Setting objective tests in mathematics with QM Designer

Martin Greenhow
Dept of Mathematical Sciences, Brunel University

martin.greenhow@brunel.ac.uk

This paper describes many of the lessons I have learned during the development of the "Mathletics" suite of online tests as recently reviewed [1]. Mathletics was written using Question Mark Designer [2,3,4,5], but many of the remarks below will be applicable to other online tests, especially those written using the closely-related QM Perception [6], and even to paper-based tests.

The initial objective was to deliver diagnostic tests to large numbers of freshers (typically, 200 students take 2-4 tests during Induction Week), give feedback to each student, advise on remedial action to address individual weaknesses, and, importantly, give the level 1 calculus and algebra lecturers a whole-class view so that common errors could be tackled in lectures early in the module [7]. This use soon expanded into regular formative, and even summative, testing, whereby some worksheets were replaced by online tests. Looking at the answer files showed that some (usually better) students were also using the tests as a self-study tool and revision aid before exams.

These answer files and student feedback at Brunel University suggest that, as part of the overall course design, objective testing with computers can be valuable in motivating and monitoring students. It is, however, important to realise not all parts of a module can be tested in this way; objective tests clearly lend themselves to factual recall and direct calculation where there is a single, or several, clearly defined-answers [8]. In mathematics this points to testing of what is often called methods at level 1, much of the A-level core material and almost all of the GCSE mathematics syllabus. What this range has in common is that one is interested in testing a *single defined skill* rather than synthesis of several skills, modelling of a problem, devising a strategy for solving a problem and, importantly, proof.

Partial fractions as topics	Partial fractions as skills
Denominator has 2 linear factors	***Assumed skill: student can factorise the denominator***
Denominator has linear and quadratic factors	Can student recognise quadratic and repeated factors?
Denominator has repeated linear factors	Can student choose correct form for the partial fraction?
Denominator has repeated quadratic factors	Can student set up equations for solving coefficients?
	Assumed skill: student can solve simultaneous equations.
	Can student check his/her answer?

Table 1. Partial fractions as topics/skills

It is important to realise that an objective test actually tests a *skill* rather than a *topic* and a test-setter will need to identify which skill is being tested and which other skills are assumed, see table 1 for an example. This is quite unlike setting a traditional worksheet on a topic that usually spans a range of skills and levels of difficulty with questions getting harder towards the end. Moreover, a traditional worksheet usually assumes a marker using his/her experience to detect missing knowledge or misunderstanding of concepts. The point is that online objective tests need to codify tested and

Supplier's contact details

Name
Address

email address

prerequisite skills *a priori*. Each tested skill will then result in a separate library, so that covering an area often turns out to be a bigger job that originally thought. Nonetheless, without doing this, it is impossible to devise effective wrong answers (distracters) in a multi-choice question, still less provide useful feedback to a student who chooses them, and it is difficult to identify why a student is failing from examination of his / her answer files and to recommend remedial action.

