Testing to a Purpose: Assessing the Mathematical Knowledge of Entering Undergraduates

Peter Galbraith, The University of Queensland
Michael Pemberton, The University of Queensland
Christopher Haines, City University, London

Two circumstances of major significance are presently impacting on undergraduate mathematics courses. These are respectively increased participation resulting in a wider spread of abilities among entering students, and the increasing use of symbolic algebra software in course delivery. This paper reports on preliminary work in projects at two Universities. Responses to sample questions are discussed in the context of their purpose which is motivated by the need to address the dual circumstances indicated above.

Studies addressing concerns about the mathematics knowledge base of entering undergraduates have been reported from time to time (Buckland, 1969; Gray, 1975; Clement, Lochead, and Soloway, 1980; Galbraith, 1982; Tall and Razali, 1993). Gray, for example, wrote of misconceptions, misguided and underdeveloped methods, and unrefined intuition that appeared to survive assignments, corrections, solutions, tutorials, and examinations. A common thread in all studies was the negative influence of fragmented learning and the apparent absence of cognitive strategies to co-ordinate and test the consistency and validity of accumulating knowledge.

The entering characteristics of undergraduates have become an even greater matter of interest because of contemporary changes in the intake and delivery of tertiary courses. Firstly the increased intake into universities means that entering students are being drawn from lower performance bands than previously. In Queensland for example selection is based on an overall position rating (OP). Whereas in the recent past undergraduates at the University of Queensland have been drawn from bands OP1 to OP9, in 1996 students from bands as low as OP15 have been awarded quota places. Hence the disparity in entering performance level has been greatly exacerbated with flow on implications for first year teaching.

The second contextual event of significance is the increasing use of symbolic algebra software (Derive, Maple, Mathematica) in undergraduate teaching. Fey (1989) captures a significant point when he notes that effective use of technology in learning involves more than projecting from a pen and paper medium into a computer based context on the basis of our intuitions. New elements are introduced into the teaching-learning situation by the use of computer technology and their effects need to be researched, not assumed.

With respect to the first challenge it becomes necessary to establish entering knowledge levels from a wider perspective than formal Year 12 test results. Many students taking undergraduate mathematics bring a background that is both diverse in its scope and variable in its recency (mature age entry). Hence screening programs need to provide for a much wider spread than the lowest common denominator previously associated with formal year 12 testing procedures. It is with the purpose of distinguishing lower levels of intake proficiency that we are partly concerned in this paper. (It is also acknowledged that some universities have been grappling with this issue for some time.)

With respect to the second issue (readiness for technology based instruction) one particular attribute stands out as central to learning based around symbolic algebra software. This is the capacity to understand, co-ordinate and apply alternative symbolic representations, in particular the algebraic and graphical representations of functionality.

Two recent studies provide evidence supporting the importance of algebraic/graphical inter-relations. Tall and Razali (1993) report on the performance of students comprising a diagnostic and remedial group of undergraduates at the Universiti Teknologi, Malaysia. Among a variety of findings they reported that students generally had difficulty in co-ordinating graphical and algebraic representations when slight variations were introduced into the context. Boers and Jones (1994) describe the voluntary use of graphics calculators in a traditional calculus examination. While the calculator was used extensively throughout the course teaching, and its use was built in to the instructional materials, analysis of the students' solutions indicated that, contrary to expectations, the graphics calculator was under-utilized by most students. The capability of the students to deal simultaneously with graphical and algebraic information from two independent sources was identified as the main obstacle to effective use. The authors make the significant point that the ability to integrate algebraic and graphical information will not manifest itself just because a student has the technology. 'It is a learnt skill which needs to be consciously nurtured.'

Knowledge base of new undergraduates

To address the problem outlined as the first issue above a multiple choice entry test was designed for administration during the first tutorial hour of the subject MP105 *Mathematical Methods and Applications* at the University of Queensland. This subject draws students from at least seven faculties and in 1996 has an enrolment of 252. It is therefore a subject for which the clientele forms a representative sample of the variety of students undertaking tertiary mathematics.

The test comprised 20 multiple choice questions designed in four clusters viz algebra, co-ordinate geometry, indices and logarithms, and simple calculus. In the main the items addressed knowledge from junior mathematics apart from several simple one-step questions devised around calculus operations. The alternatives were chosen so that the distribution of responses would provide information as to the extent of popular but incorrect approaches.

Some examples are provided below together with the percentage of correct responses, the most popular incorrect response, and an indication of the flawed approach most likely to be responsible for the incorrect response.

1. $(5a^3)^2$ was correctly given as $25a^6$ by 67% of students. The most popular incorrect response was $25a^5$ given by 15% of students. The probable reasoning behind the incorrect response involves adding indices.
2. Solutions to the equation $x(x+2)=3$ were given correctly by 64% of students. One of two equally popular incorrect responses was $x=1, 3$ given by 10% of students. The probable reasoning behind this response is that either $x=3$ or $x+2=3$.

