It also emerged as an important issue in other studies (e.g. Gosling 2002; Shortland 2004). Gosling (2002, p. 2) talks about the need for staff to be seen as 'genuine peers in which there is real mutuality and respect for each of the participants as equal'. He suggests that the process can be undermined if the observer is senior in hierarchy to the observee, although his claim is unsubstantiated. His concerns seem to be based on the premise that such a relationship is likely to result in more senior members of staff taking charge, hence threatening the equality of the interaction.

In an autobiographical study, Peel (2005) reflects on her personal experiences as a new lecturer and examines the arguments for and against peer observation. She

avers that it can be a useful means of facilitating reflection as long as it incorporates reflection on wider issues of the teaching and learning process and not just that of the observed lesson. She remarks that it was as a result of engaging in critical reflective thinking triggered by the feedback element that led to her successful CPD rather than discussion centring on the observation itself. Thus, peer observation is being used as a 'lens' to stimulate critical reflection (Brookfield 1995). Similarly, Hammersley-Fletcher and Orsmond (2005, p. 223) highlight the importance of reflection. Like Peel, they see peer observation as 'a vehicle for encouraging academics to develop their reflective thinking about their role as professional lecturers and to seek and engage in developmental processes as a result'.

### **Typologies of observation: contexts, models and purposes**

Studies exploring observation contexts, models and purposes are limited. I have chosen to focus on Gosling (2002) and Wragg's (1999) typologies as a general reference point because they provide contrasting means of categorising observation. Gosling's framework of analysis categorises observation into three *models*, whereas Wragg's is grouped in to *contexts*. Besides, they are arguably the most cited sources. However, there is also discussion of the model of 'lesson study' (Stigler & Hiebert 1999; Lewis et al., 2006) towards the end of this section.

Gosling (2002) categorises observation into three distinct models, all of which he labels as 'peer observation'. A description of each model is shown in **Table 1** below.

<b>characteristic</b>	<b>evaluation model</b>	<b>development model</b>	<b>peer review model</b>
Who does it & to whom?	Senior staff observe other staff	Educational developers observe practitioners; or expert teachers observe in department	teachers observe each other
Purpose	Identify under-performance, confirm probation, appraisal, promotion, quality assurance, assessment	Demonstrate competency/improve teaching competencies; assessment	engagement in discussion about teaching; self and mutual reflection
Outcome	Report/judgement	report/action plan; pass/fail PGCert	Analysis, discussion, wider experience of teaching methods
Status of evidence	authority	expert diagnosis	peer shared perception
Relationship of observer to observed	power	expertise	equality/mutuality
Confidentiality	Between manager, observer and staff observed	Between observer and the observed, examiner	Between observer and the observed - shared within learning set
Inclusion	Selected staff	Selected/ sample	all
Judgement	Pass/fail, score, quality assessment, worthy/unworthy	How to improve; pass/fail	Non-judgemental, constructive feedback
What is observed?	Teaching performance	Teaching performance, class, learning materials,	Teaching performance, class, learning materials,
Who benefits?	Institution	The observed	Mutual between peers
Conditions for success	Embedded management processes	Effective central unit	Teaching is valued, discussed
Risks	Alienation, lack of co-operation, opposition	No shared ownership, lack of impact	Complacency, conservatism, unfocused

**Table 1:** Gosling's Models of Peer Observation of Teaching (2002, pp. 4-5)

The accuracy and consistency of some of the terms Gosling uses are questionable. As noted previously, the term peer observation is generally agreed to have very specific



connotations and thus it is terminologically inaccurate to use it in the generic way that he does. In fact, it is only his third model, which he refers to as the 'peer review model' that can be described as a genuine example of peer observation.

Nevertheless, his 'models' still have value as a way of examining the different contexts and purposes for which observation is used.

Gosling's 'evaluation model' typically reflects the approach taken in formal QA processes such as appraisal whereby senior staff observe teachers. The observer carries out an assessment of their professional competence, culminating in a summative judgement. QA systems such as probationary periods for new staff, external inspections and internal audits fall into this category.

The purpose of Gosling's second 'developmental model' is for observees to 'demonstrate competency' and for the observer to 'improve teaching competencies'.