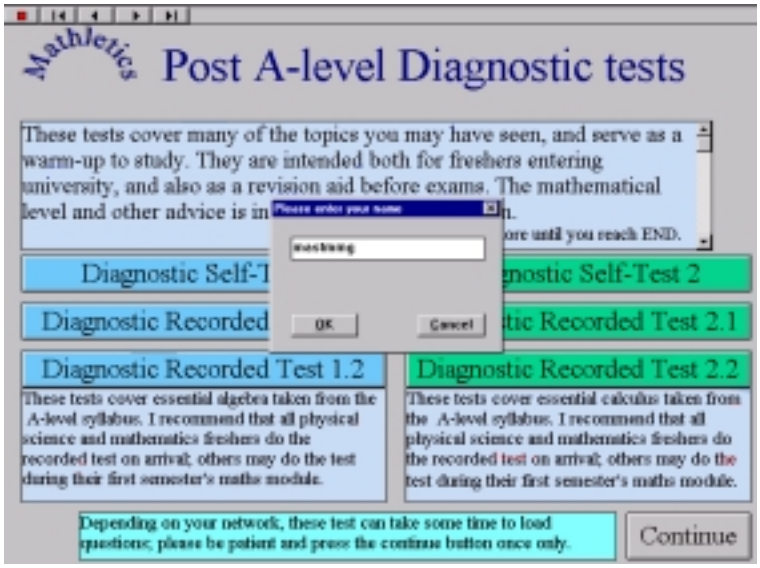

Fig 1: Part of the Mathletics shell

It is clear that a programme of objective tests can only go so far for any taught module; students still need traditional problems sheets marked by lecturers, to cover the higher-level aims of the module and, at a more mundane level, give the student practice in setting out correct mathematics when faced with a blank sheet of paper, as they will be in the exam. I see online objective tests, such as Mathletics, as being a good foundation on which to develop these higher aims, in the same way as weight training will help a sprinter run faster.

This makes such tests ideal for diagnostics and class tests early in a module where one needs to ensure that all students can do the basics. Some preparation is needed; the students should have used similar tests before, the system tested by a fictitious student login to ensure that write permissions in answer directories have been correctly set, and the lecturer should ensure that he / she can actually read the answer files and that enough information is recorded to interpret them. It is important to award some marks for the class test (say 10%) and to invigilate the test so that weak students take the event seriously and try to get good marks, which are often viewed by them as a hedge against anticipated low marks in the exam.

It is obviously important to follow up failing students; Mathletics does this by specifying an individual study programme based around Calmat, Mathwise or Transmath and a subsequent retake of the test. (Note: a nice feature of Perception is that at the end of a test, students can be routed automatically to a web site containing remedial material). The lecturer or tutor must take an active interest in this process and intervene as needed. As mentioned above, common errors should also be dealt with in lectures.

A more informal approach, whereby the students simply use objective tests as a resource, can also be taken. This suits stronger students better, and the process can be made more fun by group working and offering prizes (I offer a bottle of wine for >90% in any marathon test). Such activity can be used as extension material, but a formalised threshold of activity for all students should still be carefully monitored so that weak students make a serious attempt and therefore learn and benefit from the feedback.

Test Structure of Mathletics

Mathletics is entered via a shell which is itself a Question Mark test called by present.exe, or testview.exe if no answer files are required. The function of the shell is to:

- provide "driving instructions" for the tests, information on the test structure and a description of linked software (Calmat, Transmath and Mathwise)

- ask the user his / her name. This is not as trivial as it might seem. *What is your name?* will give first names first or last, initials, nicknames, etc and this makes sorting answer files very difficult. Mathletics asks for userids which have the advantages of being unique, allowing the lecturer to sort users into department or year group, and providing their email address for follow-up communication

- provide navigation to the various tests via multimedia buttons which call another instance of either present.exe or testview.exe (note: this must be consistent; calling testview.exe from within a shell called by present.exe, or vice versa, causes a conflict). The multimedia button specifies the test

and where to write answer files in the subdirectories Tests can cover on a single skill (a "track and field event"), a combination of related skills (eg algebra pentathlons, decathlons, marathons) or wide-ranging skills at a certain level (post-GCSE or post A-level diagnostic tests)

- give the credits; much of Mathletics has been written by final year project students who were considering a teaching career. Writing the tests gives them a significant insight into teaching and learning mathematics and, occasionally, exposes their own misconceptions in basic topics!

For portability to networks, the directory structure for Mathletics has to be flat so that calls to tests, executable files and graphics will be found in the current directory without specifying a drive and path. Calls to other resources, such as notepad.exe or calc.exe also work without a specified path but for other programs such as Word, Excel, Derive, Mathematica etc one would need to edit their calls when moving from eg stand-alone to network. For this reason, these resources are not currently used and neither are sound or video, which would have a substantial time delay over a network.

