AUDITING THE MATHEMATICS SUBJECT MATTER KNOWLEDGE OF PRE-SERVICE ELEMENTARY SCHOOL TEACHERS

Tim Rowland, University of Cambridge

This paper reports findings from a project which investigated the mathematics subject knowledge of prospective elementary school teachers in the UK, and how this related to classroom teaching performance. The project was initiated in 1998 in the context of UK government policy to introduce subject content knowledge as an explicit dimension of the ‘standards’ for the award of Qualified Teacher Status in England. The methodology combines qualitative and quantitative approaches. We review here some findings about topics that trainees found difficult and indications that the extent and security of their subject matter knowledge is related to their teaching competence.

Acknowledgement: this paper draws extensively on a paper presented at the Second East Asian Regional Conference on Mathematics Education in Singapore: May 2002. That paper was co-authored with Sarah Martyn, Patti Barber and Caroline Heal, and I have retained the ‘we’ in the original.

This paper is set within the theoretical framework set out by Maria Goulding at the first Nuffield Seminar on Mathematical Knowledge in Teaching, Cambridge, January 2007. The following report sets the scene for subsequent research within the SKIMA consortium which began in 2002.

BACKGROUND

Shulman’s categories of teacher knowledge, and, in particular, his constructs of subject matter knowledge, pedagogical content knowledge and curricular knowledge have been influential in the US and beyond. (Shulman, 1986; Shulman and Grossman, 1988). With regard to subject matter knowledge, a distinction can be made (Schwab, 1978) between substantive knowledge (the key facts, concepts, principles and explanatory frameworks in a discipline) and syntactic knowledge (the nature of enquiry in the field, and how new knowledge is introduced and accepted in that community).

An uninformed perspective on subject-matter knowledge in relation to mathematics teaching might be characterised by the belief that secondary teachers already have it and elementary teachers need very little of it. But there is evidence from the UK and beyond to refute both parts of that statement (e.g. Ball, 1990; Alexander, Rose and Woodhead, 1992; Ofsted, 1994; Ma, 1999). In an earlier study of the planning and instruction of novice secondary school teachers, Steinberg, Marks and Haymore (1985) found that teachers use various coping strategies when they lack content knowledge, including relying heavily on ‘the textbook’ as a convenient source of information. Lack of SMK also affected style of instruction, resulting in the avoidance of discussion and student questions. Ma (1999) presents compelling evidence that the adequacy of elementary teachers’ knowledge of mathematics, for their own professional purposes, cannot by any means be taken for granted. Recent government initiatives to enhance the mathematics subject-matter knowledge (SMK) and pedagogical content knowledge (PCK) of prospective and serving elementary teachers have been taken in a number of countries. The rather direct approach to the ‘problem’ in England is captured by an edict in the first set of government ‘standards’ for Initial Teacher Training (ITT) issued in 1997:
All providers of ITT must audit trainees’ knowledge and understanding of the mathematics contained in the National Curriculum programmes of study for mathematics at KS1 and KS2 [2], and that specified in paragraph 13 of this document. Where gaps in trainees’ subject knowledge are identified, providers of ITT must make arrangements to ensure that trainees gain that knowledge during the course … (DfEE, 1997, p. 27)

The process of audit and remediation of subject knowledge within primary ITT became a high profile issue following the introduction of these and subsequent government requirements (DfEE, 1998). Within the teacher education community, few could be found to support the imposition of the ‘audit and remediation’ culture. And yet the introduction of this ‘testing’ regime provoked a body of UK research on prospective primary teachers’ mathematics subject-matter knowledge (e.g. Rowland, Martyn, Barber and Heal, 2000; Goulding and Suggate, 2001; Jones and Mooney, 2002; Sanders and Morris, 2000, Morris, 2001; Goulding, Rowland and Barber, 2002). The proceedings of a symposium held in 2003 usefully draw together some of the threads of this research (BSRLM, 2003).

