

The Quality of Secondary Mathematics PGCE Courses: a critical perspective on the inspection of initial teacher education

Keith Jones, University of Southampton

Anne Sinkinson, Homerton College, Cambridge

The quality of initial teacher education courses has been the subject of adverse comment and media speculation for some time. During 1996/7, Ofsted began a comprehensive round of inspections of initial teacher education providers using an extensive framework of inspection. This paper reports the results of the first round of inspection of secondary mathematics PGCE providers. Almost three-quarters were judged to be good or better. In examining the inspection reports from a critical perspective, this paper focuses on the level of consistency in the judgements made in the published inspection reports.

Introduction

For more than ten years, courses of initial teacher education have been subjected to adverse comment. Lawlor (1990), for example, suggested that higher education should cease to have anything to do with the training of teachers, charging that, “despite the intentions of government reforms, the training discourages good candidates from entering the profession and undermines the standards of those who do” (p7). The wholesale condemnation of teacher education continued through the 1990s with argument being promulgated that if standards in schools were too low it was because the teachers were not trained properly (for an example, see the statement from the Chief Inspector of Schools in Convey 1997 p 30). Whatever the basis, claims of poor performance in initial teacher education courses have been used as part of the rationale for introducing a raft of measures (for evidence, see TTA press releases 19/96, 24/96, 45/97, and so on), including the setting up of the Teacher Training Agency (choosing the word training, rather than education), formulating new standards for the award of qualified teacher status, a ‘national teacher training curriculum’, a framework for external inspection, and ‘league tables’ of initial teacher education providers.

Most recently, Her Majesty’s Chief Inspector of Schools (Ofsted 1999) has declared that there remain “key weaknesses” in courses of initial teacher ‘training’ such that “rigorous inspection of ITT must therefore continue so that institutions of higher education can remedy the weaknesses that remain”. Yet the validity and reliability of Ofsted inspection procedures have themselves been called into question (see, for example, Matthews *et al* 1998). Such issues are of particular importance in initial teacher education as a poor Ofsted rating can lead to the rapid withdrawal of TTA accreditation, meaning course closure, while even satisfactory ratings can lead to

uncertainty over course quota, leading to a spiral of decline in course viability. This paper presents a preliminary analysis of the reports of Ofsted inspections of secondary mathematics PGCE courses carried out between September 1996 and July 1997. The analysis presented below shows that standards in the 20 providers inspected were judged by Ofsted inspectors to be good or better in almost three-quarters of cases. For only one course was there judged to be an issue of poor quality and this in only one aspect considered in the inspection process. This particular institution now no longer offers a secondary mathematics PGCE course, having decided to close its secondary mathematics course following a decision by the TTA to begin the process of withdrawing accreditation (see TES 30/10/98).

In reviewing the inspection judgements, we focus on the validity and reliability of inspection process. In doing so, we should emphasise that, as the authors of this paper, we were quite satisfied with both the conduct of the inspection at our own institutions, and the outcomes. This is not partisan research. This paper is offered as a modest contribution to the process of opening the inspection of teacher education to proper academic scrutiny with a view to informing its procedures, practice, and quality. The Ofsted inspection process is not itself above critical examination.

Validity and Reliability of Inspection

The necessity for close scrutiny of the inspection process is borne out in a number of academic publications (for instance, Graham 1997, Matthews *et al* 1998). Indeed, the role of Ofsted, including the validity and reliability of the inspection process, is the subject of a comprehensive review currently being conducted by the Education Sub-committee of the House of Commons Select Committee on Education and Employment (see Education and Employment Committee Press Notice No. 40, 21 September 1998).

Graham (1997 p1) claims that “the inspection process has left many HEIs feeling that the quality of their provision has not been fairly judged”. He argues that the methodology of the Ofsted/TTA Framework for the Assessment of Quality and Standards is unproven and reductionist, and that the “demarcations of grade boundaries have not been agreed or exemplified” (p6). It seems that Ofsted promised exemplar material for the grades (on a 1-4 scale) which has never appeared. Yet, as Gilroy and Wilcox (1997) explain, the construction of such exemplification assumes that the criteria are unambiguous and their interpretation and application straightforward. Neither, they argue, is the case. They suggest that inspectors are likely to have developed their own ‘rules of thumb’ which means that the scope for a wide variety of practice on the part of inspectors “would seem considerable ... thus rendering doubtful the notion of consistent and objective practice” (p28).

