INTRODUCING MORE PROOF INTO A HIGH STAKES EXAMINATION – TOWARDS A RESEARCH AGENDA

John Monaghan & Tom Roper

University of Leeds

We raise issues in examining what 'proof' is in high stakes examinations and what research issues may be usefully explored in considering 'proof in high stakes examinations'. Our considerations arise from practical 14-19 curriculum and assessment development work. This article is not a standard academic article but, rather, a collection of issues and ideas.

INTRODUCTION

> If a student has gone through his mathematics classes without having really understood a few proofs like the foregoing one[i], he is entitled to address a scorching reproach to his school and to his teachers. Polya (1945, 216)

The context within which we consider 'proof in high stakes examinations' is outlined in this section. In subsequent sections we consider: what is proof in high stakes examinations; problems in setting proof questions; research issues and opportunities in ***14-19 Pathways*** work; and 'players and artefacts' in high stakes examinations.

The influential Smith Report (Smith, 2004) made a number of recommendations. Two, of relevance to this article, are: a call for the commission of curriculum and assessment development studies for 14-19 mathematics (ibid., Recommendation 4.11); a call for (academic stream) students to engage more fully with proof (ibid., Recommendation 4.5). We were part of a University of Leeds team who developed a curriculum and assessment study (14-19 Pathways, phase 1) and are now, with the examination board AQA (14-19 Pathways, phase 2), trialling[ii] and preparing to pilot curricula and assessment for courses of study and examination from 2011.

In our phase 1 recommendations on proof at GCE (A-level) we noted that proof is not currently a significant feature in A-level. Syllabii referred to 'knowing' specific proofs, e.g. 'knowing' the proof to obtain the sum of an AP and GP and proving trigonometrical identities. In the examination papers the words 'proof' or 'prove' were often a call for symbolic manipulation, e.g.

> Given that $x = \tan\frac{1}{2}y$, prove that $\dfrac{dy}{dx} = \dfrac{2}{1+x^2}$.

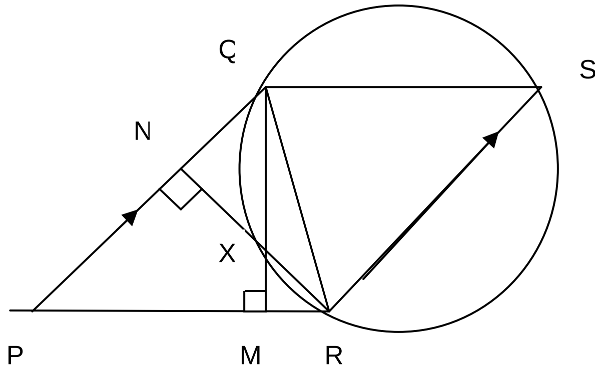
Our phase 1 report[iii] to QCA regarding proof at GCE recommended greater emphasis be given to proof and that this be reflected in more questions being set that involve proof and 'proof-theoretic ideas' in examinations equivalent to current C3/C4 examinations. This recommendation was incorporated into current phase 2 work with AQA and we have trialled (informal development and testing) proof questions with students and are preparing to pilot (full curriculum with examinations leading to a formal qualification) from September 2007.

WHAT IS PROOF IN HIGH STAKES EXAMINATIONS

The inclusion of proof in high stakes examinations is not new. It has been done before and it is worth considering what proof looked like in public examinations in the past. Whilst the curriculum under discussion is A-level, questions such as that below appeared in O-level (Joint Matriculation Board Syllabus B, November 1979 Paper II B9).

In the acute-angled triangle PQR, the point M is the foot of the perpendicular from Q to PR. The perpendicular from R to PR cuts QM at X and meets PQ at N. A circle is drawn through Q, X and R. The line through R parallel to PQ meets this circle at S. Prove that:

(i) quadrilateral PNXM is cyclic,
(ii) ∠QPR = ∠QSR,
(iii) ∠PRQ = ∠RQS,
(iv) PQRS is a parallelogram.



In this particular question the candidate is required to marshal his/her material and set it out in the form of a logical argument structured by a traditional format.

Proof at A-level was sometimes not dissimilar to what we find today:

> The first and last terms of an arithmetic progression are *a* and ℓ, respectively. If the progression has *n* terms, prove from first principles that its sum is $\tfrac{1}{2}n(a + \ell)$

This example (JMB Syllabus A June 1973, Paper 1 S1 (part)) would be standard bookwork of the kind still done in the classroom today; it would have been rehearsed via past papers with the aim of perfect reproduction in the examination.