OVERVIEW OF GOALS AND METHODS

In the UK, the vast majority of trainees follow a one-year, full-time course leading to a Postgraduate Certificate in Education (PGCE) in a University Education department, about half the year being spent working in a school under the guidance of a school-based mentor. All primary trainees are trained to be generalist teachers of the whole primary curriculum. Later in their careers, most take on responsibility for leadership in one curriculum area (such as mathematics) in their school, but, almost without exception, they remain generalists, teaching the whole curriculum to one class.

In this paper, we describe our approach to, and finding arising from, the audit of the mathematics SMK of 173 primary trainees in 1998-99. This was the first cohort of students following the one-year PGCE course to whom the requirements of Circular 4/98 applied by statute. We had, however, piloted the audit and a draft version of the ‘standards’ on a voluntary basis the previous year.

The structure of the primary PGCE under consideration is such that by the middle of January, with fully six months of the course remaining, the main content areas – number concepts and operations, data handling, mathematical processes, shape and space, measures, algebra, probability – have been ‘covered’ in lectures and workshops. A 90-minute written assessment consisting of 16 test items in mathematics was therefore administered at this point of the course. Four items are shown in an Appendix to this paper. Trainees had been given notice of the ‘test’ and a revision syllabus some six weeks earlier. Their response to each question included a self-assessment of their ability to complete it successfully. The scripts were read by course tutors and the response to each question coded as: ‘secure’, ‘possibly secure’, or ‘not secure’. Corresponding scores of 2, 1 and 0 respectively were recorded for each question and each student. In mid-February, an individual audit feedback sheet was returned to each student, with guidance (where appropriate) for further study. Those students (about 25% of the cohort) who had been found to be secure in 15 or more of the 16 topics audited were invited to become mathematics peer tutors. Following training for this task, they conducted one-to-one peer tutoring sessions with all other students (on average, three per peer tutor) in April, writing a feedback sheet on each of their tutees. Barber and Heal (2003) give further details of the peer tutoring process.

The course includes two extended ‘practicum’ placements in schools in the latter parts of the second and third terms. Given these and other demands of the course, the major SMK remediation opportunity comes between the first and second placements.
Other aspects of the taught course based in the University

<table>
<thead>
<tr>
<th>MathSMK audit</th>
<th>Other aspects of the taught course based in the University</th>
<th>School placement 1</th>
<th>School placement 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term 1 (autumn)</td>
<td>Term 2 (spring)</td>
<td>Term 3 (summer)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The chronology of the PGCE course

During school placements, each student works under the joint supervision of a school-based mentor and a university tutor. The two supervisors agreed on assessments of the student’s performance in teaching mathematics towards the end of (and in the context of) each placement, against the standards of Circular 4/98.

TRainees’ Mathematical Thinking: Generalisation and Proof

One dimension of our research has been to identify what mathematics (within the remit of Circular 4/98) primary trainees find difficult, and the nature of their errors and misconceptions in these areas. In fact, the same audit instrument was administered to a sample of 41 trainees (PGDE and Dip Ed) in Singapore. The profile of easy/hard items is more or less identical for the English and the Singapore trainees (Swee Fong Ng, personal communication). The only differences worthy of note are on two items. The first concerns inverse operations, a topic which the Singapore students found more problematic than those in our sample. It may be significant that this topic is addressed in the algebra component of our taught course immediately prior to the audit. The second was an item on reasoning and proof (discussed later in this paper) on which the Singapore students demonstrated superior syntactic subject matter knowledge. In this brief account, we focus on just two aspects of the UK students’ trainees mathematical thinking.

Generalisation

Just over half the trainees were insecure in an item designed to address the ability to observe and express a generalisation (see Rowland et al 2000 for details).

Check that 3+4+5=3x4 8+9+10=3x9 29+30+31=3x30
Write down a statement (in prose English) which generalises from these three examples. Express your generalisation using symbolic (algebraic) notation.

