Research-Led Policy Change for Technologically-Active Senior Mathematics Assessment

Kaye Stacey
University of Melbourne
<k.stacey@edfac.unimelb.edu.au>

Barry McCrae
University of Melbourne
<b.mccrae@edfac.unimelb.edu.au>

Helen Chick
University of Melbourne
<h.chick@edfac.unimelb.edu.au>

Gary Asp
University of Melbourne
<g.asp@edfac.unimelb.edu.au>

David Leigh-Lancaster
Victorian Board of Studies
<leigh-lancaster.david.d@edumail.vic.gov.au>

This paper describes a new research project, which aims to investigate how computer algebra systems can be introduced into the Victorian Certificate of Education mathematics subjects. Lessons drawn from the experiences of the introduction of scientific and graphics calculators are reported along with current moves in some other countries. Options for the role to be played by technology in assessment are canvassed. It is proposed that technology should be permitted in all components of assessment but that explicit attention will need to be given to including questions that test algebraic insight. A plan for minimising inequity is proposed.

For thirty years, new information technology has been a major force driving change in mathematics curriculum and assessment. In the 1960s the introduction of electric calculators changed the way that arithmetic was done in business. As a consequence, the goals of the mathematics curriculum needed to be thoroughly revised. Later, when hand-held electronic calculators became affordable for students at school, curriculum and assessment began to change at all levels — a change involving such a fundamental re-assessment of goals and teaching methods that it continues today. Adapting to the technology of the four-function calculator was, for example, a major theme of the 1998 NCTM Yearbook (Morrow & Kenney, 1998). School systems around Australia are now seriously grappling with the use of graphics calculators in senior school mathematics teaching, curriculum, and assessment. We see this as a transient stage, simply the prelude to the next phase of technology-driven challenges and opportunities for teaching mathematics, which will centre on computer algebra systems.

Computer algebra systems (CAS) can perform all the routine procedures of mathematics normally covered in secondary school and university, including drawing graphs, calculating with vectors and matrices and doing algebra, calculus and statistics. These systems have existed for over twenty years, even on small systems such as home and school computers, but until the recent advent of hand-held machines of moderate cost, they have had little effect on school mathematics. Even at the earliest stages, it was predicted that there would be little impact on school and university teaching until hand-held technology arrived. Wilf, in his 1982 article evocatively entitled "The Disk with the College Education", looked forward to the day when students might have a $39.95 pocket calculator with a long LCD window for display and lots of buttons such as DERIV (for differentiation) and MATRIX, performing operations from the major areas of senior secondary and early tertiary mathematics.

Although the $39.95 version has not yet literally come, CAS "supercalculators" should soon be priced such that all Australian senior mathematics students might own one. The potential effect of the CAS revolution is dramatic because of its broad application across mathematics and because it directly involves the secondary-tertiary interface. As the availability of CAS in schools increases, so does the need for the official curriculum to produce appropriate policy responses. Effective policy development on the possible use of CAS in formal examination-based assessment requires rigorous and well-grounded research, especially because this is associated with the highly sensitive credentialling arrangements for tertiary study.

Researching Possible Curriculum Change

This paper outlines some preliminary issues to be addressed in a new study funded from 2000–2002 by the Australian Research Grant Strategic Partnerships with Industry Scheme. The study aims to investigate the changes that regular access to CAS supercalculators will have on senior mathematics subjects and the associated assessment in Victoria, Australia and to explore the feasibility of offering new mathematics subjects which use CAS extensively. In Victoria, schools provide Year 11 and 12 subjects for the Victorian Certificate of Education (VCE) in accordance with the 'study design' that specifies curriculum content and assessment procedures. Final results, which are used for various purposes including selection to tertiary courses, are derived from a mix of tasks set at the school and external examinations. The Chief Investigators of the project are Gary Asp, Helen Chick, Barry McCrae and Kaye Stacey; and David Leigh-Lancaster is a partner investigator. The four industry partners are the Board and three calculator suppliers and manufacturers: Hewlett-Packard; Shriro (Casio); and Texas Instruments. The industry partners will supply CAS supercalculators to students in three schools for a three year program of classroom based research. With the cooperation of the Board, the content and formal assessment undertaken by these students for Year 12 will be altered, culminating with the trial in volunteer schools in 2002 of a VCE examination using CAS, an alternative to the current Mathematical Methods subject. The major features are summarised in Figure 1. Actual implementation of this requires the Board's continuing approval, which will be informed by the findings of the project in earlier years. The outcome will be substantial policy advice to the Board, practical feedback to the calculator manufacturers, and important research insights into questions of learning mathematics in a technology-rich environment.

