The Victorian Curriculum and Assessment Authority (VCAA) Mathematical Methods Computer Algebra System (CAS) Pilot Study Examinations 2003

Michael Evans
Scotch College
<michael.evans@scotch.vic.edu.au>

Pam Norton
Monash University
<pam.norton@sci.monash.edu.au>

David Leigh-Lancaster
Victorian Curriculum and Assessment Authority
<leigh-lancaster.david.d@edumail.vic.gov.au>

The first set of examinations, for the original cohort of 78 students from three volunteer schools, were conducted in November 2002 and some initial analysis of student performance, in particular with respect to items common with the standard course, was reported at the 2003 MERINO conference. This paper further investigates student performance on this aspect of examinations for 270 students of the expanded pilot program in November 2003.

Mathematical Methods (CAS) Units 1 - 4 is an accredited pilot study of the Victorian Curriculum and Assessment Authority January 2001 – December 2005 (VCAA, 2004). The first phase of the pilot study 2001 – 2002, involved students from three Stage 1 volunteer schools, and was implemented in conjunction with the CAS - CAT project 2000 – 2002 (DSME, 2004). In November 2002, 78 students from the three Stage 1 volunteer schools sat end of year Mathematical Methods (CAS) Unit 3 and 4 examinations, for which student access to an approved CAS calculator (TI-89, CASIO ALGEBRA FX 2.0 or HP 40G) was assumed. Some preliminary investigation of these examination results, in particular comparative analysis with respect to student performance on common questions with the corresponding standard Mathematical Methods course (function, algebra, calculus and probability), can be found in Evans, Leigh - Lancaster and Norton (2003). This paper reports on further investigation of results from the two November 2003 examinations for the standard Mathematical Methods cohort of around 17 600 students, and the Mathematical Methods (CAS) pilot cohort of around 270 students, on common questions for these studies. This investigation considers whether there are any indications that regular access to CAS may have facilitated or restricted student conceptual understanding and facility with mathematical skills and processes. As more systems and jurisdictions move to incorporate the regular use of CAS technology in at least some aspects of their curriculum, pedagogy and related assessments, such investigations are significant as part of broader consideration of the real and perceived advantages and disadvantages that might be associated with the widespread use of such technology in senior secondary mathematics education. CAS have now been used in academia, research, industry, business and commerce for over two decades.

Research on related aspects of CAS related to the pilot study can be found in (Ball, 2003) - how students structure responses to questions; (Flynn, 2003) - a model for analysis of the ‘sensitivity’ of questions to different hand-held CAS; and (Garner and Leigh-Lancaster, 2003) - change in teacher practice with classroom access to CAS. Following the increasingly widespread use of mathematically able software and more sophisticated calculators in secondary mathematics classrooms from the early 1990’s, there has been significant work on the notion of *instrumental genesis*. This is the process of an object becoming an instrument – part *artefact* and part *cognitive scheme* - through the dual processes of *instrumentalisation*, transforming the potentialities of artefacts for specific uses; and *instrumentation*, the development of schemes of action that progressively take shape as techniques that permit effective responses to given tasks (Artigue, 2002). However, there is, at present, only limited data and related analysis available with respect to student performance on examinations in school systems and jurisdictions which permit the use of CAS (see, for example, Brown, 2003; Burrill, Allison, Breaux, Kastberg, Leatham & Sanchez, 2002), although some further material in this area will soon be published (see Bohm, Forbes, Herweyers, Hugelshofer & Schomacker, 2004, in press). Given the small VCAA 2002 pilot cohort, the earlier analysis of these results by the authors was necessarily tentative in nature.

The expanded pilot study incorporates the original three schools (implementing Mathematical Methods (CAS) Units 1 and 2 from 2001 and Units 3 and 4 from 2002) and includes two additional groups: nine Stage 2 volunteer schools implementing Units 1 and 2 from 2002 and Units 3 and 4 from 2003, and a further seven Stage 3 volunteer schools implementing Units 1 and 2 from 2002 and Units 3 and 4 from 2004. The schools in the expanded pilot include co-educational and single sex, metropolitan and regional schools from government, catholic and independent sectors, using a range of different CAS. Thus, there were about 270 students enrolled in Units 3 and 4 from 11 schools of the expanded pilot in 2003, as shown in Table 1, including students using the CAS TI *Voyage* 200, *Derive* and *Mathematica* in one school for each of these CAS. Progress of the pilot study has been reported in Leigh-Lancaster (2003).