Some students did not recognise the features common to all three examples and tried to derive a generality from the first case only. For example, one wrote:

Three consecutive numbers added together equals the product of the first two numbers.

\[ n + (n+1) + (n+2) = n \times (n+1). \]

Others were able to express the generality in their own words but not symbolically. Responses such as \( a + b + c = 3b \) captured part of the picture but omitted the essential condition that the three numbers being summed are consecutive (or in arithmetic sequence).

A few students struggled to find the words to communicate what they could ‘see’ in the examples. One wrote:

Three ascending numbers may be equal, in sum, to 2 numbers that are multiplied together. The middle number of the sequence and the overall numbers are multiplied to give the same answer as those added together.
This item, then, exposed weaknesses in recognising and articulating pattern and relationships, identifying significant elements, and in formulating expressions to represent these relationships. This relates to syntactical subject knowledge, since inductive reasoning is central to the philogeny and the otogeny of mathematical knowledge. It is perhaps not difficult to share the concern of the UK government about prospective primary school teachers who, for example, find it so difficult to perceive and communicate unity of form (let alone of meaning) in the three equations. At the same time, we would question the adequacy of “guided self-study” (DfEE, 1998, p. 48) in the face of such cognitive obstacles.

Proof

Currently, there is evidence for concern in the UK about students’ facility with mathematical proof, both at school and at university level (see e.g. London Mathematical Society, 1995). One argument suggests that logical reasoning was a casualty of curriculum and assessment reforms in the 1970s and 1980s. Circular 4/98 requires that trainees demonstrate “that they know and understand […] methods of proof, including simple deductive proof, proof by exhaustion and disproof by counter-example (DfEE, 1998, p. 62)”.

The following item was designed to audit this ‘standard’.

A rectangle is made by fitting together 120 square tiles, each 1 cm$^2$. For example, it could be 10cm by 12 cm. State whether each of the following three statements is true or false for every such rectangle. Justify each of your claims in an appropriate way:

(a) The perimeter (in cm) of the rectangle is an even number.

(b) The perimeter (in cm) of the rectangle is a multiple of 4.

(c) The rectangle is not a square.

More than one mode of justification is possible for each part, but we anticipated some deductive arguments for (a), counterexamples for (b), and perhaps contradiction ($\sqrt{120}$ is not an integer) for (c). In the event, only one third of students made a secure response to the whole question and 30% either gave insecure answers to all three parts or did not attempt the question (see Rowland et al., 2001 for a detailed analysis). Most interesting, perhaps, is the fact that a significant number of students did not seem to perceive the second statement as amenable to personal investigation on their part, which (for those who did so) uncovers counterexamples to refute the statement. Some claimed, for example, that the statement must be true because 120 is a multiple of 4, or even because the rectangles have 4 sides. These prospective teachers evidence little or no sense of mathematics as an experimental test-bed, in which they might confidently respond to an unexpected student question “I don’t know, let’s find out”.

SUBJECT KNOWLEDGE AND CLASSROOM PERFORMANCE

We move on now to data that have enabled us to build on and update our earlier findings (Rowland et al., 2000) associated with the second of our project goals – investigating the relation between trainees’ SMK and their teaching competence. The level of each student’s subject knowledge (based on the audit) was categorised as low, medium or high, corresponding to the need for significant remedial support, modest support (or self-remediation), or none. In addition, assessments of the students’ teaching of mathematics were made (against the standards set out in Circular 4/98) on a three-point scale weak/capable/strong. For the 1998-99 cohort, these assessments were made (a) on both first and second placements, and (b) with respect to both ‘preactive’ (related to planning and self-evaluation) and ‘interactive’ (related to the management of the lesson in progress) aspects of mathematics teaching (following Bennett and Turner-Bisset, 1993). Tables 2 to 5 below show the four 3 by 3 contingency tables, for Placement 1 (N=167: six students had withdrawn from the course) and Placement 2 (N=164: three more students had
withdrawn), together with expected frequencies (in parentheses) based on the null hypothesis that audit performance and teaching performance are independent.

