

Why do GCSE examination papers look like they do?¹

Ian Jones

University of Nottingham

GCSE mathematics examinations have been criticised for being too structured and not adequately assessing process skills. Exam papers are produced by private awarding bodies working to government regulations. A given paper and its mark scheme is usually written by one individual, and then reviewed and revised by a Qualification Paper Evaluation Committee (QPEC). As such, QPEC meetings have a significant role in how the final published exam paper looks. Over the past year I have observed several QPEC meetings across the three awarding bodies that publish GCSE papers. I describe how what gets said in QPEC meetings bears on the structure and process skills assessed in exam papers.

Keywords: Assessment, GCSE, examinations

Background

GCSE mathematics is assessed by exam papers published and marked by private, competing awarding bodies. Several recent reports have questioned the fitness for purpose of current GCSE mathematics exams (ACME 2007; EMP 2008; NCETM 2009; Ofsted 2008). The items in exam papers are piecemeal requiring only procedural mathematics and short chains of reasoning (EMP 2008). High-stakes assessment such as GCSE has a large impact on classrooms, and so teaching and learning likewise tends to be fragmentary and procedural (Ofsted 2008). There have been calls for the design processes of GCSE exams to be changed to improve this situation (ACME 2007; NCETM 2009).

Over the past few months I have observed meetings at three English awarding bodies in order to understand why GCSE exam papers look like they do. This has taken place during a time of flux in the regulation of exams. New regulations for GCSE mathematics have come into force as of September 2010, and several new mathematics GCSE qualifications are being or have recently been piloted. This flux has shaken the awarding bodies' processes somewhat, thereby making them more observable, and so it has been an opportune time to research them.

The next section describes current GCSE exam papers by way of an example paper, and the results of an analysis of pilot GCSE papers (EMP 2008). I then briefly describe the government regulations, before focussing in some detail on the internal processes of the awarding bodies that publish and mark exam papers.

What do GCSE papers look like?

The following two questions are from a GCSE practice paper published by one of the awarding bodies that conforms to the new government regulations for September 2010. I offer them as illustrative and representative of the kinds of questions that populate GCSE papers more generally. The first question from the paper is shown in Figure 1a. (The format of the questions has been compacted to save space.) Candidates are told what to do (“reflect”) and

¹ This research was conducted as part of a Royal Society Shuttleworth Education Research Fellowship. The author would like to thank the awarding bodies and examiners involved for their openness and hospitality.

given a diagram to complete. The question is worth one mark for a correct answer. The least structured question, worth a total of six marks, is shown in Figure 1b. It is set in context, and candidates must translate each advertisement into a calculation and find the smallest number. Readers interested in further examples are directed to the awarding bodies' websites where past and practice papers are published.

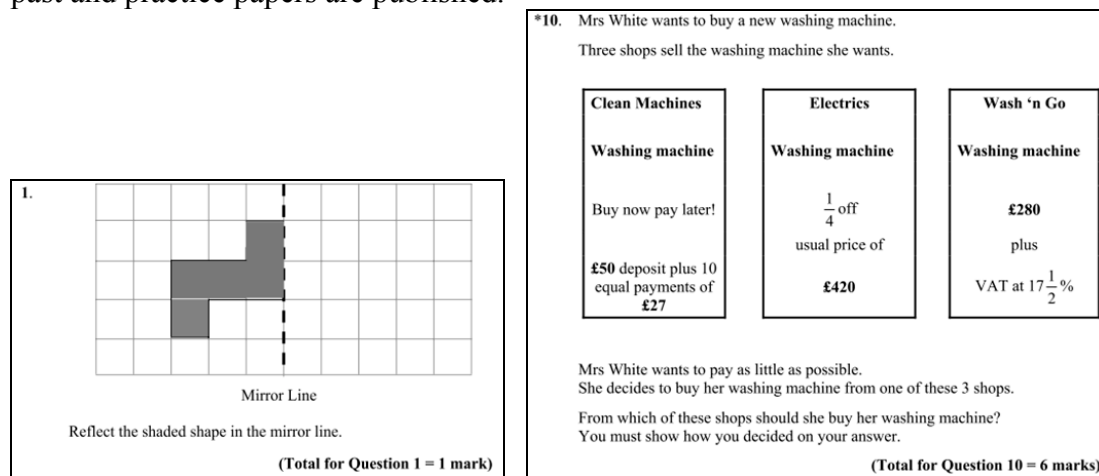


Figure 1(a): The first question from the paper. Figure 1(b): The least structured question from the paper.

