EXPLORING THE POTENTIAL, LIMITATIONS AND USE OF OBJECTIVE QUESTIONS IN ADVANCED CALCULUS

N Baruah and M Greenhow

Exploring the Potential, Limitations and Use of Objective Questions in Advanced Calculus

N Baruah and M Greenhow
Department of Mathematical Sciences, Brunel University
mapgnnb@brunel.ac.uk & mastmmg@brunel.ac.uk

Abstract

This paper describes our experiences with authoring and trialling questions in advanced calculus topics, namely ordinary differential equations, Laplace transforms and Fourier series. These topics are generally taught at the end of the first year or during the second year of a mathematics or engineering undergraduate degree. We expect that many of the lessons learned here will apply to other conceptually-advanced mathematical and scientific content. Typically, what is significant for such content is that many skills are needed from previous exposure to calculus and algebra, and that paper-based questions at this level tend to be more abstract, holistic and open-ended, requiring the sort of flexibility in marking generally associated with human markers. For objective, and therefore more constrained questions, we do not know what is feasible and whether or not questions on advanced topics will actually test the skills they are designed to test. For example, a student may carry out e.g. a Laplace transform correctly, but make an elementary algebraic mistake near the end; this would be easily recognised by a human marker, but simply marked wrong by any current CAA system which cannot assess the (generally handwritten) intermediate steps in a student's solution. Conversely, any question that can be marked by a CAA system is likely to be structured or scaffolded (e.g. by asking for intermediate steps explicitly) so that the original requirement on the student to devise a solution strategy is lost. This paper explores what can be asked effectively: facility with such questions is a necessary (but not sufficient) condition for students to master more advanced topics, so some sort of blended assessment (with human markers) may still be needed for higher-level skills. We describe the process of authoring higher-level objective and report of the experience of running the questions with our second year cohort, including an analysis of the answer files produced. Our evidence suggests that the assessments were useful to students in establishing a solid foundation of skills, mainly by being encouraged, or even forced, to engage with the extensive feedback screens.

Background

During the current academic year, online tests in the advanced calculus section of Mathletics were authored and delivered in an extended form of Question Mark Perception. Mathletics is designed to exploit the potential of computer-aided assessment, especially in formative assessment mode. Our experience over the last 5 years of trials with many hundreds of students indicates that they value the extensive feedback (generally including a fully-worked solution) as a learning resource, as well as for the marks awarded. Moreover, students' learning has been encouraged by the tests building the confidence of first-year undergraduates. The pedagogy of building tests into a module is quite well established, and various trials of the mechanics material have indicated that students move, at least partially, to a deeper approach to study (Gill & Greenhow, 2004). This paper examines whether or not the same claims can be made for assessments covering advanced calculus topics delivered to second year undergraduates.

The underpinning technology of Mathletics, whereby many thousands or millions of question *realisations* are generated by a single question *style* that encodes the algebraic and pedagogic structure of the question, is carried through to the more advanced content described in this paper. We have found that it is extremely helpful in moving students away from simple memorisation towards an understanding of the question's content and solution. The random parameters, possibly constrained according to the question's content (realism of the question and reverse engineering from a desirable solution form), are carried through to all parts of the question so that it realises with:

- dynamic MathML, giving equations in the question and in the (often extensive) solution and other content given as feedback.
- dynamic SVG, giving accurate diagrams, charts and graphs.
- dynamic wording, giving different scenarios, expressed in gender- and ethnically-balanced language.
- dynamic question functionality, such as algorithms that, when run to completion, generate, for example, HTML tables of variable length.

Accessibility (SENDA compliance) has been a key feature of the existing questions. The format of all elements may be chosen by the student and stored as a cookie. A great deal of technical effort has also gone into the writing of functions to underpin the questions. These split into two basic types: functions that return the result of a calculation, e.g. multiplying out two polynomials of arbitrary order, and functions that return display strings e.g. a MathML string to display a table of Laplace transforms or an SVG string to display a graph of a function and a few partial sums of its Fourier series, see figure 3. Exportability of the mathematical content to an ordinary web page or other web-based CAA/CAL systems is another key feature, see Ellis, Greenhow and Hatt (2006).