<table>
<thead>
<tr>
<th>SUBJECT KNOWLEDGE AUDIT</th>
<th>TEACHING PRACTICE PERFORMANCE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strong</td>
<td>Capable</td>
</tr>
<tr>
<td>High</td>
<td>17 (12.4)</td>
<td>16 (17.3)</td>
</tr>
<tr>
<td>Middle</td>
<td>31 (29.2)</td>
<td>38 (40.7)</td>
</tr>
<tr>
<td>Low</td>
<td>13 (19.4)</td>
<td>31 (27.0)</td>
</tr>
</tbody>
</table>

**Table 2: Placement 1, preactive**

<table>
<thead>
<tr>
<th>SUBJECT KNOWLEDGE AUDIT</th>
<th>TEACHING PRACTICE PERFORMANCE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strong</td>
<td>Capable</td>
</tr>
<tr>
<td>High</td>
<td>12 (8.1)</td>
<td>18 (14.1)</td>
</tr>
<tr>
<td>Middle</td>
<td>20 (18.5)</td>
<td>33 (32.3)</td>
</tr>
<tr>
<td>Low</td>
<td>7 (12.4)</td>
<td>17 (21.6)</td>
</tr>
</tbody>
</table>

**Table 3: Placement 1, interactive**

<table>
<thead>
<tr>
<th>SUBJECT KNOWLEDGE AUDIT</th>
<th>TEACHING PRACTICE PERFORMANCE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strong</td>
<td>Capable</td>
</tr>
<tr>
<td>High</td>
<td>18 (12.4)</td>
<td>12 (16.1)</td>
</tr>
<tr>
<td>Middle</td>
<td>32 (29.2)</td>
<td>37 (37.8)</td>
</tr>
<tr>
<td>Low</td>
<td>11 (19.4)</td>
<td>30 (25.1)</td>
</tr>
</tbody>
</table>

**Table 4: Placement 2, preactive**

**Table 5: Placement 2, interactive**

Each table has df=4, and values of $\chi^2$ less than 9.5 support the null hypothesis against the alternative that audit performance and teaching performance are in some way linked (p<0.05). The $\chi^2$ values for the preactive and interactive data are 8.2 (p=0.085) and 10.5 (p=0.03) respectively for Placement 1 and 17.8 (p=0.002) and 13.6 (p=0.009) for Placement 2. Thus, the association between audit score and teaching performance is significant for three of the four analyses, the exception being the preactive dimension of the first placement. For the moment, we conjecture that the assessment of preactive aspects favours not only clarity about mathematics teaching and learning, but also a certain kind of bureaucratic competence that acts as some kind of ‘leveller’ in the very first exposure to work in schools. Taken together, however, these results support our earlier findings with the 1997-98 cohort (Rowland et al., 2000) and point to the positive effect of strong SMK in both the planning and the ‘delivery’ of elementary mathematics teaching.

**FURTHER QUANTITATIVE ANALYSIS**

An unpublished independent analysis of the 1998-99 data (Proctor, 2001) incorporates hitherto unexamined variables in an attempt to ‘predict’ the mathematics teaching competence of these trainees. These additional variables are:

1. The gender (male/female) of the trainee.
2. Their chosen age specialism (Early Years 3-8 or Middle Years 7-11)
3. A subject knowledge audit self-assessment. For each audit item, trainees were asked to indicate their degree of confidence on each item, on a 5-point scale from 0 (“Terrifying, I can’t really think about it”) to 4 (“I could show someone else how to do this”).

Variables S, A and TP then denote respectively the sum of the self-assessment scores for the 16 items, the sum of the actual audit scores for the 16 items and the sum of the four teaching...
practice grades (preactive and interactive for the two placements). Thus S lies in the range 0 to 64, A in the range 0 to 32 and TP from 4 to 16.