A second example (JMB Syllabus A June 1973 Paper 1 S8 (part)) however is:

> Given that *u* and *v* are functions of *x*, prove from first principles that
>
> $$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\,du/dx - u\,dv/dx}{v^2}.$$

It is difficult to see exactly what kind of proof might have been expected for this example, perhaps the kind of demonstration given in text books and accompanied by a great deal of hand-waving by the teacher. Further the instruction "prove" can be seen to possess at least two synonyms in relation to the same piece of mathematics. Thus it is not difficult to find questions which say:

> Prove that the equation of the tangent to the curve…. at the point…
>
> Show that the equation of the tangent to the curve…. at the point....
>
> Calculate the equation of the tangent to the curve…. at the point….

Are prove, show and calculate really all the same? In many respects the answer must be 'yes' since much of what we prove in mathematics is to enable the automatic and safe calculation of certain quantities. For example, relevant theorems of analysis combine to tell us that it is safe to write that the derivative of $y=x^2$ is $dy/dx=2x$, $x\in\Re$. If, then, in doing a calculation such as finding the tangent to a curve we assume the relevant theorems of analysis, then the calculation is, in one sense, a proof of a particular property of the curve; yet we do not ordinarily think of calculation as proof. What is a proof and what is not, and the role that proof has played in mathematics, is not without dispute, "there long have been and still are conflicting opinions on the role of proof in mathematics and in particular on what makes a proof acceptable." (Hanna, 2000, p.5)

PROBLEMS IN SETTING PROOF QUESTIONS

Nothwithstanding the problems of characterising what a proof is, including proofs in high stakes examinations raises further problems. Perhaps the most important issue, simply noted in this article, is that proof may not be suited to high stakes examinations where students are under time and other pressures to 'show what they have learnt'. This important issue does not appear, in our reading of the mathematics education literature on proof, to be considered. The issue does not appear to be a straightforward case of 'you cannot produce an adequate proof in a timed examination' because, as Buxton (1981, p.89) notes, "... time pressures are not automatically bad. Some of us work better under time pressures ...".

An issue in setting proof questions is 'who sets them?' In our phase 2 work questions were set by a one of us (a mathematics education researcher) and the AQA Chair of Examiners. This was a respectful collaboration but there is an argument that neither 'player' in this activity was entirely suited to the task. In terms of their past histories and current briefs towards the trialling the mathematics education researcher was inexperienced in setting examination questions and the Chair of Examiners was influenced by matters other than just proof, e.g. developing workable markschemes, syllabus coverage and setting questions that are accessible to students at various expected grade levels. We return to such matters in the section 'players and artefacts' below but, for now, note that there is a sense in which the trial proof questions developed were a 'compromise' of the objectives of these two players.

In the BSRLM presentation we presented four ideas which had accompanying problems: use trigonometric identities; use definitions; first principle calculus ideas; use 'easy' content knowledge. Space does not allow for a consideration of all of these in this article and so we focus on one, 'easy' content knowledge.

In our work *14-19 Pathways* work on GCSE problem solving and on Functional Mathematics the Leeds team developed a rubric of *stepping back in terms of content*, i.e. using prior, possibly easier (to the student in terms of their development at the time of assessment), topic knowledge. Our evidence, not reported in this article, was that this was a largely (but not wholly) successful strategy in setting questions. It does not, however, appear to be as successful in setting proof questions. Consider the following:

Given that $f(x) = \frac{x}{x-1}, x \neq 1$, prove that $f^{2n+1}(x) = f(x)$, n a whole number.

In terms of content knowledge this does not require anything beyond Higher level GCSE but it was, for students, a difficult question. One reason for this difficulty is that the algebraic manipulation required is 'fiddly'. But there were students who did not appear to have problems with the algebraic manipulation but who could not produce a proof. We cannot state with certainty reasons for this. It may be that students had little experience in producing proofs. It may be, however, and we put this forward as a roughly formulated hypothesis, that *stepping back in terms of content* is not an effective strategy in producing proof questions.

ISSUES IN INTRODUCING MORE PROOF INTO EXAMINATIONS

Of the many issues concerned with introducing more proof into an examination we raise: the preparation of students; ethics; trialling in the present for a future examination; questions about teaching.

An important issue to us is what might be called 'the curricular preparedness' of A-level students for proof. This has at least two related dimensions – geometry (an important area of mathematics with regard to proof) and induction vs deduction. Most English students study GCSE mathematics prior to their A-level studies. Since the introduction of the National Curriculum in 1988, GCSE mathematics has included little deductive geometry and a considerable focus, in general, on inductive reasoning (see also Healy & Hoyles' (1999) comments on proof, on geometry/algebra and on induction/deduction). A reasonable question, then, is it 'fair' to impose more proof on English A-level students?