Tests and Libraries

Each test contains a title screen and longer tests give a contents page, an option to quit, and sometimes a few pages of revision material (this is not intended to be a primary teaching resource but merely to jog the students' memory on material they have already been taught). The test then composes, at runtime, a random selection from question libraries and can jump according to student's performance, usually to the end of a test if a student does poorly on the easy material, so that they are not asked questions which are beyond them. The test also specifies the mode of delivery of questions (strict or free answering sequence), whether feedback and marks are given, and what is in the caption bar at the top of the screen (this should state test and user name to give the students confidence that their marks are actually being recorded properly).

Note that for free answer sequence, one should not give feedback, marks or the running total in the caption until the end of the test since students could then change answers by guessing until they see their mark increase. Mathletics attempts to lock the student into a test, so that they will not quit immediately they get one wrong and try again. This would create many partially-complete answer files which are difficult to administer. Finally at the end of a test, students are given a percentage (not points) mark, feedback and advice on what to do next according to their marks band, and marks for each library tested. They can print this page and cross-reference to remedial material according to a displayed and printable table.

Tests do not contain actual questions; these are held in libraries and thus may be recycled to compose new tests once a substantial number of base libraries has been created. Each library contains 20-30 questions to allow for random selection and hence repeat tests. It is therefore important to ensure that each question within a library covers the same skill and is at a comparable level of difficulty. It is also helpful to have a consistent look and feel, using the same colours and avoiding non-standard fonts which may not be installed on users' machines. Lettering should be at least 20 point and contrast strongly with the background colour – remember that some users will have vision impairment and may be colour-blind. The *Continue*, *Hint*, *Formula sheet* and *Calculator* buttons should be consistent in colour and position (bottom right) with feedback appearing in a vacant place on the screen so that the original question and answers can still be read. Mathletics places this centre right, often over a cartoon or *Tip of the day* (which gives additional information on the question such as *Did you know that* e^x *and* $\exp(x)$ *are the same thing?*). Note that in formal testing, one can have question-by-question control over when a student may use a calculator eg one might expect a student to be able to estimate $\sqrt{180}$ without a calculator.

To paste equations into QM Designer, one can run the MathType equation editor or its cut-down version (eqnedit.exe or eqnedit32.exe) which comes with Word. It is convenient to create an icon or shortcut to run this without entering Word. One should then double all the default font sizes (using SIZE, DEFINE) before typing, selecting and copying the equation to the clipboard and pasting into a rich text box using PASTE, GRAPHIC. Equations cannot be resized or further edited, so it is advisable to use black text and line breaks rather than inserting them into text where alignment is not possible. Such a procedure may become redundant with a combination of Perception and MathML, which would also allow insertion of random parameter values into questions, distracters and feedback. Whilst this would be a real enhancement, students will still need to see a range of questions which are actually different.

Graphics can easily be inserted into questions, but it is worth noting that this will dramatically increase file

size. Thus black and white should be converted to two-colour bitmaps; if not, they may also display in unexpected ways on other machines (eg black on black) and be unreadable. Graphics of worked solutions or formula sheets created using the equation editor and then screen captured, may be called from buttons or shown as feedback to a question.

QM Designer and Perception support the following question formats:

Multi-choice questions

This should include the correct answer (key), at least three plausible distracters and *I don't know* and *None of these* options. There are, of course, problems with such questions constraining, or at least structuring the student's thoughts and in mathematics they can suggest various forms of "cheating", ie substituting spot values eg $x = 0$ and/or $x = 1$, elimination of other options by considering asymptotic behaviour as $x \to 0$ or $x \to \infty$, or doing the question backwards, ie differentiating the options instead of integrating the question. However, the overwhelming advantage for low-level testing is that distracters can be devised using commonly occurring "mal-rules" so that the feedback can tell the student not just that he/she is wrong, but *why*. An example might be to evaluate $10^2 \times 10^3$ offering as options 10^5 and 100,000 (both correct but the second might give the feedback *Are you using a calculator? If so, please don't!*), 20^5, 20^6, 100^5 etc all requiring appropriate feedback.