The observer adopts a formative and summative approach, as it is envisaged that the observation will be graded on a pass/fail basis and the findings will lead to recommendations for improvement and inform a future action plan. This model reflects some aspects of practice commonly found in ITT and CPD programmes, although it seems to be based on a deficit model as there is an assumption that the observee is deficient in some way and the role of the observer is to make good what s/he identifies as the deficit. McMahon, Barrett, and O'Neill (2007, p. 502) argue that 'Gosling suggests that this model can be clearly differentiated from the first but, in

practice, the boundaries are anything but clear.' Thus the distinctions between these two models are rather blurred and the summative aspect of Gosling's model appears to overwhelm the formative.

Gosling's third model is what he refers to as a 'peer review model', where tutors observe each other as part of a formative process. This is what was referred to earlier as 'peer observation'. Bennett and Barp (2008, p. 563) remark that this model 'specifically assumes an underlying dynamic of equality and mutuality of learning where feedback from the observer is non-judgemental and constructive, in a spirit of co-participation'. Thus, the value of observation is mutual and serves the dual purpose of promoting the development of observer and observee. There are parallels to Lave and Wenger's (1991) work on 'communities of practice' in terms of the core values that underpin such conceptualisations i.e. the sharing of practice, especially with experienced and novice tutors working collaboratively and cooperatively.

Equally, there are also links to lesson study as an alternative model, discussed further below.

In contrast to Gosling, Wragg (1999) chose not to categorise observation into *models*, but rather into the *contexts* in which it occurred (**Table 2** below). Wragg also chose to explore these contexts through descriptive prose in a separate chapter entitled 'Classroom observation in context' rather than in tabular format like Gosling. However, to facilitate comparative analysis between the two and to condense

Wragg's original chapter, I have chosen to present a summarised version of the 'major contexts' he discusses in the form of a table (**Table 2** below) with bullet points comprising the key features of each context.

<b>Major contexts of observation</b>	<b>Summary description</b>
<b>ITT</b>	<ul style="list-style-type: none"> <li>• Value of observing others for trainee teachers' development</li> <li>• Peer observation discussed in the context of reciprocal pair work i.e. observer acts as 'another pair of eyes'</li> <li>• Supervision and mentoring – role of mentor carried out by teacher in the school where the trainee teacher is placed and/or a tutor from the ITT institution</li> <li>• Importance of the mentor-trainee relationship, underpinned by 'trust' and 'respect'</li> </ul>
<b>INSET and CPD</b>	<ul style="list-style-type: none"> <li>• Very similar focus to ITT</li> <li>• The focus tends to concentrate on developing skills of self-reflection amongst practitioners</li> <li>• Teachers are observed by more senior staff, or by their peers and vice versa</li> <li>• Peer observation identified as the principal use</li> <li>• Can occur across institutions</li> </ul>
<b>Studying pupils</b>	<ul style="list-style-type: none"> <li>• Learner-centred focus i.e. tracking pupil behaviour, ability to work in pairs/groups etc.</li> </ul>
<b>Curriculum development and evaluation</b>	<ul style="list-style-type: none"> <li>• Focus on how aspects of the curriculum are being implemented by the teacher</li> <li>• Principal use is 'to match intent against action' i.e. 'Is there a mismatch between intention and strategies?'</li> <li>• Observer usually senior member of staff e.g. deputy head</li> </ul>
<b>Job analysis</b>	<ul style="list-style-type: none"> <li>• Audit of what teachers typically spend their time doing in the classroom e.g. how much time is spent on assessment?</li> <li>• Holistic view of teacher's role i.e. observation forms part of a collection of data for job analysis</li> </ul>
<b>Teacher appraisal</b>	<ul style="list-style-type: none"> <li>• observation forms an integral part of the formal appraisal process but part of a collection of different data</li> <li>• Appraisal of teaching competence must be both retrospective and prospective, looking back at what has been achieved and forward to what might be done in the future in order to make an impact</li> <li>• Most common form is a 'supervisor-subordinate model' i.e. more senior person observes junior colleague</li> <li>• Links with disciplinary action against 'incompetent teachers'</li> <li>• Problem of 'snapshot' observation highlighted</li> </ul>

**Table 2** – Wragg's major contexts of observation (adapted from Wragg 1999, pp. 82-101)

Wragg's first two contexts, 'ITT' and 'INSET and CPD' share many similarities in their focus. For example, in both instances there is an emphasis on the value of peer observation and its reciprocal benefits for observee and observer in their professional development. Though Wragg does not define his use of the term peer observation at any point it is generally used in a way that is consistent with previous definitions i.e. collaborative, constructive and non-judgemental (Bennett & Barp 2008). For Wragg, distinctions between the use of observation in ITT and INSET are governed by its underpinning purpose. In the ITT context there is a tendency for trainee teachers to observe experienced others' teaching (live or recorded lessons) as an important means of seeing particular skills and methods in action that they will be expected to master and subsequently demonstrate in their own teaching when observed by a school-based mentor and/or an ITT course tutor in a formative and summative capacity. In an INSET context, Wragg argues that the emphasis is on developing the skills of self-reflection among teachers, implying a movement towards greater self-sufficiency and empowerment for the experienced teacher in shaping their own professional development.