Before the construction of the questions, the learning levels of the questions were categorised from a pyramidal (rather than hierarchical) version of Bloom's taxonomy (Hatt,J. & Baruah, N. 2006), with the six chambers in three learning levels. The pedagogy of each of the subtopics of Laplace transforms and Fourier series was analysed to specify the tested, and prerequisite, concepts and skills. Concept maps (Turns et al 2000) were drawn for this purpose. The questions are mostly from the first two levels comprising the *remember, understand, apply, analyse* and *evaluate* chambers. *Create*-level questions were designed for only a few topics, see figure 2.

One of the main objectives of the study was to develop questions at higher-learning levels. Several different question types (multiple-choice, multiple-response, hotline, true-false, numerical input and responsive numeric input) have been utilised. For effective and targeted feedback, mal-rules encapsulating the essence of an incorrect solution method or error, are needed for multiple choice, responsive numeric input and hotline questions. To discover such mal-rules, the answer files of previous elementary calculus CAA tests were analysed and answer scripts of past examinations were examined. From such evidence, and from the works of Orton (1983), Schechter (1994) and Greenhow (1996), an error taxonomy has been developed. Not only is this useful in question design, but it also greatly facilitates the interpretation of students' answer files.

For more advanced topics, the choice of question type needs specific attention from both pedagogic and technical standpoints. Some multiple-response questions were designed to test students' understanding of general mathematical properties, but the form of multi-choice questions may be ineffective due to guessing. To overcome this, new four-optioned yes/no and true/false question types have been designed for testing identification of general properties and theorems, see figure 1. Such questions are scored dichotomously to reduce drastically the probability of rewarding guessing. Whilst the question in figure 1 is quite static (in that other realisations will look very similar) other versions are made more dynamic by replacing the unspecified general functions by particular randomised functions.

The hotline and responsive numeric input questions were extended in some of the topics by recording the students' certainty in their answers, along the lines given by Gardner and Gahan (2003), but without negative marking.

As some of the problems solvable by the Laplace transforms naturally require the inclusion of the diagrams, dynamic drawing objects have been developed using Scalable Vector Graphics (SVG). These build elementary drawing objects like lines, rectangles and ellipses, to form new objects such as graphs. Such types of questions may be helpful in the presentation of the question stem, say by the inclusion of circuit diagram, or in the feedback, see figure 3.

An example of a question at the higher *create* level is shown in figure 2 where students need to obtain the limits of integration by correctly interpreting the diagram. Thus students are being tested on concept of periodicity. In the feedback, the general form of the Fourier series is written before being applied to this particular question; this exposes students to the underlying concepts (deep learning) as well as purely procedural skills (surface learning). The feedback is reinforced by providing a graph, plotted using a high-level function due to Ellis (2006).