This uncertainty in the inspection judgements means that there is a possibility that a course that is actually satisfactory may be awarded a grade 4, the lowest grade, which results in course closure. Tymms (1997) constructed a mathematical model to simulate thousands of inspections and calculated the likelihood of various results for different levels of security of Ofsted judgement. In particular he looked at the likelihood of a Grade 4 being given to institutions which are in fact performing satisfactorily. His calculations revealed that a satisfactory course faces *a 50% chance of being judged to be failing*.

Problems of reliability and validity in the inspection process prompted Ofsted inspectors to investigate the matter. Matthews *et al* (1998), report on the judgements of 100 pairs of trained school inspectors who independently observed the same school lesson. Agreement occurred in 80% of cases. In 3% of cases, the judgement differed by as much as two grades. Matthews *et al* found “less agreement on grades .. at a sensitive area of the grading scale” (p184), the boundary between satisfactory and unsatisfactory. They also found more agreement about *weaknesses* than about strengths and suggest a problem with the inspection procedure is that “inspectors take more care to record all the weaknesses .. than they do to record all the strengths”. They conclude that “the anomalies here call into question the reliability of the judgments made by a very small proportion of inspectors on one occasion” (p186). Given that secondary subject courses of initial teacher education are generally carried out by a single inspector, an instance of unreliability could have devastating results. What is more, the study by Matthews *et al* was of “the more confident and experienced inspectors” of schools (p186). The comprehensive inspection of PGCE courses is a more recent development, with UCET claiming to have evidence of “some ill-prepared inspectors” who were “insufficiently qualified to inspect higher education provision” (UCET 1998).

Methodology

The available literature suggests that, in a critical review of the Ofsted reports, we need to look for evidence of the demarcation of grade boundaries, particularly what distinguishes one grade from another, examine the tone of reports to see whether weaknesses outweigh expressions of course strength, and investigate the reliability of the judgments, by searching for possible inconsistencies across equal grades. We note that Ofsted reports are Crown Copyright and that a condition of use is that “extracts quoted are reproduced verbatim without adaptation and on condition that the source and date thereof are stated”. In reporting our analysis we are conscious of the ethical issues involved in identifying individual institutions. All the extracts we reproduce below are quoted verbatim without adaptation from the 20 Ofsted

inspection reports published by Ofsted as a result of the round of inspections of secondary mathematics PGCE courses carried out during 1996/7 (Ofsted 1998).

Results

The inspection framework (Ofsted/TTA 1996) set out 14 ‘cells’ that could be inspected. For the 1996/7 round of inspections, six ‘cells’ were inspected, those selected being shown in table 1, below. The tables give an overview of the inspection grades for each ‘cell’ inspected (source: Ofsted 1998, n=20). Grade 1 signifies ‘very good, with several outstanding features’, grade 2 ‘good, with no significant weaknesses’, grade 3 ‘adequate, but requires significant improvement’, and grade 4 is ‘poor quality’.

Cell	Grade	1	2	3	4
S1	selection procedures	35%	55%	10%	0
T2	quality of training	30%	40%	30%	0
T4	assessment of student teachers	15%	60%	20%	5%
C1	student teachers’ subject knowledge	35%	55%	10%	0
C2	student teachers’ planning and teaching	25%	50%	25%	0
C3	student teachers’ assessment of pupils	15%	60%	25%	0

Table 1: Grade profile of secondary mathematics PGCE courses, 1996-97 inspections

		Percentage of courses rated ‘good or better’	Percentage of courses rated ‘adequate or better’
S1	selection procedures	80	100
T2	quality of training	70	100
T4	assessment of student teachers	75	95
C1	student teachers’ subject knowledge	90	100
C2	student teachers’ planning and teaching	75	100
C3	student teachers’ assessment of pupils	75	100

Table 2: Profile of ‘good or better’ and ‘satisfactory or better’ by cell, 1996-97 inspections

The tables illustrate that standards in all inspected ‘cells’ were judged by Ofsted inspectors to be *good or better* in almost three-quarters of cases. For only one provider was there judged to be any issue of ‘non-compliance’ and this in only one ‘cell’ from those inspected. In the data presented below, we focus on cells T2 (the quality of training) and C2 (the student teachers’ planning and teaching), both because these are particularly important aspects of a PGCE course, and because there appears to be more variation in the grades awarded in these cells.