An analysis of pilot papers written to the new regulations was undertaken by the Evaluating Mathematics Pathways project (EMP 2008). EMP's purpose was to analyse differences between the then current GCSE papers and the new pilot qualification that has now come into force. It found the new pilots contained more questions set in context (e.g. Figure 1b), and were slightly less structured. Overall, however, GCSE mathematics papers, including those written to the new regulations, are highly structured and contain typically around 80% one or two step items (e.g. Figure 1a). An analysis of the "process skills" required by GCSE items showed an over-emphasis on procedure in pure and artificial contexts, with little opportunity for interpreting, communicating and representing mathematics (Figure 2). In sum, the change in regulations had only a minimal impact on the design of exam papers.

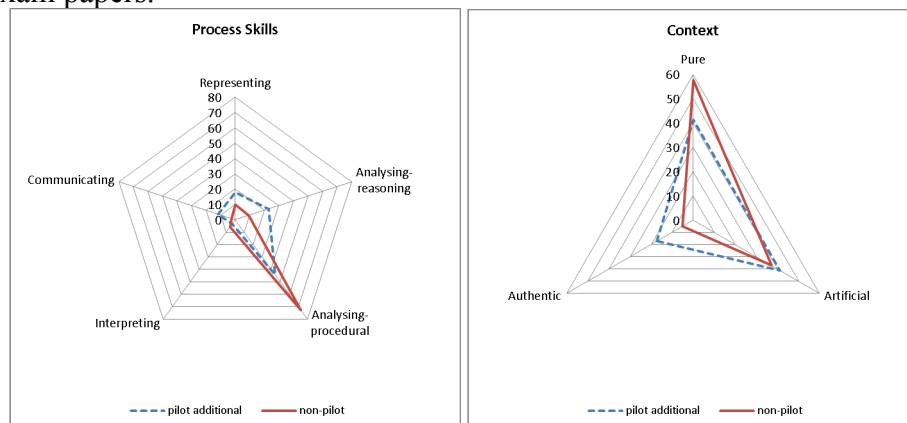


Figure 2: Process skills and context in one awarding body's live and pilot additional mathematics GCSEs for 2009.

How are GCSE papers made?

GCSE papers assess the Programme of Study for Key Stage 4 mathematics, and conform to government regulations. Readers interested in the details of the regulations are directed to the QCDA website. Particular to the new regulations are the three process-based "Assessment Objectives" (AOs). Around half the questions in a paper must be recall and application of

knowledge (AO1, e.g. Figure 1a), and the other half must be set in context, or involve problem solving, or both (AO2 and AO3, e.g. Figure 1b). Some of the questions must involve “functional elements” of mathematics (for a discussion of functional mathematics, and its ambiguities, see Roper, Threlfall, and Monaghan 2005)

For each GCSE the awarding bodies submit content specifications to the government regulators for accreditation. A small excerpt from a specification document is shown in Figure 3. Interested readers are directed to the awarding bodies’ websites to see their specification documents.

Ref	J517 ref	Subject content – Candidates should be able to...	Notes and examples
Number			
HIN1	N6.4	Multiply and divide simple fractions. Add and subtract mixed numbers.	This does not include multiplication and division of mixed numbers.
HIN2	N5.5	Express one quantity as a fraction or percentage of another.	
HIN3	N5.4	Increase and decrease quantities by a percentage.	

Figure 3: A small excerpt from an awarding body’s specification document.

For each exam paper that is written the awarding bodies produce an assessment grid to help ensure papers are balanced and conform to government regulations and the awarding bodies’ own specifications. Assessment grids vary across specifications and awarding bodies, and an excerpt is shown in Figure 4. Each row is a question, and the columns show its content (Number; Manipulative Algebra; non-Manipulative Algebra; Geometry; Statistics), processes (AO1; AO2; AO3; Quality of Written Communication; Functional elements) and difficulty (grades D-A*).

Q	Topic	N	MA	nMA	G	S	AO1	AO2	AO3	QWC	Func	D	C	B	A	A*
1	Ratio	3					3					3				
2	Sum of digits	3							3			3				
3	Depreciation			5			5				5	5				

Figure 4: An excerpt of an assessment grid showing the content, processes and difficulty for each question.

