RECENT DEVELOPMENTS IN SETTING AND USING OBJECTIVE TESTS IN MATHEMATICS USING QM PERCEPTION

E Ellis, N Baruah, M Gill & M Greenhow

Recent Developments in Setting and using Objective Tests in Mathematics using QM Perception

Nabamallika Baruah, Edward Ellis, Mundeep Gill & Martin Greenhow

Mathematical Sciences, School of Information Systems, Mathematics and Computing, Brunel University, Uxbridge, UB8 3PH

Abstract

This short paper builds upon work described at the last CAA Conference, Greenhow & Gill (2004), in setting objective tests in various areas of mathematics using Question Mark Perception. Current activities continue to exploit the QML language and template files, coupled with MathML mathematics mark-up and the Scalable Vector Graphics (SVG) syntax for producing diagrams. There are many advantages to using such mark-up languages, primarily the use of random parameters at runtime that thereby produce dynamic equations, distracters, feedback and diagrams. An unlooked for, but welcome, advantage, is that one can also resize and recolour these elements by reading the preferences that have been set up in a user-defined cookie. This means that "reasonable provision" for disabled students as required by the SENDA legislation, is built-in.

The MathML and SVG technology can be exported to any web-based system, or indeed ordinary web pages that can provide an inexhaustible set of realisations at the click of the reload button. Being central to the display of mathematics on the web, MathML's WebEQ applet has recently been considerably extended to include graphing of MathML expressions, naturalistic input of equations with syntax checking and math-action <maction> tags. These math-action tags can be used to define a specific part of an equation, and mouse actions can then be acted upon, for example to provide a commentary on that part of the equation, toggle to another equation (perhaps a derivation of the tagged term or similar) or, possibly, to set a variable that can be used for marking (as in a hot spot question). The first part of this paper will show how these new facilities can be input into new question types for effective questions and feedback design.

It is clear that much useful technology already exists, but setting effective questions that benefit students' learning requires equal attention to their content and pedagogy. The second part of this paper looks at a possible

methodology for setting much more advanced questions than hitherto, looking closely at an example from the ordinary differential equations section of Mathletics.

The third part of this paper looks at a series of experiments with a first year mechanics group at Brunel University, as part of the Formative Assessment and Feedback (FAST) project. Students' reactions were studied, especially the effect of the feedback on their subsequent behaviour when faced with similar/dissimilar questions after a variable time delay. Students spent a lot of time and energy considering the feedback provided, sometimes copying it down or printing it out. Somewhat surprisingly, it seems that a "learning resource" has actually been written, whose formative nature is of equal or more importance than the assessment function originally intended. It can be concluded that plentiful formative feedback is of great importance in the students' ability to learn mathematics from the tests, rather than simply get their grades or marks in an efficient manner.

Part one: recent technical developments

Design Science (2005) have for some time been developing a product they call WebEq. This consists of a family of java applets, each of which all focused on interpreting MathML. MathML, SVG, JavaScript and Java applets may be exported to any system based on web pages. This may be as part of a larger site/system or as stand-alone pages. Indeed, one does not even require a web server.

WebEq version 3.6 features many new developments. The family of java applets has been extended. They now include:

WebEq applets		
Input	*Output*	*Processing*
Input applet	Viewer applet	Comparator applet
	Graphing applet	Evaluator applet.

Previously only the viewer applet was used in Mathletics. The inclusion of the newer applets has allowed a number of new question types to be supported and for more detailed feedback to be given.

The WebEq graphing applet will plot up to ten different equations at a time. Each graph may be of a different colour. All graphs are limited to one-dimensional rational functions. This allows more constructive feedback in areas like calculus and linear algebra, as functions and their tangents may be easily plotted.

The WebEq viewing applet has been enhanced to support more of the MathML version 2.0 specifications. Most notably this includes the math action tag. Whilst it does not support the menu or tooltips from the MathML version

2.0 specifications, Design Science have included some of there own actions. These include:

- display text in the browser status bar on a mouse over.
- highlighting parts of the equation on a mouse over.
- toggling parts of an equation on a mouse click.
- embedding a link in an equation.