For each project school CAS support from one industry partner

2000	2001	2002
One Yr 10 Maths class ➔	➔ Yr 11 Maths Methods ➔	➔ Yr 12 New CAS Maths Assessment: New Board CAS paper
One Yr 11 Maths Methods class ➔	➔ Yr 12 Maths Methods Assessment: Standard VCE + trial CAS paper	
	One Yr 11 Maths Methods class ➔	➔ Yr 12 New CAS Maths Assessment: New Board CAS paper

Figure1. Time-line of research and assessment changes for each project school.

For the calculator industry partners, the project enables their products to be tested seriously within an Australian curriculum context, investigating the suitability of the products' capabilities, interfaces, notations and physical characteristics. Materials for training teachers to use the CAS supercalculators will be enriched by experiences in the project schools. For the Board, the project will result in advice to support the development of policy for curriculum and assessment, covering issues such as:

- which parts of Year 12 assessment could permit CAS, forbid it or require it;
- the protocols for the use of CAS supercalculators in assessment;
- how (and if) examination questions can be set tc be fair to users of CAS supercalculators of different brands and models, given that they have different capabilities (the involvement of different industry partners is critical for this); and
- subsequent redevelopment of lower secondary mathematics curricula and suggestions for post-secondary mathematics and mathematics-related studies.

Beyond these practical issues, the scientific importance of the study lies in its advancement of our detailed knowledge of how students learn mathematics and how it can best be taught. This new learning environment provides us with new opportunities for research into the teaching of mathematics. When students have CAS in class and in examinations, the need for memorising routine procedures may be enormously reduced, yet the need for conceptual and structural understanding is almost certainly undiminished and possibly enlarged.

Effective policy development (in particular for assessment used for university selection and pre-requisites) requires rigorous and imaginative research, especially as the possible changes are likely to be substantial and subject to robust debate. We know of no similar study being undertaken anywhere in the world. Thus, the results of this research are likely to have implications for education systems throughout Australia and, indeed, the rest of the world.

In the rest of this paper, we present our preliminary thinking on one of the major policy issues to be resolved through the project: the way in which assessment will use or not use CAS. We recognise that this is likely to be a sensitive and possibly political issue and so careful detailing of arguments and analysis of data will be needed to guide policy.

Senior Mathematics Assessment with Technology

In this section, we review the experiences of adapting assessment in Victoria to graphics calculators and report on some current policies from other countries on adapting assessment to the presence of CAS.

The Graphics Calculator Experience

The CAS revolution (if that is what it proves to be) for senior mathematics comes hard on the heels of the introduction of graphics calculators. Out of reach because of price just a few years ago, schools in many places have now enthusiastically taken them up. Systemic education curriculum and assessment authorities, internationally and nationally have, over the past decade, grappled with policy issues related to graphics calculators (Stacey, Dowsey, McCrae, & Stephens, 1998).

Within Australia, the state education systems have now specified whether graphics calculators are permitted, are banned or are compulsory in examinations and other assessment. As of March 2000, Victoria and Western Australia have adopted graphics calculators, while the other states with an external examination system have not. In Victoria, the policy of the Board is one of "assumed access" to graphics calculators for all year 12 examinations by 2000. The adoption of graphics calculators by the Board required change in the syllabus for each mathematics subject and change in the type of examination questions used. Questions that are "trivial" with a graphics calculator have been replaced by questions that probe conceptual understanding and problem solving with technology. The above changes are being introduced smoothly at least in part because of extensive professional development and strong support of the Mathematical Association of Victoria. However, these changes are just a prelude to the much larger policy issue of CAS.

In Victoria, the transition to using graphics calculators in Year 12 external examinations was made in two stages: firstly (in 1997 and 1998 for the main tertiary entrance subjects) the examinations were set in a "technology neutral" fashion. This meant that a student without a graphics calculator should not be disadvantaged on any of the questions. So, for example, in 1996 before graphics calculators were permitted, one multiple choice question on the Specialist Mathematics examination (CAT 2) asked about the implied domain of the function $f(x) = 1 + \mathrm{Sin}^{-1}(2x)$. In 1997, a similar question asked about the function $f(x) = \mathrm{Cos}^{-1}(x - a)$; calculator neutral but somewhat harder. The transition has now been made to "technology-active" assessment, where not using a graphics calculator on some questions may definitely disadvantage a student. The Board (1999) provided examples of the new range of questions which could be asked, including sketching the graph of the function $v = 100 \tan(\mathrm{Tan}^{-1}(3) - 25000t)$.