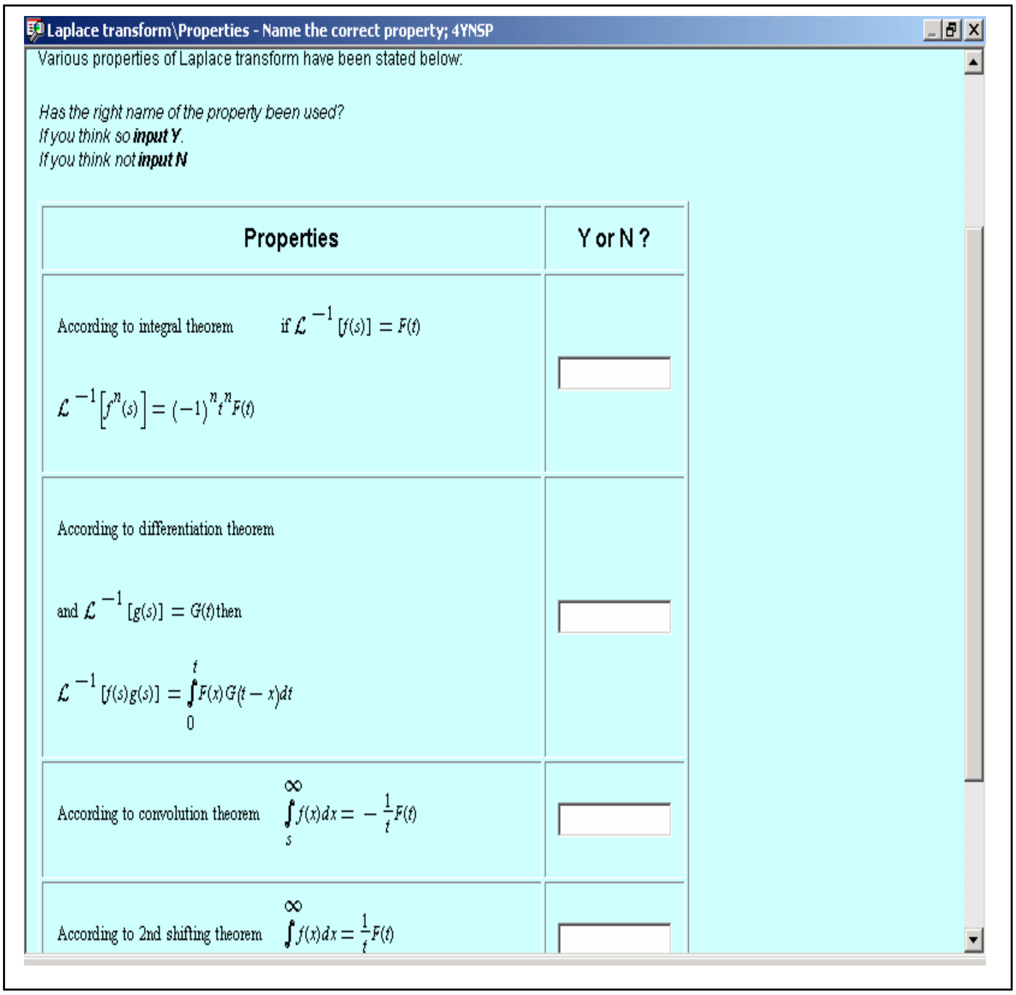

Figure 1 Screen shot of a yes/no question for assessing general theorems and properties.