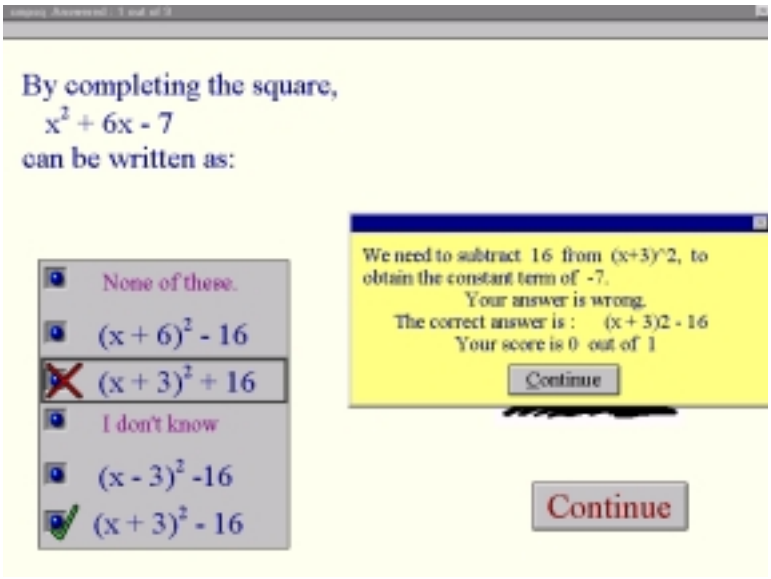

Fig 2: A multi-choice question with feedback

The subsequent whole-class view for a lecturer is also very helpful. It is, moreover, possible to give distracters which detect students doing questions backwards or incompletely eg a partial fraction which, when multiplied up, gives the correct denominator but incorrect numerator would generate feedback to the student that he/she has not completed the problem but is taking a shortcut. Similarly, distracters that do not arise from any plausible mal-rule would give the message *You must be guessing – please don't!* and the student will be deterred from guessing by knowing that this is being recorded. Note that *None of these* should always appear and should sometimes be correct, giving feedback confirming the correct answer; if not, then feedback for this and for *I don't know* should either give the correct answer or a hint and allow a second go. Finally if the student selects the correct answer, feedback can sometimes still be helpful eg *Correct; I hope you spotted it was the difference of two squares.* Scoring should be simple; 1 for correct, 0 otherwise; negative marking will discourage students from committing themselves and they will then miss the feedback.

Multi-response questions

You will need to tell students that they should click on all options they think are correct. An example might be a question on strategy, eg

To find $\frac{d(e^x \sin(x))}{dx}$ I would need to use:

- *Product rule*
- *Chain rule*
- *Quotient rule*
- *Logarithmic differentiation*

Another example could be which property (symmetric, continuous, bounded, etc) applies to $\sin(x^2)$.

Feedback with Designer is blanket rather than specific, ie the same whatever the student response. Perception can be specific but it is lengthy and complicated to cover all possible answer combinations, so this may not be very useful. Scoring is of more concern; to avoid rewarding a streetwise student who selects everything, the following scheme was adopted in Mathletics; 1 mark for choosing each correct answer and also for avoiding each wrong answer, -1 mark for choosing each wrong answer and avoiding each correct answer. This marking scheme prohibits the *I don't know* option and students should simply abstain from answering in this case. Furthermore at least one correct answer must be displayed rather than "None of these" being correct. Thus a random guesser will, on average, get zero marks. Note that the above question is potentially worth 4 marks now; on the

other hand it is harder that a straight multi-choice question. However, the above scheme is not a hard and fast rule; it will not make sense if the distracters are mutually exclusive for example, since knowing the correct answer is then the same as knowing other answers are incorrect.

Numeric

The student is asked to give the correct answer, to within a tolerance if appropriate. Since the student is not led by the options, a correct answer demonstrates mastery. However because it is difficult to specify all plausible input, incorrect answers usually can only generate blanket feedback which is limited in usefulness. Scoring should be 1 mark for correct; zero otherwise.