The account which follows is inevitably selective and omits the technical details to be found in the original. Proctor’s report begins with some exploratory data analysis. This establishes that the normality of A is compromised to the extent that the maximum score (32) constrains the distribution because it acts as a buffer. (We have in fact modified the audit coding system in line with Proctor’s recommendations). Calculation of correlation coefficients shows significant correlation between the three pairs of variables, although scatter graphs indicate that neither A nor S is going to be a useful predictor of TP.

A comparison of audit scores over gender (female F, male M) and course (early years EY, middle years MY) indicates that there is little difference between the four groups F/EY, F/MY, M/MY. The M/EY scores are high relative to the other groups, but the group is very small (2 individuals) and this group is excluded in subsequent analysis.

A comparison of self-assessment scores for the same groups, however, suggests a much higher level of confidence among the MY students irrespective of gender (omitting the M/EY pair).

A similar comparison of Teaching Practice scores points to the superiority of F/MY as a group. The mean TP score of M/MY is the highest (i.e. the worst) and this group contains a long ‘tail’ of weak students as regards mathematics teaching performance.

Comparison of the M/MY, F/EY and F/MY groups by multivariate analysis of variance (MANOVA) is consistent with these indications.

Further MANOVA analysis seeks to assess the relative usefulness of audit score A and confidence score S in predicting the teaching practice score TP, irrespective of gender and course. Specifically, the sample of 153 students who completed both teaching practices were separated into two bands, the upper band consisting of those with a mean TP score of 1 or 2, the lower those with a mean TP score of 3 or 4. The two bands were similar in size (upper 83, lower 70). Borderline averages were rounded down, as university course assessors would tend to do so in practice. The upper band can be viewed as those clearly demonstrating a secure level of teaching competence. The lower band includes those who are merely satisfactory along with those who would be assessed as borderline or unsatisfactory. It turns out that individuals in both bands are represented across the range of audit and confidence self-assessment scores. However, there is evidence for a difference in the mean (A, S) vectors of the high/low bands (Hotellings $T^2$ statistic). Further analysis shows that the contribution of the
audit score A to this difference is much greater than that of the confidence score S.

**Figure 1: 95% confidence regions for the two teaching bands A (high) and B (low)**

This is nicely illustrated in Figure 1, in which the 95% confidence regions of the mean for both A and S are represented by the ellipses $\mathcal{A}$ for the high band and $\mathcal{B}$ for the low band. Although these confidence regions themselves do not overlap, the projections onto the relevant axes show no overlap between the audit score confidence intervals for the two groups compared with a clear overlap between the self-assessment scores.

A final section of Proctor’s analysis uses regression analysis in order to predict the TP score from A and S, together with dichotomous variables for gender G (M=0, F=1) and course Y (MY=0, EY=1). Ordinary least squares (OLS) regression analysis results in the model:

$$TP = 16.7 – 0.166A – 0.0376S – 1.178G + 0.365Y$$

The overall significance of the model is high, and the greater contribution of the audit score A to this difference is much greater than that of the confidence score S. However, the p-value of the course variable Y is found to be 0.35, indicating that there is little justification in including it in the model. The OLS model then becomes:

$$TP = 16.95 – 0.164A – 0.044S – 1.67G$$

An OLS model to predict the sum of the two scores for the second teaching placement only – the one that ultimately matters most in the PGCE course assessment – gives constants roughly half of those shown immediately above, as might be expected, although the contribution of the audit score and gender is even greater.

Perhaps the most startling result comes from a comparison of separate OLS models for the Early Years (excluding the two males) and Middle Years students.

For the Early Years trainees:  
$$TP = 14.5 – 0.0646A – 0.0818S$$

For the Middle Years trainees:  
$$TP = 16.8 – 0.233A – 0.0045S – 1.82G$$

This indicates that for the MY group, the predictive effect of the self-assessment score is negligible compared with that of the audit score and the gender effect, whereas for the EY trainees confidence is the more powerful predictor. (This may not be self-evident since the coefficients of A, S in the EY model are in the ratio 3:4 approximately. However, it must be remembered that the range of values of S is twice that of A. In fact, the correlation (Pearson’s) between both TP and A and TP and S is significant ($p<0.05$)).