Table 1
Classification of Schools for Stage 2 Expanded Pilot Examinations

	Metropolitan	Regional
Government	2 (TI-89, *Derive*)	2 (TI-89, CASIO 2.0)
Catholic	2 (TI-89)	1 (*Mathematica*)
Independent	3 (TI-89, *Voyage* 200)	1 (CASIO 2.0)

For students in each VCE study, the VCAA computes a study score in the range 0 to 50, from a truncated normal distribution with mean 30 and standard deviation 7. For VCE Mathematics, this study score is based on two examinations, each worth 33% of the final weighting, and a school based coursework assessment score, worth 34% of the final weighting, and statistically moderated with respect to the examinations. The Victorian Tertiary Admissions Centre (VTAC) re-scales these study scores to take into account differences in relative difficulties of studies (based on analysis of how students perform across studies). These are then used to compute a national tertiary entrance score on a scale of 0 – 100 from a combination of best subject scores. Table 2 shows the 2002 and 2003 scaling for Further Mathematics, Mathematical Methods and Mathematical Methods (CAS) and Specialist Mathematics.

Table 2
VTAC Mathematics Study Scaling 2002 and 2003 Examinations

VCE Mathematics studies - VTAC scaling 2002									
study	mean	sdev	20	25	30	35	40	45	50
FM	27.2	6.7	18	22	27	32	37	43	50
MM	36.6	6.9	26	32	38	42	46	49	50
MM(CAS)	38.8	5.6	25	32	38	43	47	49	50
SPM	41.2	7.1	30	36	42	47	50	53	55
VCE Mathematics studies - VTAC scaling 2003									
study	mean	sdev	20	25	30	35	40	45	50
FM	27.23	6.7	18	22	27	32	38	43	50
MM	36.00	6.9	25	31	37	41	45	48	50
MM(CAS)	36.46	5.9	27	33	38	42	45	48	50
SPM	41.2	7.1	30	36	42	47	50	53	55

These show a slightly higher (re-scaled) pilot mean score with a slightly smaller standard deviation, with the Stage 2 mean score close to the Mathematical Methods mean. Thus, while the Stage 2 schools are not a stratified random sample, they do represent a broad range of backgrounds, and their overall mean performance on examinations is similar to that of the larger cohort.

The use of CAS has been permitted in some components of the US College Board's *Advanced Placement Calculus* examinations since 1995 (hand-held graphics calculator/CAS calculator in a common question, technology active, but graphics calculator/CAS neutral examination); the French *Baccalaureate Générale Mathematics* examination since 1999 (hand-held graphics calculators or CAS calculators in a pure mathematically oriented technology neutral examination) and the Danish *Bacclaureat Mathematics* examination since 1997 (graphics calculator or hand-held/computer based CAS in an open book, technology active, format with common questions and some questions with non-CAS/CAS alternative versions). The unique feature of the Victorian CAS pilot examinations is that access to an approved CAS is assumed for *all* components of these examinations. While the Mathematical Methods (CAS) study has been developed from the Mathematical Methods study it also has significant distinctive curriculum content and related pedagogy; and these features are reflected in the corresponding assessments. Mathematical Methods and Mathematical Methods (CAS) are thus parallel and alternative courses, for two distinct, but like, populations. While differences in these two populations are significant and amenable to qualitative analysis in their own right, common examination questions provide an important basis for some comparison and analysis of the performance of the two cohorts.

Mathematical Methods (CAS) Examination 1 – 2003

In 2003, 269 students sat for the Mathematical Methods (CAS) pilot study Examination 1, and 17 620 students sat for the corresponding Mathematical Methods Examination 1. The papers both comprised 27 multiple choice questions, each worth one mark, and 6 short answer questions worth a total of 23 marks. Twenty of the multiple choice questions (about 74 % of this component) were common to both papers. One short answer question (worth two out of a total of 23 marks) was common to both papers, while another question (worth 3 marks on the CAS paper and 4 marks on the standard paper) was very similar on both papers (about 22 % of the short answer component).