Systemic Adoption of CAS

Examining mathematics with access to CAS presents more challenges than examining mathematics with graphics calculators. McCrae (1996) found that access to graphics calculators would impact on only 6% of a 1994 VCE Specialist Mathematics paper (the hardest mathematics subject in Victorian schools), but that about 60% would be affected by the availability of CAS. Similarly, Shumway (1989) reported that about 90% of the exercises in most U.S. textbooks could be computed directly by CAS. The adoption of CAS in teaching is therefore inherently associated with the adoption of CAS in assessment.

As yet, only a few countries around the world have national policies permitting the use of CAS in examinations. In Denmark, CAS will soon be permitted in all mathematics examinations for 15-19 year old students. In France, any calculator, including those with CAS, is permitted in examinations. In parts of Germany, teachers can decide whether CAS is permitted in lessons and examinations. Elsewhere around the world, some countries have now permitted graphics calculators but not CAS (as in Victoria), some (such as Italy and Ireland) are about to introduce pilot projects, and some, often in Asia but also including poorer nations, do not use calculators at all. The unique nature of each country's assessment regime and the national priorities makes research in individual countries important.

One policy likely to be influential in Australia is that of the USA College Board Advanced Placement (AP) Calculus (see website). The long list of calculators permitted includes both graphics and CAS supercalculators, although those with QWERTY keyboards, pocket organisers and pen-input computers are not allowed. From 2000, both multiple-choice and free-response sections will be in two parts: one where some parts require the use of a graphics calculator and one where the use of any calculator is not permitted. The AP web-site notes that: "This change in format is an effort to respond to heightened concerns with equity as more students may use graphing calculators with computer algebra system (CAS) features." It is claimed that the two part format will provide flexibility in the types of questions that can be asked and also ensure greater fairness to students, regardless of calculator used. In addition, it is specified that students can only use a calculator for three operations: solving an equation; finding a derivative; or calculating the value of a definite integral. In these cases, the student must indicate the set-up of the problem (e.g., to write down the relevant definite integral for finding an area before evaluating it by calculator). In all other cases of calculator use, the student must show the "mathematical steps necessary to produce the results". For example, to find a local minimum of a function, normal calculus procedures must be followed to establish the derivative and set it to zero. Many of the capabilities of CAS therefore cannot be used.

Technology-free, Technology-neutral or Technology-active?

Stephens and Leigh-Lancaster (1997) write that there are three possible positions that should be investigated for the use of CAS technology in examinations:

- that assessment (of at least some areas of mathematics) should be *technology-free*, i.e., that students should not use technology (or at least advanced or "new" technology) in the assessment;
- that assessment should be *technology-active*, i.e., that students should be permitted to use specified advanced or 'new' technology in assessment and that some questions should be designed to require its use; and
- that assessment should be *technology-neutral*, i.e., that students not using technology in an assessment should be able to answer questions as easily as those using technology.

As noted above, the Victorian mathematics subjects have in recent years passed from technology-free (with regard to *graphics* calculators, although scientific calculators have long been permitted), to technology-neutral, to technology-active external assessment. This experience has led us to believe that the technology-neutral position is certainly not sustainable in the long-term, as it is very hard to set questions which are genuinely technology-neutral (McCrae, 1996). In this we agree with Kemp, Kissane and Bradley (1996) who note in regard to graphics calculators that:

> "... the use of calculator neutral examinations is an unwise long-term strategy, although it may be seen as helpful in the short term to allay concerns about disparities in student access to graphics calculators. In the long term, such a strategy would send a clear (and incorrect) signal that graphics calculators are not of importance in mathematics, and would discourage both students and their teachers from acquiring either hardware or expertise in its use."

In designing its examinations, AP Calculus has combined the three assessment positions. Technology-free assessment has been chosen for part of the assessment and the other part is technology-active assessment, partially neutralised. It is technology-active in the sense that a CAS calculator is permitted and some parts of some questions require its use. On the other hand, the use is constrained to three operations. They have given two reasons for this choice (http://www.collegeboard.org/ap/calculus/html/exam002.html): firstly, the recent introduction of graphics calculators has placed a large burden on teachers, who need time to adjust their courses. Secondly, the College Board notes that it can develop fair examinations with any specified type of technology, but it cannot develop exams that are fair to all students if the spread of capabilities of the technology is too wide. The College Board has endorsed the use of any technology and promoted it in teaching—accepting that some students will have CAS whilst others have basic graphics calculators—but constrained their use in assessment in order to increase equity.