3. The equation to a circle centre (1,2) and radius 3 was correctly given by 43% of students. The most popular incorrect response was $(x+1)^2 + (y+2)^2 = 9$ given by 23% of students. The most probable reason behind this response is flawed rote learning.
4. The form $\ln y = \ln 4 + \ln x$ was correctly equated with $y = 4x$ by 30% of students. The most popular incorrect response was $y = 4+x$ given by 25% of students. The most probable reason involves cancelling ln.
5. The derivative of $e^{2x} + 1$ was correctly obtained by 53% of students. The most popular incorrect response was $2xe^{2x-1}$ (23%) which represents a grotesque attempt to make use of the form nx^{n-1}.
6. The value of $\int_1^2 x^2\, dx$ was given correctly by 50% of students. The most popular incorrect response was 3 given by 21% of students. This result is achieved by substitution of the terminals directly into the integrand.

These illustrations verify that broken knowledge has fundamental roots beyond senior school mathematics courses. A question of interest is the extent to which symbolic algebra software, by performing many routine tasks, provides remedial instruction to the learner, or whether its use produces, for these students, an even wider disparity between their mathematical world and the world of professional mathematics.

Assessing algebra and graphical linkages

In designing items to probe algebraic and graphical linkages we accept the view that knowledge may be usefully classified in terms of concepts and procedures (Anderson, 1990; Hiebert and Lefevre, 1986) and that retrieval of knowledge from memory storage may be understood similarly.

Conceptual knowledge is assumed stored as a linked network of individual units, where the more elaborate the network the more points of access there are for activation to occur. Inadequate conceptual knowledge means that needed information cannot be located (blank response) or that some inappropriate version will be activated (1/A written as the inverse of matrix A).

Procedural knowledge involves the execution of a routine (rule) in response to an activating condition.

e.g. If $ax_1 = ax_2$ and $a \neq 0$ (condition) then $x_1 = x_2$ (rule)

Flawed procedural knowledge may be a consequence of either a mis-applied or wrongly remembered condition (losing a solution from $x(x-1) = x(2x-3)$ through cancelling x) or an incorrect rule $(x - a)^2 = x^2 - a^2$.

In performing mathematics students may simply need to complete a *mechanical* routine. At a higher level they may need to *interpret* information to facilitate a conceptually based conclusion. At a higher level again they may need to *construct* a solution that involves the creation of links and integrations between concepts and procedures that must be themselves generated by the solution process.

Consequently we believe that mathematical knowledge can usefully be considered in terms of these three conceptions 1. mechanical 2. interpretive 3. constructive.

In using this classification resulting test items can play several roles:

- as a measure of entering mathematical competence. In this role they serve a *diagnostic* purpose in which performance patterns identify areas of strengths and weaknesses on entry.
- as measures (pre and post) to assess the overall impact of a teaching program in deepening the understandings and capacities of students on basic concepts and procedures that underpin its structure.
- when designed (as in the present case) to address concepts and procedures that are specific targets of computer based learning sequences, to assess the initial state of knowledge with respect to these concepts and procedures.

Sample test items

Mechanical: the items in this group require the performance of a standard procedure that is cued in the wording of the question.

M. The graph with equation $y = 2x^2 - bx^3$ cuts the X-axis at X=4. The value of b is

A 4 B 2 C 0 D 1/2 E none of these.

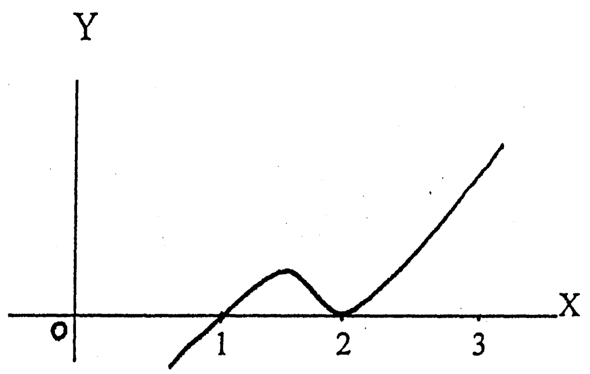
The specific cue to 'find b' is provided. It will be recognised that the items discussed previously from the University of Queensland test are mechanical items.

Interpretive: the items in this group require the retrieval of conceptual knowledge and its application to identify a correct alternative. Procedures as such are not involved.