Despite the differences in analytical frameworks between Wragg and Gosling, Wragg's contextualisation of observation can be regarded as more extensive than Gosling's. A good example of this is how he includes contexts that extend beyond the standard conceptualisation of observation as centring on the observed teacher. For example, contexts three and four, 'studying pupils' and 'curriculum

development and evaluation' illustrate how the focus of observation does not always have to be on the teacher's performance and/or development as is the case with Gosling's three models. It is difficult to find any evidence of these two contexts of observation in either the FE or HE sector, which suggests that they may be specific to schools and even primary schools in particular. In fact, there appear to be many similarities between these two contexts referred to by Wragg and the recent development of the lesson study model in schools (see below for more discussion).

Wragg's sixth context of observation, 'teacher appraisal', is similar to Gosling's 'evaluation model' inasmuch as a teacher's professional competence is what is judged. Both also highlight the significance of the hierarchical nature of the power relationship between observer and observee as the former is usually a senior member of staff. Wragg's conceptualisation of observation in the context of appraisal differs to Gosling's in two respects. Firstly, it situates the use of observation as part of a 'collection of data', thus recognising that it is one of several sources of evidence in the formal appraisal process. Secondly, it highlights the shortcomings of adopting a 'snapshot' approach to appraisal through observation previously discussed by encouraging both a 'retrospective' and 'prospective' approach. Wragg's decision to categorise observation with reference to contexts rather than models also suggests that this is a more meaningful way of configuring it as it avoids the blurred boundaries between Gosling's models.

Gosling's 'evaluation' and 'developmental' models share two noticeable characteristics. The first is that they are observer-led and the second is that they are indicative of Freire's (1977) 'banking concept of education' i.e. that the observed teacher is an account waiting to be filled with deposits of wisdom from the observer, thus suggesting an imbalance of power and expertise in the observer-observee relationship and the degree of autonomy afforded the latter. Tilstone (1998) refers to this as a 'non-collaborative' approach that was very much rooted in early practices within ITT and the assessment of novice teachers whereby the observer would sit at the back of the room and make a list of prescriptive judgements or 'sitting next to Nellie'. Yet in his conceptualisation of the role of observation in ITT, Wragg (1999) infers a more equitable mentor-trainee relationship underpinned by the importance of mutual 'trust' and 'respect' in order for it to be successful.

It is only Gosling's third 'peer review' model that reflects a more balanced, equitable relationship, which offers teachers the opportunity to take ownership of their own professional development. As McMahon, Barrett, and O'Neill (2007, pp. 502-3) remark, 'it is the only model in which the reflection does not take an inevitable second place to concerns about managerial/inspectorial judgements'.

An essential ingredient of the success of any observation scheme is that the method/approach employed should suit its purpose(s). Wragg (1999, p. 4) argues that in theory 'the purpose, timing and context of an observation should largely

determine its methods'. For example, if the focus of the observation is meant to be collaborative and developmental, as discussed above with peer observation, then it would seem inappropriate for the observer to undertake a summative assessment of his/her peer. It would appear, however, that the theoretical delineations between different models, contexts and purposes of observation have become blurred in practice. As McMahon, Barrett, and O'Neill (2007, p. 503) warn, one of the dangers of this is that it allows:

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.

Finally, as commented earlier, one of the most recent developments in the use of observation as a form of professional development has been that of 'lesson study', which seems to build on many of the features of models of peer observation previously discussed. Lewis *et al.* (2006) have argued that one of the defining characteristics of lesson study is the way in which it approaches the observation process. The lesson is seen as a collaborative research context in which to collect data that will provide the platform for classroom-based inquiry:

The observed lessons, called "research lessons", are regarded not as an end in themselves but as a window on the larger vision of education shared by the group of teachers, one of whom agrees to teach the lesson while all the others make detailed records of the learning and teaching as it unfolds. These data are shared during a post-lesson colloquium where they are used to reflect on the lesson and on learning and teaching more broadly (Lewis *et al.* 2006, p. 3).