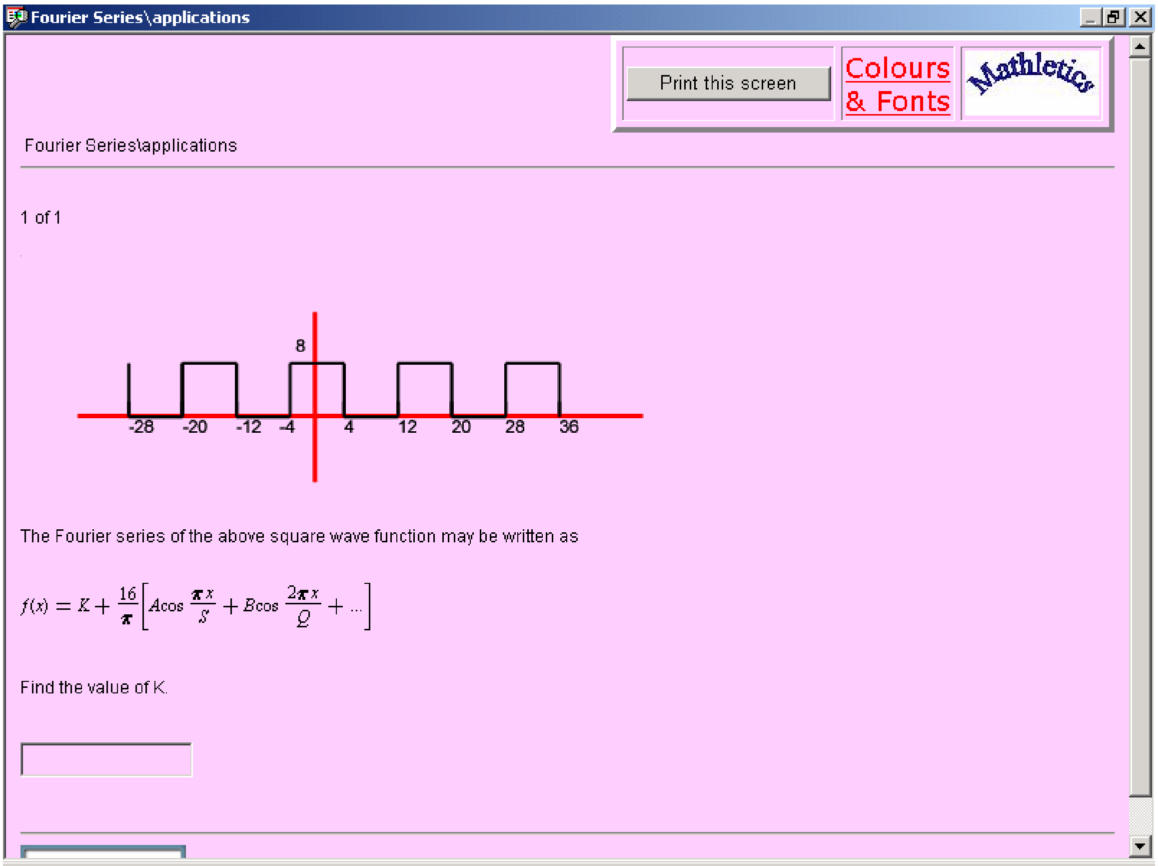

Figure 2 Screenshot of a question requiring analysis of a diagram.

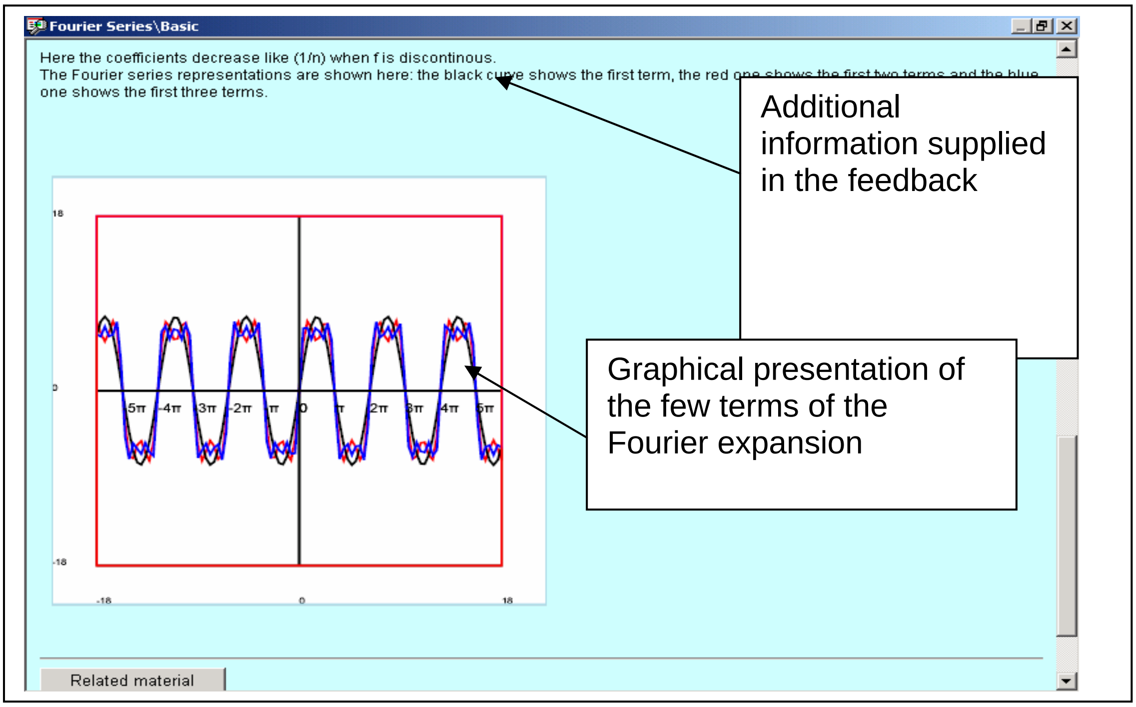

Figure 3 Screenshot of part of the feedback using a graph. In contrast to the diagram in figure 2, which is really a schematic, this graph is accurately drawn.