Discussion of Multiple Choice Questions

As in 2002, the multiple choice component was well done by the pilot cohort, with a mean of 19.1 marks out of a possible 27. In general, the Mathematical Methods (CAS) cohort performed comparably, or better, than the Mathematical Methods cohort on common multiple choice questions. Table 3 summarises the difference in percentage of correct responses. A positive difference indicates that a higher proportion of CAS pilot students selected the correct response. The questions have been classified as technology independent (I); technology of assistance but neutral with respect to graphics calculators or CAS (N); or use of CAS likely to be advantageous (C). Those items for which technology is of assistance, but that are likely to be answered efficiently by conceptual understanding, pattern recognition or mental and/or by hand approaches have been indicated with an asterisk (*).

Table 3
Summary of Differences Between Percentages of Correct Responses to Common Examination 1 Multiple Choice Items: Question Numbers.

	Negative difference	No difference	Positive difference		
Item type	Up to 3%	Same	Up to 5%	6 to 10%	More than 10%
I	5, 12, 26, 27	20	4, 11, 15, 23	7, 24	14
N			1*	8*, 16*	6*
C			3*	13*, 21*	22

On 15 of the 20 common multiple choice questions, a higher percentage of the CAS cohort obtained the correct answer. For questions which are technology *independent*, the CAS cohort performed better on seven out of the eleven questions for which there was a difference (with no difference on one question). For the four questions where technology was of assistance but *neutral* with respect to graphics calculators or CAS, the CAS cohort performed better. If a 95 % confidence interval is calculated for the percentage of correct responses to these twenty multiple choice items from a simple random sample of size 269 from the Mathematical Methods cohort, then on no question is the percentage of correct responses from the CAS pilot cohort less than the corresponding lower bound, but for 9 questions the percentage of correct responses from the CAS cohort was higher than the upper bound. Although the CAS pilot cohort cannot be considered as a simple random sample of the Mathematical Methods cohort, these comparisons indicate where the CAS pilot students differ markedly from the Mathematical Methods cohort. Not surprisingly the CAS pilot cohort also performed better on the four questions where CAS use was classified *a priori* as likely to be *advantageous*, in particular, Question 22, solving the equation: $2\log_e(x) - \log_e(x + 2) = 1 + \log_e(y)$, for y. While the relatively high proportion of correct responses, 80% for the CAS cohort compared with 52% for the standard cohort on this indicates access to CAS substantially increased accuracy and reliability on certain routine questions, there are still students for whom this is not the case. This may be due to either error in procedural understanding (what to do), and/or technical understanding (how to do it using CAS, or how to relate CAS output to the algebraic forms of the given alternatives). An important issue here for pilot teachers is to identify those pedagogical practices that are most likely to lead to *very high* levels of accuracy and reliability on these, and similar, sorts of problems.

Discussion of Short Answer Questions

The only identical common item, Question 2, asked students to find the exact solutions of the equation $\sin(2\pi x) = -\sqrt{3}\cos(2\pi x)$, $0 \le x \le 1$. Only 55% of CAS students managed full marks for this question, compared to 38% for the non-CAS students. This less than robust result for both cohorts likely reflects conceptual difficulties arising from the $2\pi x$ term, but also, for the CAS cohort, the way CAS solve equations involving circular functions (some give a parametric, or principal domain, form of solution; others require preliminary transformation by hand to an equivalent tan formulation). Thus, while CAS provides *some* assistance, it does not guarantee automatic one-step solutions to such problems but typically requires additional complementary analysis.

Question 4 was similar on both examinations. In the first part, students had to find the value of the x coordinate, $x = k$, for the point of intersection of two curves, $y = -x + 1$ and $y = 1 - e^{-x}$, numerically, correct to three decimal places ($k = 0.567$). While 87% of CAS students did this correctly, only 62% of non-CAS students did, and this difference in results is surprising, as the numeric equation solving functionality is the same on corresponding models of both CAS calculators and graphics calculators. The related functionality is not easier to use on the computer based CAS *Derive* and *Mathematica*. This should be a routine skill whatever technology is used. Perhaps the difference in cohort performance is a reflection of the explicit use of technology in the design of the CAS course, and the generally positive orientation of pilot teachers towards the active use of technology. However, there is also a strong hypothesis that it may substantially arise because CAS require and display input and output expressions in true and correct mathematical form. In natural combination, these effects would likely lead to a high level of correct entry of mathematical expressions for evaluation and interpretation. The rest of the question involved the *writing down* and *evaluating* of a definite integral. The non-CAS cohort was required to show an anti-derivative, whereas the CAS cohort was not. Of the CAS cohort, 76% were able to write down the correct definite integral for the area, $\int_0^k (-x + 1 - (1 - e^{-x}))dx$, and evaluate it (numerically, or analytically, with subsequent substitution of numerical approximations to exact real values) correct to two decimal places; compared with 34% of the non-CAS cohort.