Status line text can be used to comment on significant features in an equation. Toggling equations can be used to provide dual definitions (for instance, defining a complex number in both Cartesian and Polar form). The embedded link can be used to provide more extensive information, by navigating to a page with more details on a particular stage of a proof or derivation.

Each of the above has implications for people developing online tutoring materials, not only to those in the CAA community.

New question types

The new applets have also been combined to produce new question styles. Multiple choice and multiple response question type templates have been reworked to support use with the WebEq graphing applet. For example, a multiple response question was created where the student is requested to identify which of a number of functions is odd or even, given only the graphs of the functions.
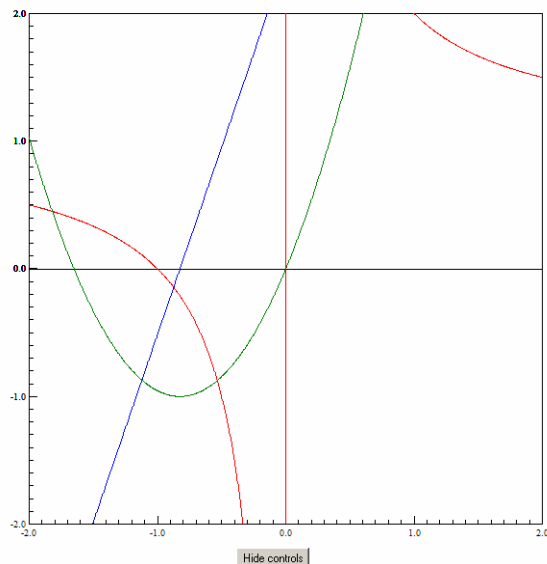
Using the WebEq comparator applet makes a loose mathematical comparison of two MathML strings fairly trivial. Similarly the WebEq evaluation applet will calculate the value of a one-dimensional MathML function at a point and is also fairly easy to use.

These become really useful when combined with the WebEq input applet. This allows student to input MathML expression using a fairly natural GUI. That expression can then be obtained from JavaScript in MathML. When this MathML is used with the comparator or evaluation applets, one is able to produce question styles in which a student may be asked to input a polynomial or other one-dimensional mathematical function and have it correctly marked by the system. Figure 1 shows a question has been developed where a student is asked to differentiate a polynomial. To respond they must enter their 'differentiated' polynomial. This is then compared to a trusted answer calculated by the question and rendered in MathML. The students MathML is then sent to a copy of the WebEq viewer for the student to verify. The answer MathML and the students MathML expressions are then sent to the comparator applet for checking. If they are similar enough, the student gains the marks otherwise they do not. The input typed in Figure 1 makes no mathematical sense (!) but was typed in to show how the system can alert a student to a common input error.

Differentiate the polynomial below with respect to x. Enter your expression in the applet provided. Before clicking submit, press the confirm button to double check your answer.

The polynomial to be differentiated:	$2.42306760358614x + 1.4678141528871697x^2$
Enter your equation here:	$1/x + 1$
Click the confirm button to check your response:	Confirm
Check your equation before submitting:	$\frac{1}{x} + 1$
Finally submit your response.	

Submit

These functions are plotted on the graph below. Your response is plotted in red and the original function is shown in green. The correct response is drawn in blue.

Hide controls

Related material

0 out of 1

Figure 1. A mathematical input question and a graphical part of the feedback

This type of question moves away from tutor provided answers (eg. Multiple Choice, Multiple Response) toward student provided answers (numerical input, mathematical input). This allows one to emulate, in some limited fashion, far more powerful systems (e.g. STACK) but without the need for an underlying computer algebra system.

We also look forward to question types using hotspots embedded in equations. This would be implemented using the embedded link to refresh the document with get-style options appended to the page URL. In principle, this combination could be used wherever numerical input options are used (e.g. Fill in the blank, Multiple Mathematical Input, Responsive Mathematical Input).

All above WebEq applets have JavaScript wrapper functions (functions that provide the HTML to include the MathML applet along with MathML and options). These are fully compliant with existing accessibility code.