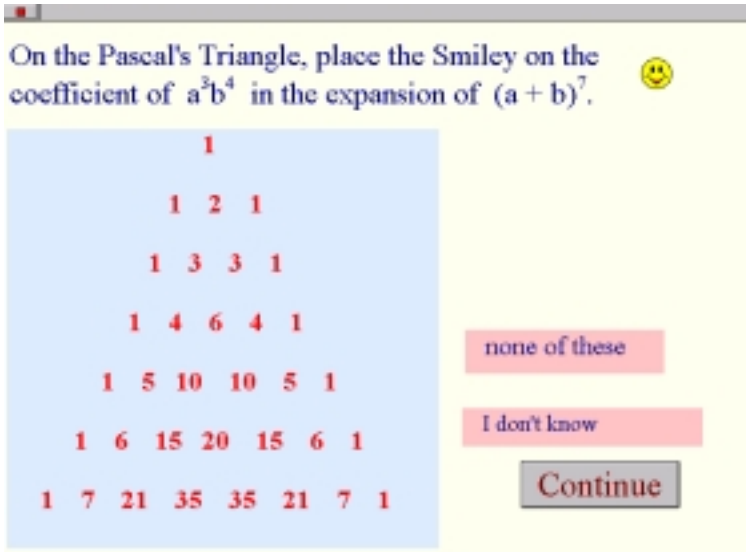

Fig 3: A hotspot question

Text match

Mathematics input problems usually mean that this is not a useful question format. For example a text match on ½ would mark, 0.5, .5, 0.50 etc as wrong.

Fill in the blanks

This format is helpful when one wants to input a multi-part answer such as a pair of coordinates. As for numeric questions feedback can be hard to define. Also one should turn off "find text in input"; otherwise, for example, the answer 1234567890 would be marked correct for any single digit true answer. Scoring should be 1 mark for correct; zero otherwise.

Hotspot

This is a very powerful type of question in mathematics. One can, for example, display a page of mathematics and ask the user to point to the error, or choose *I don't know* or *There is no error*. It is important to score, label with descriptors and provide feedback for all feasible parts of the screen (eg covering each line of the mathematics). For example,

Where is 2-(-3) on the number line?

would have –5, -1, 1 and 5 all hotspotted; similarly when asking the student to point to $x = -3\pi/2$ on a sketch of the sine curve without scales shown, one would hotspot other multiples of $\pi/2$. Scoring should be 1 mark for correct; zero otherwise.

Selection

This is a good way to ask many questions on the same topic by matching the items to questions or statements. For example, one could ask whether statements about the secant function were true, false, undefined or meaningless. Another question might ask for a sequence of techniques needed to solve a problem, ie

1) put into standard form, 2) find the integrating factor ...

Because of the multiple nature of selection questions, accurate feedback is difficult (beyond giving the correct answer). Scoring should be 1 mark for each correct selection; zero otherwise.

Future development

Mathletics can be downloaded from the WWW at www.qmark.com / questionbank / mathletics.html or on CD (version 2.4.3 - again free) on request. I hope that schoolteachers and lecturers will use this resource and provide me with feedback to improve and develop Mathletics. Setting such objective tests certainly makes a valid, challenging and useful project for either mathematics or education staff and students, and I would welcome contributions to include within Mathletics. As well as increasing the number of questions in existing libraries, I plan to expand Mathletics other areas; in particular mechanics, discrete mathematics, linear algebra and descriptive statistics would be suitable and relevant to a large number of students in science and technology departments. It is also planned to convert Mathletics to the WWW and exploit the additional capabilities of QM Perception. A start on this rather large task has been made by Daugherty and McCabe at Portsmouth University, see http: / / L62.csm.port.ac.uk

References

[1] Kyle J, "Mathletics – a review", *Maths&Stats* v10 n4 Nov 1999

[2] McCabe M & Greenhow M, "Designer software for mathematics assessment?" *Maths&Stats* v6 n1, Feb 1995