Instead of the summative assessment of the teacher's performance being the *raison d'être* of the lesson, thus leading to a *fetishisation* of the one-off observed lesson as

O'Leary (2011) has argued, lesson study adopts a longitudinal approach that concentrates on collecting data about a particular aspect of the curriculum with the ultimate aim of reflecting on how student learning can be further enhanced. Thus the spotlight is not shone on the individual teacher's competence and performance but on the student experience. In short, unlike performance-driven models of observation where the observations and grades are an end in themselves, the lesson study model uses observation as a method of data collection whilst avoiding making judgements about individual teachers. Thus it could be argued that lesson study has the potential to make a significant contribution to developing understanding and improvement of teaching and learning but with the caveat that the organisational elements are carefully considered (Grierson & Gallagher 2009) and that further research is needed in order to explore its potential (Lewis *et al.* 2006).

### **Concluding remarks**

There are clearly recurring themes surrounding the use of observation in all three sectors. For example, its value as a means of stimulating reflection on practice by engaging in professional dialogue with colleagues, who act as 'critical mirrors' (Brookfield 1995) seems to be a shared interpretation among researchers and practitioners in all three sectors, albeit with the caveat that specific ground rules need to be established for this to work successfully i.e. notions of mutual trust, respect, ownership etc. At the same time, there are divergences between the three that partly reflect their historical status and the history of policy in each sector. In FE



and schools observation appears to have operated principally to satisfy policy driven agendas of performance management systems. In HE, however, its role is less prescribed, thus allowing lecturers more autonomy and control over its use.

Observation has invariably been situated in performance management systems in FE and schools and much of the literature also links this to related discussions about 'teacher effectiveness' and 'good practice'. Such approaches to observation have tended to adopt a technicist model of teaching and learning, seeking to reduce it to a list of measurable competencies and failing to appreciate its complexities, particularly the need to consider context (James & Biesta 2007; Ollin 2009).

The use of graded observation has triggered debate regarding the reliability of observation as a form of assessment in schools (Wragg 1999) and more so recently in FE (e.g. O'Leary 2011). Graded observation is perhaps the single most contentious issue relating to the topic. Its performance focus is something that seems to have provoked strong reactions across both sectors. Brown (in Brown, Jones and Rawnsley 1993, p.51) compared such models of observation to the traditional examination in that both teacher and student are required to produce 'peak performance under stressful conditions with little opportunity for dialogue with the examiner and no real chance to gain meaningful feedback on how things are going.'

Some of the existing research in the field has highlighted how the performative nature of graded observation has resulted in a decline in the creativity and innovation of teachers' work in the classroom. There is a reluctance to want to 'take risks' for fear of being given a low grade. Teachers are aware of the need to 'play the game' (Jeffrey & Woods 1998), which can result in them following a collective template of 'good practice' during observation. According to Elliott (1990, p. 83), this is an example of 'management exercising control over performance by preventing teachers from reflexively developing new understandings of the nature of teaching and learning tasks'. In the FE context, Peake (2006) has illustrated how even in the context of ITT trainees avoid taking risks during observation. Yet recent research into the use of lesson study among qualified teachers in schools seems to suggest a counterbalance to this.

With regard to the use of observation as a formative tool for CPD, there would appear to be a commonality across much of the literature in terms of some of the key concepts discussed i.e. collaboration, equality, autonomy, ownership, trust etc. Much of this work has focused on the use of lesson study in the schools' sector and peer observation, particularly in the HE context. Referring to Ramsden (1992), Jones (in Brown, Jones and Rawnsley 1993, p. 31) comments that in order for observation to work it needs to be part of a teacher's professional development and not something that is 'done to them':

Ownership of observation needs to be devolved down as much as possible to the participants in the teaching process. The closer the ownership of the process is located to the actual participants, the more likely it is that the aims will be achieved and the outcomes accepted by all concerned (Brown, Jones and Rawnsley, 1993, p. 10).

To conclude, across all three sectors previous studies reveal that observation is regarded as an important means of evaluating, reflecting on and improving the quality of teaching and learning as well as contributing to a greater understanding of these processes. Whether this occurs as part of QA systems or CPD programmes, the central role that observation has to play in the professional practice of teachers seems incontestable. Where the contestations start to emerge, however, is in relation to the stated aims behind its use, the extent to which the outcomes match these aims and the way in which the process of observation is operationalised.