Our initial position (which may be modified in the light of experience and data) is that our project should aim for technology-active assessment only. The reasons for this initial decision are explained in the next section. However, we expect that this position will be feasible provided only a narrow range of (high) technology capabilities is permitted and if teachers are provided with adequate professional development before they begin to teach the new subjects. Instead of the wide range of calculators permitted by the College Board, we expect to recommend machines with strong commonalities. With time, the capability of these machines will increase, as more advanced technology becomes affordable. Furthermore, if the new subjects are initially alternatives to the current subjects, teachers can decide to move their classes across to the new syllabus when they are themselves professionally ready and when typical students in their school community can afford to purchase an adequate machine. We see serious equity considerations arising from the cost of new technology which cannot be solved, only reduced, in the foreseeable future.

Although we do not intend to investigate technology-neutral assessment (for reasons given above), one of the features of the research design is that it will allow us to fully explore the practicality of *brand-neutral* assessment. As hand-held machines become more and more powerful, the divergence between the capabilities of different models may increase markedly (alternatively it may decrease with "product maturity"). Current experience with graphics calculators shows that questions that are challenging on one machine can be simple on another. For example, an important point of an assessment item may be to test students' understanding of the shape of a graph. Sometimes this can be done by using a graph that does not appear in a standard viewing window, so that the student needs to know properties of the graph to locate a desired feature. However, facilities such as determining complex arithmetic or automatically locating zeroes and turning points are available on some calculator, making certain questions straightforward. The capabilities of different models will need to be fully appreciated in order to set brand-neutral assessment items.

Why Plan only Technology-active Assessment?

Our initial position is to create technology-active assessment, rather than technology-free or technology-neutral assessment. This does not mean that every question will require use of technology, although it would be available for all. Our primary reason relates to values and beliefs about mathematics: mathematics at school should be like mathematics as used outside school. As mathematics outside school changes and the methods of choice change, so the methods of choice at school should also change — to the extent that this is possible. Furthermore, we believe that assessment should be aligned with teaching as closely as possible. This principle has been endorsed by national associations for many years (see, for example, AAMT/CDC, 1987) We believe that a technology-free component of assessment would endanger mathematics remaining a sensible subject where students learn to use up-to-date methods. On balance, we believe that the decision to make senior school mathematics arithmetic-calculator-active has been the right one and that this will also be the right decision for algebra. Our final argument for technology-active assessment only is that this maximises the assessment time that is spent on assessing higher-order thinking skills, rather than routine procedures.

What are the arguments for a technology-free component of assessment? One is the need to provide fair assessment for students with different types, models and brands of calculators and we have outlined above our approach to minimising this concern. The research project can provide some data to judge this approach. A second argument is that tertiary mathematics courses may not use CAS and hence students must not rely on it. We expect that this will not be the case for much beyond the time frames of the proposed curriculum change. A third argument is that the existence of technology-free assessment would encourage the acquisition of important mental or by-hand skills and a fourth is that technology-free assessment is better able to test 'true understanding' — what students really know and really understand. These two final arguments are substantial, getting to the heart of values and beliefs about mathematical activity. How can we respond to them?

How can Mental and by-hand Skills be Encouraged?

We take it as axiomatic that an important function of an external examination system is to encourage good learning and teaching. We agree that there are skills (especially algebraic skills) over which students should have personal mastery. Moreover, we agree that the existence of technology-free assessment would encourage the acquisition of these important mental or by-hand skills. Years of teaching with scientific calculators have shown that to achieve the goal of sensible and powerful mathematics, students cannot be only taught how to carry out arithmetic on a machine. They must develop a strong number sense, which enables them to operate quickly and effectively in the world of informal arithmetic (e.g., "Am I being charged about the right amount here?") and also to operate a machine competently, guarding against errors by effortlessly monitoring the results of calculations. At the same time, there has been a re-assessment of what students should know and by-hand procedures such as taking square roots and division using log tables have been consigned to history. Parallel moves will need to be undertaken with CAS — students will need well-developed symbol sense (Arcarvi, 1994), a mastery of some simple procedures (e.g., expanding with the distributive law, doing the same to both sides for equation solving) and some currently taught procedures will not be required (is factorising one of these?) Our initial position, however, is that separate technology-free assessment is not required. If algebraic insight (symbol sense) is made a goal of the assessment, we expect that it can be assessed as well with calculators as without, by clever question design. On the other hand, we expect that mastery of simple procedures is absolutely essential for smooth CAS use and therefore, students who do not have it will not do well on a technology-active examination. This will be implicitly tested, rather than explicitly. It will therefore be important to be quite explicit about these issues in advice to teachers.