Part two: a possible methodology for setting questions on advanced material

Although it is established in part three that objective questions can be set effectively for testing factual recall and simple, but non-trivial, solution procedures, (multi-skills) questions much beyond A-level mathematics seem to require a much more structured approach. Assumed and tested skills need to be carefully specified if one is to interpret answer files beyond the simple marks stored. Furthermore, with random parameters within questions, the question and outcome descriptions need metadata that characterises (both mathematically and pedagogically) the class of question and tested skill(s) and the distracter mal-rules.

For undergraduate level 1 ordinary differential equations (ODEs), the mal-rules were chosen according to the errors made by the students on 88 Brunel University exam scripts from 2003-04. (The information in chief examiners' reports from examination boards were useful in identification of problem areas, but not sufficiently detailed to identify actual mal-rules that can be encoded in questions.) Examples of mal-rules include:

communication errors such as: mis-reading of initial conditions/instructions, incorrect input of initial condition(s), incorrect rounding, lack of attention to the restrictions on the problem, misinterpretation of the initial condition(s), inability to map a real-world problem to mathematics (i.e. apply problem solving skills)

algebraic errors such as: division by zero, bad/lost/assumed parenthesis e.g. $(4x)^2=4x^2$, improper distribution e.g. $5(2x^2-10)=10x^2-10$, commuting operations e.g. $\sqrt{a+b}=\sqrt{a}+\sqrt{b}$, $\dfrac{1}{a+b}=\dfrac{1}{a}+\dfrac{1}{b}$

function manipulation errors such as: $\cos(x+y)=\cos(x)+\cos(y)$, $\ln(y)=x^3-x+c$ resulting in $y = \exp(x^3)-\exp(x)+\exp(c)$; $e^{ab}=(e^a)(e^b)$

recurring calculus errors such as: improper use of the integration formulae, improper substitutions

procedural errors such as failure to: distinguish the type of the differential equation, use correct method for the type of ODE, find the integrating factor, find the particular integral, obtain the complementary function (this sometimes involves in factorisation), express answer in required form.

The above identification process becomes increasingly difficult when one sets questions in levels 2 and 3 at undergraduate level where one must assume the student possesses more underlying skills. At present, we do not know where the boundaries for effective objective questions lie; testing the ability of students to make constructions, such as mathematical proofs and models, is

currently not feasible via objective questioning alone. Nevertheless, we feel that much can still be done in testing some of the more algorithmic parts of advanced mathematical methods courses, especially if a clear skills/sub-skills tree structure for the whole question bank could be constructed. This might provide a deeper insight into a student's understanding by looking at results across a wide range of tests and automatically identifying errors that manifest themselves in different areas; for example, failure to understand negative fractions will cause problems in algebra and calculus as well as arithmetic. Whilst this lies in the future, we have recently constructed libraries of questions that cover Laplace transforms and ODEs with this development in mind. We illustrate these rather abstract ideas below with a concrete example from one of the ODEs libraries for which assumed pre-requisite skills involve a good grounding in algebra, differentiation and notation.

For many of the mathematical topics (e.g. exact equations or integrating factor type) we have subtopics that organise the questions according to the different levels of understanding/ cognitive skills, identified using Bloom's taxonomy as revised by McCormick and Pressley (1997). We have considered three levels of learning:

knowledge: the ability to know specific facts, common terms, basic concepts, techniques, principles and theories. ODE questions in this class might test students' abilities to distinguish between linear and nonlinear differential equations, find the integrating factor, write down the general solution and determine a particular solution. A question stem might be: Which of the following options will represent the particular solution of $2\dfrac{dy}{dx} + y = e^x$, $y(0)=1$?

comprehension: the ability to understand, differentiate between concepts and terms, explain, rewrite in new way and interpret. ODE questions in this class might test students' abilities to identify linear differential equations, transform into linear differential equation by substitution and find the initial condition(s) from the particular solution. A question stem might be: The function *y=f(x)* satisfies $2\dfrac{dy}{dx} + y = e^x$ and $3ye^{x/2} = 2 + e^{3x/2}$. What is the initial condition for *y*?

application: the ability to apply facts and concepts to new situations, to solve problems and apply techniques to the real world problems. ODE questions in this class might test students' abilities to use their techniques of linear differential equation to solve a physical problem and to identify and use initial conditions from the physical situation. A question stem might be: Identify the transient term in the solution of $2\dfrac{dy}{dx} + y = e^x$.