Wragg (1999) argued that the *purpose* of observation should largely determine how it is used, but evidence suggests that the boundaries between different models, contexts and purposes have become blurred and contested. At the heart of these contestations lies a conflict between the forces of 'structure' and 'agency' and related notions of power and control, which manifests itself in the sometimes paradoxical agendas of policy makers, the institution and its teaching staff. This conflict is epitomised by the way in which the developmental needs of staff and the requirements of performance management systems are often conflated into a 'one-size-fits-all' model of observation in the case of schools and colleges. Whilst it could be argued that there is evidence of convergence between these two forces in relation

to the *ends* i.e. the improvement of teaching and learning, the *means* by which such improvement can be achieved remains more contentious.

## References

Adshead, L., White, P. T., & Stephenson, A. 2006. Introducing peer observation of teaching to GP teachers: a questionnaire study. *Medical Teacher*, 28, 68-73.

Armitage, A. Byrant, R. Dunnill, R. Hammersley, M. Hayes, D. Hudson, A. & Lawless, S. 2003. *Teaching and training in post-compulsory education*. 2<sup>nd</sup> ed., Buckingham: OUP.

Association of Teachers and Lecturers (ATL) 2008. *ATL Advice – Classroom*

*Observation*. Available online at:

<http://www.atl.org.uk/Images/ADV19%20classroom%20observation.pdf>. Date accessed 23/3/2010.

Bennett, S. & Barp, D. 2008. 'Peer observation - a case for doing it online'. *Teaching in Higher Education*, 13(5), 559 – 570.

Brookfield, S. D. 1995. *Becoming a Critically Reflective Teacher*. San Francisco, CA: Jossey-Bass.

Brooks, V. 2009. 'Marking as judgment', *Research Papers in Education*, First published on: 11 November 2009 (iFirst).

Brown, S., Jones, G. & Rawnsley, S. 1993. *Observing Teaching – SEDA Paper 79*. Birmingham: SEDA.

Burrows, J. 2008. *Trainee Perceptions of Observation*. Huddersfield: Huddersfield Consortium.

Campbell, J., Kyriakides, L., Muijs, D & Robinson, W. 2004. *Assessing Teacher Effectiveness – Developing a Differentiated Model*. London:Routledge Falmer.

Cockburn, J. 2005. Perspectives and politics of classroom observation. *Research in Post-Compulsory Education*, 10(3), 373 – 388.

Cope, P., Bruce, A., McNally, J. & Wilson, G. 2003. 'Grading the practice of teaching: an unholy union of incompatibles'. *Assessment & Evaluation in Higher Education*, Vol. 28(6), 673 – 684.

Dudley, P. 2007. Lessons for Learning: Using lesson study to innovate, develop and transfer pedagogic approaches and metapedagogy. London: TLRP. Available online at: <http://www.bera.ac.uk/lesson-study/>. Date Accessed: 4/10/2011.

Dudley, P. 2008. *Lesson Study Development in England: practice to policy*, paper presented at the World Association of Lesson Studies Annual Conference, Hong Kong, December 2008. Available at: [http://www.tlrp-archive.org/tlrp/upload/assets/1231760840\\_PDWALS08.pdf](http://www.tlrp-archive.org/tlrp/upload/assets/1231760840_PDWALS08.pdf) Date Accessed: 4/10/2011

Ewens, D. & Orr, S. 2002. *Tensions between evaluation and peer review models: Lessons from the HE/FE border*. London: LTSN Generic Centre.

Fawcett, M. 1996. *Learning Through Child Observation*. London: Jessica Kingsley.

Foster, P. 1996. *Observing Schools – A Methodological Guide*. London: Paul Chapman Publishing.

Freire, P. 1972. *Pedagogy of the Oppressed*. Harmondsworth: Penguin.

Fullerton, H. 2003. Observation of teaching, in Fry, H., Ketteridge, S. & Marshall, S. (Eds) *A handbook for teaching and learning in higher education* (2nd edition). London: Kogan-Page.

Gipps, C. 1994. *Beyond Testing: Towards a Theory of Educational Assessment*. London: Falmer Press.

Gosling, D. 2002. *Models of Peer Observation of Teaching*. London: LTSN Generic Centre.

Grierson, A. & Gallagher, T.L. 2009. Seeing is believing: Creating a catalyst to teacher change through a demonstration classroom initiative. *Professional Development in Education*, 35(4), 567-584.

Grubb, W. N. 2000. Opening Classrooms and Improving Teaching: Lessons from School Inspections in England. *Teachers College Record*, 102(4), 696-723.

Hammersley-Fletcher, L. & Orsmond, P. 2004. Evaluating our peers: is peer observation a meaningful process? *Studies in Higher Education*, 29(4), 489-503.