In the argument above, we have assumed that only simple procedures will be important to be able to do by-hand as a mathematics professional or user in the future. Exactly what constitutes a simple procedure will vary from person to person, as it does now with arithmetic. This however, may be a vigorous area of debate for many years. Kutzler (1999) in proposing the idea of assessment of "intellectual fitness" is beginning this debate.

Does True Understanding Depend on being able to Carry out the Steps?

The fourth argument for technology-free assessment is that it is better able to test "true understanding" than is technology-active assessment. In fact, it is tempting to believe that what a person 'really understands' what he or she can do from memory or "by-hand". Pierce (1999) has shown how students themselves often believe this. Unfortunately, however, dissatisfied mathematics students across generations have attested to the fact that they did not understand procedures that they learned to carry out by hand successfully. Technology-free assessment therefore cannot claim to reliably identify 'true understanding': the question is whether technology-active assessment can identify it and encourage it any better. Experience with assessing with graphics calculators leads us to expect that careful question design may be able to achieve this. However, the question of what will constitute understanding (with possibly different answers for different types of users of mathematics) is a much larger question and we cannot address it adequately in this paper.

Conclusion

Year 12 mathematics is critical to prepare Australian young people for a technological future. Mathematics subjects are the principal pre-requisites for tertiary study in the sciences, engineering and economics, on which Australia's economic competitiveness depends. It is clearly important that our students learn to use the mathematical tools of the future, but also that a cautious approach needs to be adopted so that the right balance between traditional by-hand algebra skills and technology use can be achieved.

References

Australian Association of Mathematics Teacher/Curriculum Development Centre (1987) *A national statement on the use of calculators for mathematics in Australian schools*. Adelaide: AAMT.

Arcarvi, A. (1994). Symbol sense: informal sense-making in formal mathematics. *For the learning of Mathematics*, *14*(3), 24-35.

Board of Studies (1999) *VCE Bulletin Supplement 1.* Melbourne: Publisher as author.

Kemp, M., Kissane, B. and Bradley, J. (1996). Graphics calculator use in examinations: Accident or design? *Australian Senior Mathematics Journal*, *10*(1), 36-50.

Kutzler, B. (1999). The algebraic calculator as a pedagogical tool for teaching mathematics. (Keynote address) *Investigating with information technologies*. Australian Association of Mathematics Teacher Virtual Conference 1999.

McCrae, B. (1996). The use of calculators in Victorian Certificate of Education examinations: Looking ahead. *Australian Senior Mathematics Journal* *10*(1), 65-71.

Morrow,L.J. & Kenney, M.J. (1998) *The teaching and learning of algorithms in school mathematics*. (1998 Yearbook) Reston, VA; National Council of Teachers of Mathematics.

Pierce, R. (1999). Computer algebra systems facilitate positive learning strategies. In Truran, J. and Truran, K. (Eds) *Making the difference*. Proceedings of the 22nd annual conference of the Mathematics Education Research Group of Australasia pp. 431-438. Adelaide: MERGA.

Shumway, R. (1989). Supercalculators and research on learning. *Proceedings of the 13th conference of the International Group for Psychology of Mathematics Education*, Vol.3, pp. 159-165. Paris: Editions GR Didactique et Acquisition des Connaissance Scientifique.

Stacey, K., Dowsey, J., McCrae, B., & Stephens, M. (1998). *Review of senior secondary mathematics curriculum.* Sydney: NSW Board of Studies.

Stephens, M., & Leigh-Lancaster, D. (1997) Using mathematically able software in assessment. *Australian Senior Mathematics Journal, 11*(1), 5-19.

USA College Board (1999) *Advanced placement (AP) calculus.* (http://www.collegeboard.org/ap/calculus/index.html) accessed March 24, 2000.

Wilf, H.S. (1982). The disk with the college education. *American Mathematical Monthly, 89*, 4–8.