So far, Mathletics has focussed on testing of skills. For more advanced material, it may be possible to infer students' understanding by mapping skills-based question content to underlying and more general concepts. For ODEs, the concepts involved might include: order, linearity, integrating factor, general solution, initial condition, particular solution. We intend to add these items as

required to the question or topic metadata. A useful tool for the display of such concept mapping is a Novak's (1991) Vee diagram. For the above problem see figure 2, which can be a useful structure when designing parallel questions in the three sub-topic areas of knowledge, comprehension and application.
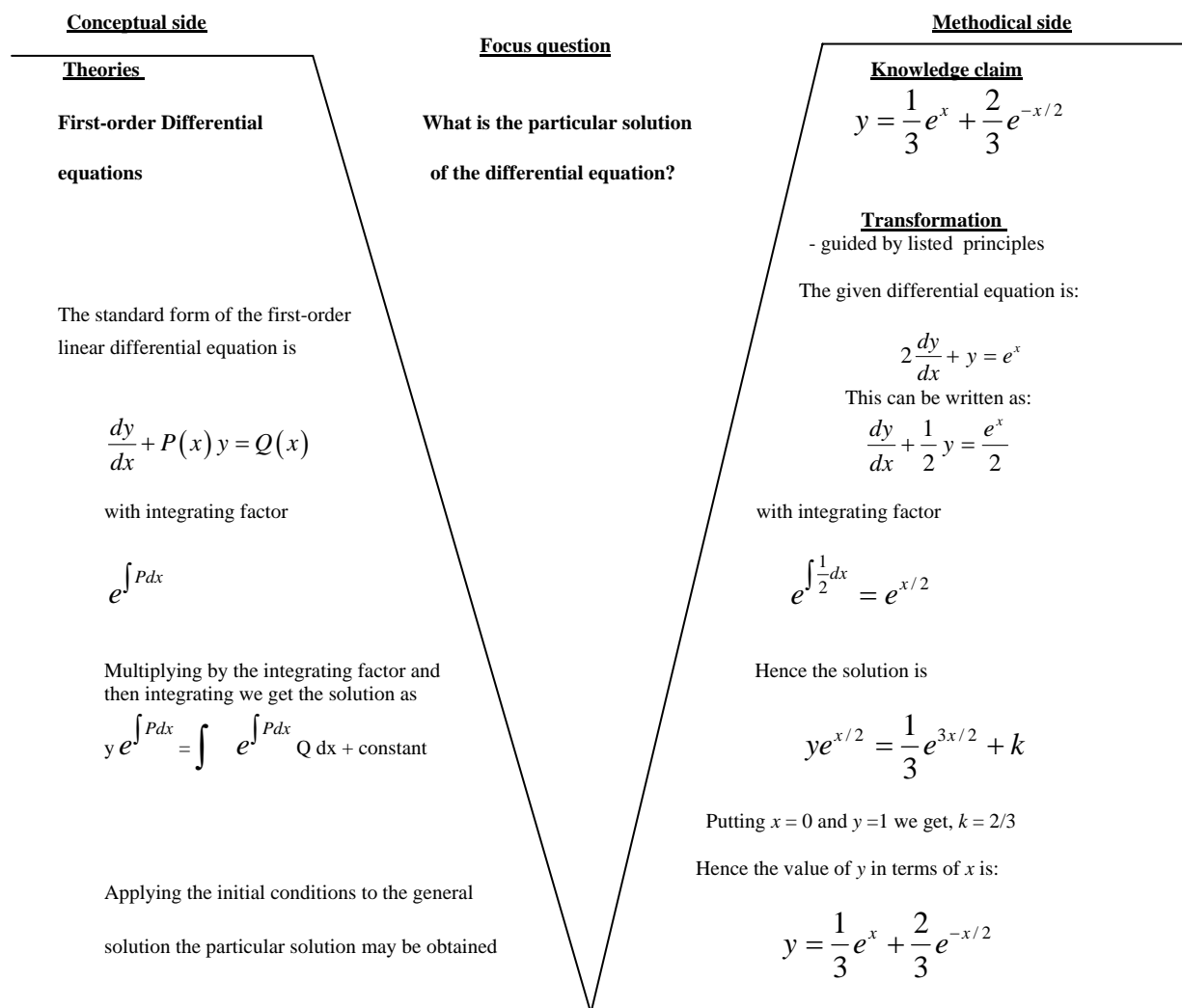
Conceptual side