A given exam paper and its mark scheme is usually written by a single person (the “setter”). The setter completes and refers to the assessment grid during this process to ensure balance across the paper. Setters work to tight deadlines and are usually practicing or retired teachers and are often experienced exam writers. Accordingly they draw on their knowledge of past papers and questions, often recycling questions literally or with some tweaking. The first draft of the paper is then reviewed by other examiners and feedback sent back to and incorporated by the setter. Finally the paper is tabled at a meeting of a Question Paper Evaluation Committee (QPEC) for further review and revision. The whole process is iterative and a given paper may be QPECed (to use an awarding body verb) more than once. The process described in this paragraph is illustrated in Figure 5. Note that GCSE questions are not trialled before final publication.

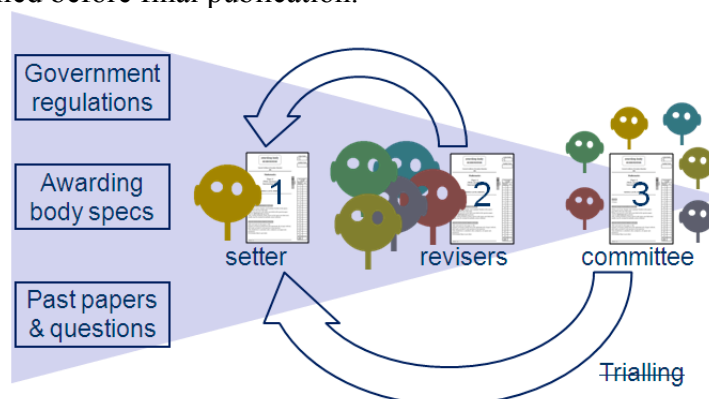


Figure 5: The iterative exam paper writing process.

QPEC meetings

Over a period of ten months I attended twelve QPEC meetings at three awarding bodies. Each meeting lasted a day, during which typically two or three papers were reviewed. They were attended by up to a dozen people, mostly examiners, including the setter and revisers for each paper, and chaired by a principle examiner. For each paper, revisers first provide some overall feedback about the balance, difficulty and appropriateness of the paper. The QPEC then works through the paper question by question, checking everything from punctuation and formatting to mathematical content and coherence. This question proofing is demanding work and takes up the bulk of time and energy in QPECs. The assessment grid is checked and often amended for each question, and then checked again at the end of the paper for balance across content, processes and difficulty. The next paper is then tabled.

The meetings are fast, efficient events in which a lot of work is achieved in short amounts of time. They are somewhat intimidating events for setters as their papers are picked apart and they have to decide when to concede and when to defend their work. Things sometimes get heated, and personalities clash at times, but QPEC meetings are generally underlined by a sense of comradeship and good humour.

In the remainder of this article I will report on some of the things I have heard said in QPEC meetings. My role was as a non-participatory observer, and I made handwritten notes made up mostly of verbatim quotes. The fast-paced nature of meetings meant it was not possible to get everything down that was said, so the resulting notes were a somewhat idiosyncratic sampling. Moreover, as I attended more meetings I became more orientated towards the nature and pace of QPEC meetings, and so more attuned to getting down the more insightful incidents and comments.

I present here some comments and incidents from QPEC meetings that bear on the structure and (lack of) process skills assessed by GCSE papers. This narrow focus cannot and is not intended to do justice to the range and richness of discussion in QPECs. Analysis is ongoing to the broader aim of capturing and reporting the full impact of QPECS on exam papers.

What gets said in a QPEC?

Much of what is said in QPEC meetings relates to the balance of content, mathematical processes and difficulty across papers, as recorded in assessment grids. Overall, but by no means always, this focus on balance tends to drive questions towards being structured and procedural. For example, after all the questions in one paper had been reviewed and the assessment grid was being checked, the committee decided the paper was short of AO1 and AO2 marks, and had too many AO3 marks. The committee went back and reduced the number of AO3 marks (i.e. questions like Figure 1b) and increased the number of AO1 marks (i.e. questions like Figure 1a) and AO2 marks (AO2 questions lie somewhere between Figure 1a and 1b). In other words, the paper was fragmented a little by the QPEC process in order to ensure it conformed to government regulations. Most examiners seem to find this fragmentation process frustrating, and even absurd. In one QPEC a reviser said “It seems odd we have to have so much AO1 on an ‘Applications’ paper”, to which the QPEC chair replied “That’s the spec. Must have half the marks on recall of knowledge. Otherwise centres won’t do it.”