Mathematical Methods (CAS) Written Examination 2 - 2003

In 2003, 268 students sat for the Mathematical Methods (CAS) pilot study Examination 2, and 17 620 students sat for the corresponding Mathematical Methods Examination 2. The papers both comprised four extended answer analysis questions: a probability manufacturing sampling context involving normal, binomial and hypergeometric distributions; a circular function and calculus skateboard ramp modelling context; a pure mathematical functions and calculus question involving a product function and tangents to its graph; and a drug in the bloodstream modelling context: worth 12, 13, 16 and 14 marks respectively. Question 1 was common to both papers, however, in each of the other questions while there were some common components, there were also related but distinctive formulations that reflect different emphases in each course (for example, exact and numerical answers), and, for the standard Mathematical Methods paper, the availability of the various quasi-CAS graphics calculator supplementary programs. Examination marking schemes for each study also reflect these differences, so the following comments are principally in relation to those parts which can reasonably be compared and are based on the percentage of students achieving full marks for each component of a question, as a robust measure of achieved competency. For Question 1, as shown in Table 4 below, the similarity is marked (as in 2002), despite initial apprehension by some pilot teachers that the requisite functionality on CAS calculators and computer based CAS was not as 'good' or 'easy to use' as for graphics calculators. Thus the two cohorts performed comparably on this question as expected. Indeed, some pilot teachers argue that this is a reasonable achievement in itself, as the CAS students had less curriculum time available for this material – they also studied transition matrices and Markov sequences and general probability distributions for continuous random variables.

In Question 3, the graph of the function f: $R \rightarrow R$, $f(x) = x^3e^{-2x}$ was provided on scaled axes, with the stationary point of inflection at the origin clearly shown. Part 3ci was similar, based on the equation of a tangent at $x = 1$. For the Mathematical Methods cohort, this equation was given, and students were asked to *show that* this was true, while the CAS cohort were not given the equation, but asked to find it. In 3cii both cohorts were asked to write down an equation of the tangent ($y = 0$) to the curve at the origin, while in 3ciii both cohorts were asked to show that the tangents found in parts i and ii of the question are the only tangents to the curve that pass through the origin. A comparison of performance on these parts is shown in Table 5.

Table 4

Comparison of Performance on Question 1 - Common Probability Question

Part marks	1a 2	1b 2	1c 2	1d 2	1ei 1	1eii 1	1eiii 1	1eiv 1	Total marks 12
MM (%)	72	21	60	53	90	58	48	13	MM mean
mean	1.59	0.51	1.28	1.35	0.90	0.57	0.48	0.13	total 6.81
MM(CAS)(%)	69	24	62	55	91	66	50	15	MM(CAS)
mean	1.52	0.58	1.38	1.38	0.91	0.66	0.50	0.15	mean total 7.08

Table 5
Comparison of Performance on Common or Very Similar Parts of Question 3 – Pure Mathematical Function and Calculus Question.

Part marks	3ci 3	3cii 1	3ciii 3	Part total Marks 7
MM (%)	38	48	4	MM mean
mean	1.43	0.47	0.18	part total 2.08
MM(CAS)(%)	67	72	5	MM(CAS)
mean	2.31	0.72	0.29	mean part total 3.32

Question 4 was based on the function $x = \frac{3t}{5 + t^2}$, $t \geq 0$, where t is the number of hours after the injection of the drug. The question was varied for the cohorts in terms of formulation involving exact or numerical values for solutions, however both cohorts were required to draw the graph of the inverse function on the same set of axes as the original function (an un-scaled blank set of axes was provided). This is a graphical question, with the same functionalities available for graphics calculators and CAS – both will draw the inverse of a function *without* needing to have its rule determined explicitly. Indeed, this question can perhaps be best answered without the use of technology at all: using a 1 – 1 scale, a fold along the line $y = x$, and some careful tracing. Only 4% of the Mathematical Methods cohort answered the question correctly (mean score of 0.76 marks out of 3 marks); while only 15% of the Mathematical Methods (CAS) cohort answered the question correctly (mean score of 1.01 marks out of 3 marks). In both cases this indicates that where the functions and graphs involved relate to more complex modelling, even the *graphical* determination of the inverse function present challenges for most students. For the CAS cohort only, the last section of this question required analysis based on a parameter p instead of the constant 5 in the denominator, in relation to minimizing undesirable side effects of the drug.