Methodical side

Focus question

Theories

Knowledge claim

$$y = \frac{1}{3}e^x + \frac{2}{3}e^{-x/2}$$

First-order Differential

What is the particular solution

equations

of the differential equation?

Transformation
- guided by listed principles

The given differential equation is:

$$2\frac{dy}{dx} + y = e^x$$

This can be written as:

$$\frac{dy}{dx} + \frac{1}{2}y = \frac{e^x}{2}$$

The standard form of the first-order linear differential equation is

$$\frac{dy}{dx} + P(x)y = Q(x)$$

with integrating factor

$$e^{\int P\,dx}$$

with integrating factor

$$e^{\int \frac{1}{2}dx} = e^{x/2}$$

Multiplying by the integrating factor and then integrating we get the solution as

$$y\,e^{\int P\,dx} = \int e^{\int P\,dx}\,Q\,dx + \text{constant}$$

Hence the solution is

$$y e^{x/2} = \frac{1}{3}e^{3x/2} + k$$

Putting $x = 0$ and $y = 1$ we get, $k = 2/3$

Hence the value of y in terms of x is:

$$y = \frac{1}{3}e^x + \frac{2}{3}e^{-x/2}$$

Applying the initial conditions to the general solution the particular solution may be obtained

Figure 2. Vee diagram for Event / Object "Find the value of y in terms of x from

$$2\frac{dy}{dx} + y = e^x$$ **and y(0)=1"**

Part three: testing objective questions in mechanics

For the academic year 2004/2005, part of the question database covering the Edexcel M1 A-level module, see Gill and Greenhow (2005), was tested on students in a mathematics undergraduate level 1 mechanics module. This was a pilot study and was carried out as part of the Formative Assessment in Science Teaching (FAST) project (2005). The aim of the FAST project was to find out how effective the feedback was:

- in terms of helping students do similar/dissimilar problems after a variable time delay,
- in getting students to spend time on the task at hand and properly engage with the material to effect a positive change in their learning behaviour.

For this 12 week module, students had 3 contact hours per week, 2 hours of lectures and 1 seminar. Additional to these scheduled contact hours, one-hour lab sessions were set up for students to attend every other week. During these sessions students were required to complete computer-aided assessments. On average, each assessment comprised 5 or 6 questions, based on topics that students had recently covered in lectures.

P   T

8g N   ↑T

Q

2g N

↓ a ms²

SVG diagram with correct question parameters

Using F = ma on particle Q:

2g - T = 2a ...(1)

MathML equations with correct question parameters

Using F = ma on particle P:

T = 8a ...(2)

Substituting equation (2) into equation (1):

Use of units in final answer

2g - 8a = 2a

2g = 10a

so, a = 1.96

Therefore the acceleration of particle P is 1.96ms^{-2}.

Figure 3. Annotated screen shots of a typical multi-choice question and feedback

Results collected

For all assessments that students attempted, answer files were recorded showing which questions students had answered correctly/incorrectly. For Multiple-Choice and Responsive Numerical Input type questions, the distracter chosen/number input indicates the class of mistake(s) students were making and the mal-rules they were probably applying. (A mal-rule is a logical, but incorrect way of doing the problem, see figure 3's distracters and Nichols, Gill and Greenhow (2003) for examples.) The effect of different question types can also be identified from the answer files. For example, students answered more Multiple-Choice questions correctly compared with Numerical Input, despite the difficulty level of the mathematics being the same.