Trials and student feedback

Questions spanning various learning levels for Laplace transforms and Fourier series were administered to second-year undergraduate students in three different tests in the months of October and November of 2006. The answer files have been analysed in an attempt to understand the questions' impact on the students' learning. Some new mal-rules were also identified through the analysis and were used in the construction of further questions on Fourier transforms.

All the questions have discrimination indices above 0.2, which indicates that no questions were invalid. The average facility values were around 0.5, indicating that the questions were of medium difficulty. However a certainty-based numerical input question and a true/false question had very low facility in comparison to the overall facility of the tests. The average facility value of the multiple-choice questions was more than that of the other type of questions. This probably reflects the effect of the information displayed on screen that allows students to check their answers against the options before clicking 'submit'. Whilst this casts doubt on using this question type for summative or mastery testing of students, for formative testing it is felt that, coupled with the very full feedback available, multi-choice questions are an effective was of building students' confidence.

Not surprisingly, students performed better in the lower-level questions than in the higher-level questions; those who were less certain scored lower than those who were more certain; and students did better in questions that tested a single concept than multi-concept questions. An exception to this appears to be that students were less able to identify general properties than apply them in specific examples. This may be due to such general properties being stated in a more abstract and mathematically terse way, or it may indicate deficiency in the conceptual learning of the topic. At the other end of the taxonomy, most mistakes occurred due to procedural errors, especially in the lower-level questions.

Results from a questionnaire suggest that students found the tests, and especially the feedback, useful. The marking scheme of some of the multiple-response questions has been set so that marks are obtained only if all the correct answers are chosen without choosing any incorrect options. About a quarter of the students considered this was not fair.

Conclusions

Whilst questions at the lower levels of a modified Bloom's taxonomy can be created and shown to be effective in testing basic, albeit necessary, skills, any course in advanced calculus involving such topics as Fourier series or Laplace transforms will need the assessment of higher-, or *create*-level, questions. We give examples of how this can be done and the sort of feedback that should be offered to reinforce the learning of conceptually-difficult material. Generally multi-choice or numerical input type questions (which serve well at lower levels) need to be augmented with other question types and/or question stem design that requires students to extract relevant material themselves, for example from a diagram. Trials have shown that whilst all of our questions were valid, some were perceived as unfair by students. Moreover, the success rate of different question types (as measured by question facility) was variable, with a new type of yes/no or true/false question testing general concepts or theorems proving to be challenging. Thus the choice of question type is important, especially in high-stakes assessments.

References

Ellis, E. 2006 Coding for an SVG graph plotter (private communication).

Ellis, E., Greenhow,M., Hatt, J. 2006 Exportable technologies: MathML and SVG objects for CAA and web content *Proc 10th CAA Conf, Loughborough, July*. http://www.caaconference.com/

Gardner-Medwin AR & Gahan M. 2003 Formative and Summative Confidence-Based Assessment, Proc. 7th International Computer-Aided Assessment Conference, Loughborough, UK, July 2003, pp. 147-155 http://www.caaconference.com

Gill, M. & Greenhow, M. 2004, Setting objective tests in mathematics using QM Perception *Proc 8th CAA Conference, Loughborough, July* http://www.caaconference.com

Greenhow, M. 1996 Computer based diagnostic tests and assessment at Brunel University, Published: Quarterly Newsletter of CTI Maths and Stats Vol 7 no 3 Aug 1996, pp 20-24.

Hatt,J. & Baruah, N. 2006 The Reconceptualisation of The Revised Bloom's Taxonomy for Use in Mathematics and its Implementation into QM Perception and Mathletics, PRHE Conference, Liverpool, May. http://hopelive.hope.ac.uk/PRHE/

Orton, A. 1983 Students' understanding of integration. Educational Studies in Mathematics, 14, 1-18

Schechter, E. 1994 The Most Common Errors in Undergraduate Mathematics, http://www.math.vanderbilt.edu/~schectex/commerrs/

Turns, J., Atman, C. J., and Adams, R. 2000 " Concept Maps for Engineering Education: A cognitively motivated tool supporting Varied Assessment Functions" IEEE Transactions on Education, v. 43, no 2, pp. 164-173.