More detailed statistical information about student performance for both cohorts, and related commentary, can be obtained from the assessment reports for these studies from the VCAA website. What is clear from a more general scrutiny of these results, is that, as in 2002, access to CAS appears to enable students to engage, and continue to engage, in extended response analysis questions, with a comparatively good level of success.

Conclusion

The CAS cohort generally scored better on common questions, including those which would be classified as technology independent (I) or technology of assistance but neutral with respect to graphics calculators or CAS (N), and certainly where CAS use would a priori be recognised as of advantageous. This is likely to be due to a natural combination of several factors: the technology active design of the CAS pilot study and that CAS provide a strong model for correct mathematical conventions, forms and notation (context); the open or positive orientation of pilot teachers towards the use of technology (affective); and that access to CAS has helped students understand key mathematical concepts and processes (cognitive).

The 2003 data and this analysis indicate several areas in which CAS access appears to have had beneficial effects on student performance with respect to certain types of question. They also point to some possible areas for investigation of student approaches to the use of *graphics* calculators, with a view to improving this performance.

References

Artigue, M. (2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning*, *7*, 245 – 274.

Ball, L. (2003, June). *Communication of mathematical thinking in examinations: A comparison of CAS and non-CAS student written responses*. Paper presented at the 3rd Computer Algebra in Mathematics Education (CAME) Symposium, Reims, France.

Bohm, J., Forbes, I., Herweyers, G., Hugelshofer, R., & Schomacker, G. (2004, in press). *The Case for CAS.* University of Munster, Austria.

Brown, R. (2003, June). *Comparing system wide approaches to the introduction of Computer Algebra Systems into examinations.* Paper presented at the 3rd Computer Algebra in Mathematics Education (CAME) Symposium, Reims, France.

Burrill, J., Allison, J., Breaux, G., Kastberg, S., Leatham, K., & Sanchez, W. (2002). *Handheld graphing technology in secondary mathematics: Research findings and implications for classroom practice*. Texas Instruments. Michigan State University.

DSME. (2004). *The CAS-CAT Project*. Department of Science and Mathematics Education, the University of Melbourne. Retrieved 23 March 2004 from the World Wide Web: http://www.edfac.unimelb.edu.au/DSME/CAS-CAT/

Evans, M., Leigh-Lancaster, D., & Norton, P. (2003, June). *The Victorian Curriculum and Assessment Authority Mathematical Methods computer algebra pilot study and examinations*. Paper presented at the 3rd Computer Algebra in Mathematics Education (CAME) Symposium, Reims, France.

Flynn, P. (2003, June). *Using assessment principles to evaluate CAS-permitted examinations.* Paper presented at the 3rd Computer Algebra in Mathematics Education (CAME) Symposium, Reims, France.

Garner, S., & Leigh-Lancaster, D. (2003). A teacher-researcher perspective on CAS in senior secondary mathematics. In L. Bragg, C. Campbell, G. Herbert & J. Mousley (Eds.), *Mathematics education research: Innovation, networking, opportunity.* (Proceedings of the 26th annual conference of the Mathematics Education Research Group of Australasia, Geelong, pp. 372-379). Sydney: MERGA.

Leigh-Lancaster, D. (2003). The Victorian Curriculum and Assessment Authority (VCAA) Mathematical Methods (CAS) pilot study 2001 – 2005. In M. Goos (Ed), *Mathematics~Making Waves* (Proceedings of the 19th Biennial Conference. Australian Association of Mathematics Teachers, Adelaide).

VCAA. (2004). *Mathematical Methods Computer Algebra System (CAS) Pilot Study Website*. VCAA. Melbourne. Retrieved 23 March 2004 from the World Wide Web: http://www.vcaa.vic.edu.au/vce/studies/mathematics/mathpilot/casmathpilotindex.html