Government regulations are somewhat subjective. The EMP analysis (see above), conducted largely by academics, found that papers over-emphasise factual recall and application (“AO1”), but this is incongruent with the common QPEC judgement that a given paper needs more AO1 items to fit the regulations. Neither are examiners themselves in

unison about the meaning of regulations, however, and QPECs often robustly debate what they mean. For example:

Chair: [Someone else's] paper had the structure. When it didn't have structure it was AO2, AO3.

Examiner 1: Part a has no structure and that's AO1.

Examiner 2: For me AO1 is not so much a question of structure.

This subjectivity enables examiners to adjust assessment grids rather than the questions when the balance of a paper is judged to lie outside of the regulated balance. This is notably so for the three new Assessment Objectives and sometimes an item will be re-allocated from one Assessment Objective to another without any amendment just to balance the grid. As one examiner said of a question: "Don't mind muddling it up between AO2 and AO3 do we?" This is somewhat analogous to drawing targets around bullet holes, and enables awarding bodies to produce similar looking papers to different looking regulations.

Another constraint examiners work under is the need to ensure exam papers are teachable and learnable within a competitive market. As one QPEC chair said of the new regulations, "I'd imagine every Awarding Body will make AO1 as high as possible [up to 55%] because it's straightforward, and we ought to be setting a similar standard. Because then it's more accessible." Another examiner said of an unusually unstructured "Applications" pilot item "It's something great for kids in a lesson for 45 minutes, but it's not a 6 minute exam question". This constraint can have a fragmentary effect on the content of a paper, with QPECs often looking to increase the number of stand-alone easy questions, or add more structure to questions. This is usually done with genuine concern for "weaker candidates".

The pressure to maintain high marking reliability can also act to increase the structure of questions. As one examiner said of a relatively unstructured question, "Last time we set a question like this we had problems with marking. It needed five possible mark schemes just for two items. This has three items and should be scrapped because we can't reliably mark it ... This is five marks and most we can have for this sort of question is three marks."

Despite this, GCSE pilot exams do offer opportunities for examiners to set less structured, more process-oriented questions to a limited extent. One examiner expressed concern that a pilot "Applications" paper was too similar to existing papers: "We'll end up destroying this pilot. If we put all the structure back in it's not an application." The production of "Sample Assessment Materials" and "Practice Papers" for pilots is seen as an opportunity to send messages to teachers that assessment is changing. For example, one examiner said of a finance question that would be unfamiliar to teachers "That's good. We need to ensure such questions are in the mocks so teachers think 'I better do that'". However, this effect towards more open, less predictable item types is limited as awarding bodies do not want to lose market share by scaring teachers off when publishing sample assessments.

Discussion

GCSE mathematics exam papers are highly structured and procedural. They are produced by awarding body examiners working to government regulations. In this paper I have reported some of things said and done in Question Paper Evaluation Committee meetings that bear on the structure and proceduralism of exam papers.

Examiners produce papers using their knowledge of past papers and their understanding of the (new and changing) regulations. In particular they need to ensure a balance of content, Assessment Objectives and difficulty across a paper. In QPECs it is commonly judged that a paper contains too few structured, procedural (i.e. AO1) items and that their number needs to be increased. However, regulations are also somewhat subjective

allowing examiners to re-allocate questions within the assessment grid without amending them. It is perhaps unsurprising that exam papers across different regulations tend to look only slightly different.

Examiners are also under pressure to produce predictable papers that will not scare teachers away to competitors, will not be unfairly harsh on low achieving candidates, and that can be marked with a high degree of reliability. This stifles innovation and concerns for validity. However, pilot qualifications give examiners some limited scope for innovation and to send messages to teachers that assessments will be less structured and more contextualised than previously. For many examiners such pilot qualifications are a welcome opportunity to try and produce slightly more creative, less predictable exam papers.

References

- ACME. 2007. *Position statement on functional mathematics*. Advisory Committee on Mathematics Education.
- EMP. 2008. *Evaluating Mathematics Pathways project: Interim report stage 3*. Nottingham, University of Nottingham.
- NCETM. 2009. *Mathematics matters final report*. London, National Centre for Excellence in the Teaching of Mathematics.
- Ofsted. 2008. *Mathematics: understanding the score*. London, HMSO.
- Roper, T., J. Threlfall, and J. Monaghan. 2005. Functional mathematics - what is it? *Proceedings of the British Society for Research into Learning Mathematics* 25: 91–96.