Hammersley-Fletcher, L. & Orsmond, P. 2005. Reflecting on reflective practices within peer observation. *Studies in Higher Education*, 30(2), 213-224.

Hardman, J. 2007. *The use of teaching observation in Higher Education: An exploration of the relationship between teacher observation for quality assurance and quality improvement in teaching in higher education, in the light of further education sector experience*. Report produced for Escalate August 2007. Available online at: <http://escalate.ac.uk/3311>.  
Date accessed 20/1/2010.

- James, D. & Biesta, G. eds. 2007. *Improving learning cultures in Further Education*. London: Routledge.
- Kincheloe, J. L. 2004. The Knowledges of Teacher Education: Developing a Critical Complex Epistemology, *Teacher Education Quarterly*, 31(1), 49-66.
- Lave, J. & Wenger, E. 1991. *Situated Learning – Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Lee, J. 2007. *Are you being observed?* London: Times Educational Supplement, 15 June 2007.
- Lewis, C., Perry, R. & Murata, A. 2006. How should research contribute to instructional improvement? The case of lesson study, *Educational Researcher*, 35(3), 3–14.
- Lowe, R. 2007. *The Death of Progressive Education: How teachers lost control of the classroom*. Oxford: Routledge
- Marriott, G. 2001. *Observing Teachers at Work*. Oxford: Heinemann.
- McMahon, T., Barrett, T., & O’Neill, G. 2007. Using observation of teaching to improve quality: Finding your way through the muddle of competing conceptions, confusion of practice and mutually exclusive intentions. *Teaching in Higher Education* 12(4), 499-511.
- Metcalf, C. 1999. Developmental Classroom Observation as a Component of Monitoring and Evaluating the Work of Subject Departments in Secondary Schools. *Journal of In-service Education*, 25(3), 447-459.

Montgomery, D. 2002. *Helping Teachers Develop through Classroom Observation – Second Edition*. London: David Fulton Publishers.

National Union of Teachers (NUT) 2006. *A Classroom Observation Protocol: Guidelines for NUT School Representatives*. Available online at:  
<http://www.teachers.org.uk/node/1335> Date accessed: 22/1/2010.

National Union of Teachers (NUT) 2007. Classroom observation: No to grading of lessons. *NUT News*, 17, September 2007. Available online at:  
<http://www.teachers.org.uk/node/6298> Date accessed: 22/1/2010.

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*, 11(2), 191 – 198.

O'Leary, M. 2011. *The Role of Lesson Observation in Shaping Professional Identity, Learning and Development in Further Education Colleges in the West Midlands*, unpublished PhD Thesis, University of Warwick, September 2011.

Ollin, R. 2009. *The grading of teaching observations: implications for teacher educators in Higher Education partnerships*. Huddersfield. Huddersfield Consortium.

Peake, G. 2006. *Observation of the Practice of Teaching*. Huddersfield. Huddersfield Consortium.

Peel, D. 2005. Peer Observation as a transformatory tool? *Teaching In Higher Education*, 10(4), 489-504.

Ramsden, P. 1992. *Learning to Teach in Higher Education*. London: Routledge.



Sadler, D.R. 2005. Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education* 30, 175–94.

Sharp, S. 2006. [The grading of placement in initial teacher education in Scotland](http://www.scotedreview.org.uk/pdf/208.pdf). Available online at: <http://www.scotedreview.org.uk/pdf/208.pdf> . Date accessed: 17/5/2010.

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

Simons, H. & Elliott, J. eds. 1990. *Rethinking Appraisal and Assessment*. Milton Keynes: Open University Press.

Stigler, J. & Hiebert, J. 1999. *The teaching gap: best ideas from the world's teachers for improving education in the classroom*. New York: The Free Press.

Tilstone, C. 1998. *Observing Teaching and Learning – Principles and Practice*. London: David Fulton.

University and College Union (UCU) 2009. *Lesson Observation: UCU Guidelines*.

[Accessed 22/02/2010] Available online at:

<http://www.ucu.org.uk/index.cfm?articleid=2969>.

Wilcox, B. and Gray, J. 1996. *Inspecting schools: holding schools to account and helping schools to improve*. Buckingham: Open University Press.

Wragg, E. C., Wikeley, F. J., Wragg, C. M. & Haynes, G. S. 1996. *Teacher Appraisal Observed*. London: Routledge.

Wragg, E.C. 1999. *An Introduction to Classroom Observation – 2<sup>nd</sup> Edition*. London: Routledge