The answer files, evidence from class observations, students' notes and doodles at the time and subsequent exam scripts, represent an extensive and sometime bewildering body of data. To simplify matters, and to find out if students were engaging with/attending to the feedback, the following indicators were used: vectors, units, diagrams and presentation of solution. In all questions vectors were indicated in bold, units were used throughout, diagrams were used where necessary and all solutions were presented in a step-by-step way where any assumptions made or formulas being used were stated. The question is; does any of this rub off on the students and *feed forward* affecting their subsequent behaviour?

At the end of several randomly chosen sessions, notes that students had made while attempting the questions were confiscated, to find out what students were actually writing down, what methods were being used, whether any of these four indicators were being used correctly, and whether any new mal-rules could be identified. Students were assessed with a 2-hour unseen written exam at the end of the course. Students' exam scripts were also analysed, in particular to find out how many students were correctly using the four indicators.

In the very last lab session, all students were given an Assessment Experience Questionnaire (AEQ) to complete, see FAST (2005). The AEQ enabled students to give their feedback on the overall lab sessions, the assessments and questions they completed and on the feedback they received.

Summary of results

From the recorded answer files and class observation, it was immediately obvious that some assessments contained too many (5 or 6) questions for the given time (50 minutes). Some students answered only 2 or 3 questions, spending a most of the time reading the feedback. Indeed, many students were inputting random numbers in the numerical input boxes or selecting the "*I don't know*", option in the Multiple-Choice type questions, just to read the feedback. This implies that many students are using the questions as a primary learning tool and are engaging with the feedback. When repeating an assessment immediately after finishing the first attempt, students were generally able to retain the feedback and make use of it to answer questions correctly at the second attempt. Moreover, it was found that some students were able to retain some of the feedback much longer and were able to answer questions correctly after a period of about 4 weeks.

Based on the marked increased percentage of students indicating vectors correctly and making use of diagrams in their solutions in exams this academic year compared with previous years, we can conclude that the feedback had a positive effect on students for these two indicators. However, no definite conclusions could be derived about the use of units and the presentation of the solution, but the lack of units in students' final answers was identified as an outstanding and persistent issue.

The results obtained from the AEQ were very encouraging. Students had a positive attitude towards the computer-based questions, even though they found them challenging. Students themselves felt they were engaging with the feedback by taking the time to read it and trying to understand what is was saying. Students also felt that they learnt something new from the feedback.

Some recommendations

Overall, the lab sessions ran smoothly with very minor technical problems. The lab sessions that were held for this particular pilot study were not made

compulsory since the assessment criteria had been finalised before the lab sessions were set up. For future testing, the lab sessions should be made compulsory so that all students attend and complete all assessments set for credit towards the module mark. This became a problem when analysing the answer files; since not all students attended all the sessions, it was difficult to make comparisons of how students were progressing as they interacted with the software.

One of us (MG) initially worried that so much feedback was being made available to students that they would simply ignore it. The results of this study clearly show that extensive feedback is welcomed by, and has a positive effect on, most students. Some students requested even more feedback. In effect, the questions are being used as a learning tool alongside, or even instead of, lectures and seminars. This could have rather far-reaching consequences: question designers should focus much of their attention on feedback, the curriculum needs to make time for students to attend to it and the assessment criteria need to reward such student engagement.

References

1. Design Science (2005) http://www.dessci.com/

2. Formative Assessment in Science Teaching FAST (2005) http://www.open.ac.uk/science/dftl/

3. Gill M. and Greenhow M. (2005) "Learning via online mechanics tests" Science Learning and Teaching Conference, Warwick University.

4. McCormick C. B. and Pressley M. (1997) "Educational psychology: Learning instruction, assessment." Longman, New York.

5. Nichols D., Gill M. & Greenhow M. (2003) "Pedagogic issues in setting online questions" MSOR Connections Volume 3, Number 4: Nov. Available on http://ltsn.mathstore.ac.uk/newsletter/vol4.shtml#no4

6. Novak, J. D. (1991) "Clarify with Concept Maps: A tool for students and teachers alike," *The Science Teacher*, 58 (7), pp. 45-49.