...continued overleaf

Courseware News

Question Mark THINKS on the web: the University of Northumbria at Newcastle and Question Mark Computing have got together to publish a suite of mathematical software written in Question Mark for Windows to enable students from semi or non-numerate disciplines to improve their mathematical abilities. The Tutorial Help in Numeracy Key Skills (THINKS) suite of software is now available to download from Question Mark's web site. It consists of fifteen tutorials, each containing around a dozen questions and revision material. The tutorials include graphics and sound, and comprehensive feedback. At the end of each tutorial, students receive a percentage score. Advice is also given on other software titles that might help them to improve their understanding of the topic. The web site is at www.qmark.com/questionbank/thinks.html

MAPLE 6 will ship from the end of February. It will be released simultaneously on all platforms, plus Student and Trial (time limited) versions. The pricing structure is not currently available, but it is aimed at advanced students, researchers and industry, while not forgetting undergraduates.

What does MAPLE 6 offer the teacher?

- Worksheets written in earlier versions of MAPLE will run under MAPLE 6.
- MAPLE 6 worksheets can be output in TeX as before, or in RTF which will go straight into any standard word processor. It includes page numbers, page breaks and page layouts.
- There is a URL-linking feature which allows a MAPLE worksheet to be set up as a browser.
- There is a powerful link to Excel 2000, allowing symbolic, numeric and graphical analysis inside an Excel worksheet. This will be potentially valuable for the teaching of financial mathematics.
- Visualisation - a new structured data browser allows an overview of large matrices and vectors with the ability to zoom in and out on individual entries. Input and output of data is possible in many more formats.
- It incorporates over 100 matrix computation functions from the NAG library. The use of these numerical routines can be switched on and off.
- Task based help system - right click on any MAPLE object to see the possibilities.
- MAPLE Application Centre - Web pages containing many examples of good practice

There is more about new Maple 6 in the latest issue of The Maple Reporter, which can be read online at www.adeptscience.co.uk/products/mathsim/maple5/reporter/issue4/

DIAGNOSYS is an expert-system testing shell for maths and other subjects, widely used in UK universities and abroad. QEDIT is a test design tool for use with it. MATHINPUT is a general-purpose maths input tool used by DIAGNOSYS and by other CAL developers. DIAGNOSYS v3.23, MATHINPUT v1.31 and QEDIT v1.02 are available from the DIAGNOSYS web site. All files and the Mathinput DLL are currently in Windows 3.1 format. A Win95/98/NT version of the Mathinput DLL will be mounted very shortly for developers who wish to use it with their own programs. The Web site is at www.staff.ncl.ac.uk/john.appleby/diagpage/diagindx.htm

Word2TeX version 1.3 from Chikrii SoftLab is described as the complete solution for Microsoft Word to LaTeX document conversion. Word2TeX 1.3 is an intelligent 32-bit external converter for Microsoft Word 6.0/95, Word97, Word2000 (or any other 32-bit Microsoft Word); also compatible with Japanese Word and mostly all native 32-bit Microsoft Word's. Produces LaTeX, LaTeX2e, AMS-LaTeX, AMS-LaTeX2e documents. Translates equations, document formatting, tables, figures, font families, does equation auto-numbering, auto-referencing, and a lot of new and improved useful features. Available at the official product support site www.vcom.kiev.ua/~chik; further info from chik@mw.kiev.ua

...continued from previous page

[3] Greenhow M & McCabe M, "A question mark over Maths Assessment?" *Maths&Stats* v7 n2 May 1996
[4] Correll M & Shaw P, "Question Mark Designer - not only for assessment" *Maths&Stats* v8 n2 May 1997
[5] Lawson D, "Some experiences using Question Mark Designer" *Maths&Stats* v9 n4 Nov 1998
[6] Strickland P, "Short review - QM Perception" *Maths&Stats* v10 n3 Aug 1999
[7] Greenhow M, "Computer-based diagnostic tests and assessment at Brunel University" *Maths&Stats* v7 n3 Aug 1996
[8] Heard S, Nicol J & Heath S, "Setting objective tests" MERTaL Publications, University of Aberdeen, 1997, ISBN 1 873154 43 7