For further discussion of Proctor’s analysis, see Rowland (2001).

**CONCLUSION**

We have drawn attention to the problematic nature of generalisation and proof as a component of the mathematics SMK of pre-service elementary teachers, adding further weight to the doubts of Goulding and Suggate (2001) that much can be done to remedy trainees’ difficulties with proof within initial training, especially given the multiple demands on them in all areas of the curriculum in an intensely pressured course. We would expect that clarity of understanding of the nature of proof and refutation in mathematics would inform the trainees’ approach to questioning and enquiry with their students, and we are struck by the robustness under replication of our earlier finding (Rowland et al., 2000) that effective classroom teaching of elementary mathematics is associated with secure SMK at a level beyond the elementary curriculum. It may be that, even within the constraints of PGCE
courses, greater priority could be given to syntactic dimensions of SMK, although inevitably this would be at the expense of substantive elements.

Proctor’s (2001) analysis could be interpreted in a number of ways. One conjecture would be that, for teachers of very young children, the actual audited SMK is less important than a confident attitude to mathematics, although this would be in conflict with the findings of Aubrey (1997). Perhaps the major contribution of quantitative analysis, in this case, was to motivate the next phase of our work, which involved qualitative analysis of classroom observations, in an attempt to find out ‘what’s going on’ – how subject knowledge could be seen to come into play when these trainees teach elementary mathematics.

Methodological/Epistemological Comment

It may be important here, however, to say that that the Shulman-inspired conception of teachers’ SMK and PCK is essentially cognitivist. Put crudely, these aspects of teacher knowledge are perceived to be relatively stable, explicit, portable from place to place, and generalisable to different situations. We recognise that it may leave unanswered some questions about teachers’ decision-making and activity in the workplace. Situated perspectives on knowledge, in particular, may offer additional, or competing, insights into these matters. Adler writes, for example, that “Knowledge about teaching is … not simply in individual teachers’ heads: it is tied to their identities and evolves in and through co-participation in the practices of the teaching community.” (1998, p. 168). Drawing on a four year longitudinal study into professional change (Johnson, Hodgen and Adhami, 2004), Hodgen (2007) describes a teacher, Alexandra, who demonstrates ‘more’ knowledge in practice than she is able to in an interview with a researcher. Hodgen argues that her professional knowledge is best understood, not only as situated, but also as social and distributed, and that such a perspective “places greater emphasis on the communities in which mathematics teachers are engaged rather than on individual knowledge” (p. 7). If one were to take the view that individualist enquiries into teacher knowledge and classroom practice are fatally flawed, then attempts to investigate them without reference to social context will be viewed with suspicion.

REFERENCES


Appendix:

PRIMARY PGCE: EXAMPLE ITEMS FROM THE 1999 MATHEMATICS AUDIT

Item A
Work out 39 + 127 mentally.
Briefly describe your mental strategy. Point out whether (and where) it makes use of the commutative and associative properties of addition.

Item B
A rectangle is made by fitting together 120 square tiles, each 1 cm². For example, it could be 10 cm by 12 cm. State whether each of the following three statements is true or false for every such rectangle. Justify each of your claims in an appropriate way:
(a) The perimeter (in cm) of the rectangle is an even number.
(b) The perimeter (in cm) of the rectangle is a multiple of 4.
(c) The rectangle is not a square.

Item C
Check that 3 + 4 + 5 = 3 × 4 8 + 9 + 10 = 3 × 9 29 + 30 + 31 = 3 × 30
Write down a statement (in prose English) which generalises from these three examples. Express your generalisation using symbolic (algebraic) notation.

Item D
(a) Is pi (π) equal to 22/7?
(b) Explain the claim that the number pi (π) is an ‘irrational’ number?
(c) Is 0.6666... (recurring) an irrational number? Explain your answer.