There is a variation in emphasis given to cell T2 (the quality of training) in the reports, from three to eight paragraphs, a variation that does not appear to be related to the grade awarded for the cell. The cell is characterised in the framework for inspection by nine criteria, one of which relates to the quality of the training sessions. In only one case were the university-based training sessions identified as being anything other than good or very good. Despite this, *six courses were judged to be grade 3* overall in cell T2. Another criterion relates to students' subject knowledge. Here there appears to be considerable similarity between comments made on this aspect, yet the cell grades differ markedly. For example, a Grade 1 report says "There is no formal audit of students' subject knowledge after selection but the training sessions encourage students to work in groups", while a Grade 3 report says "There is no formal subject knowledge audit and the responsibility for filling gaps in subject knowledge rests with the student. This is unreliable". Finally, not all the criteria for cell T2 are addressed in the published reports. For example, the criterion that "the training is differentiated" is not mentioned in at least 40% of the reports.

Cell C2 covers the students teachers' planning, teaching and classroom management. The sections of the Ofsted reports on this cell vary in length from 223 words to 692, a variation that does not appear to be related to the grade awarded. Given that there are 11 criteria for this cell, no particular criterion appears to be more critical than any other, as the following examples illustrate: Grade 1 course, "Students select appropriate objectives and content for most of their classes"; Grade 3 course, lesson plans "usually contain explicit objectives"; Grade 1 course: "Students taught whole classes well", Grade 3 course, "They teach whole classes well". Only one report mentioned contributing to spiritual and moral development and this was in the negative: "opportunities are not taken, however, to promote spiritual and moral development" (course graded a 3 on this cell).

Discussion

The modicum of results presented above illustrate that there is considerable variation in the lengths of the inspection reports (from 1683 to 3335 words, excluding identical preface and annex). At the level of individual cells there is also variation in the particular criteria applied and in how cell judgments are expressed. It is particularly noticeable how cell grades can differ markedly, yet statements relating to individual criteria can be virtually identical. Given the complexity of the framework for inspection (fourteen cells, specified by a series of criteria, varying in number from 4 to 11), it is impossible, given the current model of inspection report, to properly distinguish between consistency of application and the loading given to any particular criterion. The judging of cell grades is no doubt a complex matter. As

Wilcox and Gray (1996 p78) comment with regard to school inspection, “we know little about how complex judgments are arrived at and even less about how inspectors actually deal with the numerous sets of criteria of which they are required to take account”.

Courses of initial teacher education are usually inspected by single inspectors. The awarding of a grade four in any cell leads to course closure, a grades three leads to quota uncertainty and a spiral of decline in course viability. More transparency in the system would result if grades for each criteria were made public. As currently implemented, the inspection procedure does not seem a safe way of determining course viability. Further analysis will be made when reports of the remaining higher education providers of secondary mathematics PGCE courses are published during 1999.

References

- Convey, M (Ed), 1997, *The Role of Universities in the Education and Training of Teachers*. London: UCET.
- Gilroy, P and Wilcox, B (1997), OFSTED, Criteria and the Nature of Social Understanding: a Wittgensteinian critique of the practice of educational judgement. *British Journal of Educational Studies*. **45**(1), 22-38.
- Graham, J (1997), A Quality Experience? The TTA/Ofsted quality framework in initial teacher education 1996/7. In: *Initial Teacher Education: TTA/Ofsted quality framework*. London: UCET.
- Lawlor, S (1990), *Teachers Mistaught: training in theories or education in subjects?* London: Centre for Policy Studies.
- Matthews, P, Holmes, J R, Vickers, P and Corporaal, B (1998), Aspects of the Reliability and Validity of School Inspection Judgements of Teaching Quality. *Educational Research and Evaluation* **4**(2), 167-188.
- OFSTED (1998), *Reports of the Round of Inspections of Secondary Mathematics PGCE Courses carried out during 1996/97*. London: HMSO.
- OFSTED (1999), *The Annual Report of Her Majesty's Chief Inspector of Schools 1997/8: standards and quality in education*. London: HMSO.
- OFSTED and TTA (1996), *Framework for the Assessment of Quality and Standards in Initial Teacher Training 1996/97*. London: OFSTED and TTA.
- Tymms, P (1997), *The Security of Inspection of Initial Teacher Training*. Durham: University of Durham. Available on Education-line:
<http://www.leeds.ac.uk/educol/documents/000000382.htm>
- Universities Council for the Education of Teachers (1998), *UCET Evidence to the House of Common's Select Committee Inquiry into Ofsted*. London: UCET.
- Wilcox, B and Gray, J (1996), *Inspecting Schools*. Buckingham: Open University Press.