I. Which of the following could be the equation of the graph shown (Figure 1)?

A. $y = (x-2)^2 (1-x)$ B. $y = (2-x)^2 (1-x)$
C. $y = (x-2)^2 (x-1)$ D. $y = (x-1)^2 (x-2)^2$
E. none of these

Figure 1. Diagram for item I



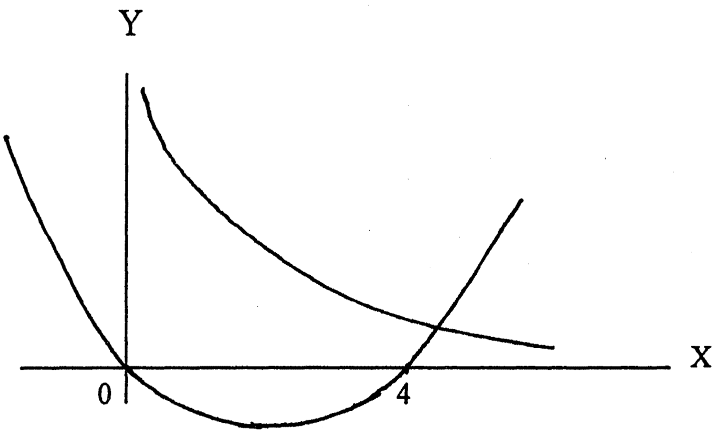
Reasoning such as the following is required. Since a double root occurs at x=2 and a single root at x=1 the equation is either $y = (x-2)^2 (1-x)$ or $y = (x-2)^2 (x-1)$ noting the equivalence of A and B. Since for large x the graph behaves like $y = x^3$ (not $y = -x^3$) alternative C is selected.

Constructive: the items in this group involve the use of both conceptual and procedural knowledge but in this case necessary procedures must be identified and introduced by the student - they are not cued. Responses involve the construction of a solution rather than the selection of an alternative.

C. The equations of two graphs are $y = 3/x$ and $y = x^2-4x$. Obtain a cubic equation where solution gives the x-coordinate of the point(s) of intersection of these two graphs. How many positive roots does this equation have?

The solution involves: recalling that the required equation is obtained by equating $3/x$ and x^2-4 (concept), simplifying $3/x = x^2-4x$ to provide a cubic equation in some form (procedure), sketching a graph such as Figure 2 (procedure), and recognizing that one intersection to the right of 0 means one positive root (concept).

Figure 2. Diagram for item C



More examples of all three item types are provided in Galbraith and Haines (1995). All items used content from preparatory secondary mathematics courses.

Preliminary Outcomes

Results from 244 entering mathematics undergraduates at City University, London were obtained for six questions of each type on an entry text. It was predicted that student performance would follow the pattern score (mechanical) > score (interpretive) > score (constructive). The results are summarized in Table 1 below after removing one a-typical outlier from each set.

Table 1. Summary response data (N = 244)

	Mechanical (5 questions)	Interpretive (5 questions)	Constructive (5 questions)
Mean score	0.59	0.39	0.28
Standard deviation	0.15	0.09	0.15

The table entry 0.59 means that on average 59% of students obtained the correct answer to mechanical questions and similarly for the others. Looked at another way we can say that overall an entering undergraduate student has a probability of 0.59 of scoring correctly on the mechanical questions etc.

Given that the challenge increases from mechanical through constructive as predicted, the greater demands on the capacity to relate graphical and algebraic representations found in the interpretive and constructive questions suggest that this type of item development will be useful in identifying knowledge factors important when teaching occurs per medium of symbolic algebra software.

The test items discussed in this paper have been developed during the preliminary stages of two projects. The items from the Queensland test will be recognised as *mechanical* in terms of the definitions above. Design, refinement, and testing will continue in the course of efforts to address the two issues facing undergraduate teaching described in this paper.

So far the design of *mechanical, interpretive,* and *constructive* questions in terms of constructs of conceptual and procedural knowledge have proved useful, both for the targeting of specific attributes involved in the co-ordination of algebraic and graphical representations, and by way of explaining the choice of incorrect solutions in terms of inadequate knowledge networks, improperly learned connections, or faulty production rules. These themes will be elaborated further in the presentation.

References

Anderson, J.R. (1990). *Cognitive psychology and its implications* (3rd ed). New York: W.H. Freeman and Company.

Boers, M.A., & Jones, P.L. (1994). Students' use of graphics calculators under examination conditions. *International Journal of Mathematics Education in Science and Technology, 25*(4), 491-516.

Buckland, P.R. (1969). The mathematical background of teachers in training. *The Mathematical Gazette, 386*, 357-362.

Clement, J., Lochead, J., & Soloway, E. (1980). Positive effects of computer programming on the students' understanding of variables and equations. *Cognitive Development Project.* Department of Physics and Astronomy, University of Massachusetts.

Fey, J.T. (1989). Technology and mathematics education: A survey of recent developments and important problems. *Educational Studies in Mathematics, 20*, 237-272.

Galbraith, P.L. (1982). The mathematical vitality of secondary mathematics graduates and prospective teachers. *Educational Studies in Mathematics, 13*, 89-112.

Galbraith, P.L., & Haines, C.R. (1995). Students' mathematical characteristics: Some developmental skills for undergraduate programs. *Bulletin of the Institute of Mathematics and its Applications, 31*, 175-178.

Gray, J.D. (1975). Criticism in the mathematics class. *Educational Studies in Mathematics, 6*, 77-86.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and Procedural Knowledge: The case of mathematics*, (pp.1-27). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tall, D., & Razali, M.R. (1993). Diagnosing students' difficulties in learning mathematics. *International Journal for Mathematical Education in Science and Technology, 24*(2), 209-222.