This is an ethical question and a further ethical issue is carrying out pilot work in the context of high stakes examinations where students must, for their future careers, perform to the best of their abilities. All parties we work with on the development of this project regard students' grades as paramount but this, in turn, generates a certain 'conservativeness' with regard to the introduction of more proof into examinations. NB Schools/students following the pilot A-level can, if they want take both the old and the new examinations if they are concerned about grades.

Another issue which raises ethical, as well as practical problems, concerns trialling examinations questions for a future course of study on students not following this course of study. It is essential that proof questions developed by the AQA-Leeds team are tested on students but we expect that these students are going to do less well on these questions than future students following the pilot course because future students should be more prepared for proof questions in their studies. With regard to ethics we risk making students feel bad about their mathematical prowess if they cannot make a reasonable attempt at the question and, with regard to practical matters, we are left guessing what a certain correct response rate now might mean for correct response rates of future students.

There are a number of issues/questions concerning teachers/teaching and proof. How does one ‘teach’ proof? What approaches will teachers on pilot courses adopt? What CPD can/should be provided for these teachers? How can we evaluate these questions? As difficult as these questions are to address it is crucial that they are addressed so that information can inform future CPD for all teachers if the pilot course becomes the new GCE.

‘PLAYERS AND ARTEFACTS’ IN HIGH STAKES EXAMINATIONS

In the section on setting proof questions above we mentioned an issue concerned with two ‘players’ (a mathematics education researcher and the AQA Chair of Examiners) working together. This was not an issue concerned with these two players ‘not getting on’ (they both felt they worked well together) but of the players coming from different communities of practice with different objectives and and rules (one may consider these issues within discourse on ‘communities of practice’ or ‘activity systems’ or ‘actor network theory’ but we are not concerned with this level of detail at this stage in our thoughts). We now take a wider consideration of players and artefacts.

There are a number of players and artefacts involved in the production of a high stakes examination (in this case a pilot GCE which includes proof questions). Some of these are:

- communities, e.g. LMS, IMA, ACME, AQA, DfES, QCA;
- individuals with specific responsibilities within these communities;
- artefacts including reports from communities and/or individuals, e.g. the Smith Report, and GCE ‘specifications’ including aims, objectives, schemes of assessment and subject content.

There are also relationships between these players and artefacts, e.g. it is clear to us that the DfES has a strong influence on QCA (but no reverse influence is apparent), that QCA demands artefacts such as specifications from examination bodies such as AQA and that AQA regards the production of the specification as a major activity (and these specifications are ‘boundary objects’ between different communities).

What is clear to us is that the introduction of more proof into GCE mathematics is not the product of the work of the mathematics education researcher and the AQA Chair of Examiners but the product of a network of players and artefacts. Researching the interrelations in this network in the development of this new GCE, which includes more proof questions in its examination, would be a difficult undertaking but one that could shed important light on how ‘proof in a high stakes examination’ is created.

POSTSCRIPT

Many of the issues highlighted above would not have occurred prior to 1988, the introduction of the National Curriuclum and its associated testing regime. Prior to 1988 the introduction of more questions associated with proof would have been at the whim of the examination boards, perhaps prompted by a local or national curriculum development initiative or the predilections its chief examiner. However, whilst a highly regulated and high stakes system such as exists in England offers a degree of openness and a clear entitlement, curriculum development is then fraught with difficulty since the knock-on effects throughout the system of even the smallest change have to be considered. The introduction of more questions focusing on proof might appear to be relatively straight forward; such questions surely reflect the nature of the subject? Clearly they do, but an important question is, can such a development be made to fit within the system without challenging any part of the system? There appears to be no simple answer to this question.

NOTES

[i] The "foregoing one" being *that the sum of the three angles of a triangle is equal to two right angles.*

[ii] 'Trialling' is UK education-speak for developmental testing of materials and 'piloting' refers to the development of a full curriculum and assessment together with an award.

[iii] This does not, unfortunately, appear to be in the public domain.

REFERENCES

Buxton, L. (1981) *Do You Panic About Maths?* London: Heinemann.

Hanna, G. (2000) 'Proof, explanation and exploration: an overview'. *Educational Studies in Mathematics*, 44, 5–23.

Healy, L. & Hoyles, C. (1999) 'Student's performance in proving: competence or curriculum?', in I. Schwank (Ed.) *Proceedings of the First Conference of theEuropean Society for Research in Mathematics Education.* Available from http://www.fmd.uni-osnabrueck.de/ebooks/erme/cerme1-proceedings/cerme1_contents1.html

Polya, G. (1945) *How to Solve It.* New York: Anchor books.

Smith, A. (2004) *Making Mathematics Count*. London: HMSO.