The 2007 Common Technology Free Examination for Victorian Certificate of Education (VCE) Mathematical Methods and Mathematical Methods Computer Algebra System (CAS)

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

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

Peter Jones
Swinburne University of Technology
<pjones@swin.edu.au>

Magdalena Les
Victorian Curriculum and Assessment Authority
<les.magdalena.m@edumail.vic.gov.au>

Michael Evans
Australian Mathematical Sciences Institute
<michael@amsi.org.au>

Margaret Wu
Victorian Curriculum and Assessment Authority
<wu.margaret.l@edumail.vic.gov.au>

In November 2007 around 14 000 Mathematical Methods students and around 1 500 Mathematical Methods (CAS) students sat a common one hour technology-free examination. The examination covered the same function, algebra, calculus and probability content with corresponding expectations for key knowledge and key skills for both studies. This paper provides some analysis of the examination and items and student performance with respect to both cohorts.

Background

For the period 2006 - 2009, Mathematical Methods Units 1 – 4 and Mathematical Methods (CAS) Units 1 - 4 are parallel and equivalent Victorian Certificate of Education (VCE) mainstream function, algebra, calculus and probability studies, with an approved graphics calculator, and an approved CAS, as their respective assumed enabling technology (see VCAA 2005). Units 3 and 4 are typically studied in Year 12, and have corresponding end-of-year examinations. There is a common one hour short answer question and some extended answer question technology free examination (see VCAA 2007a); and a two hour multiple choice and extended response technology assumed examination, which has substantial common questions as well as some distinctive questions (see VCAA 2007b, VCAA 2007c). From 2009, all students undertaking the mainstream function, algebra, calculus and probability study in Victoria will be enrolled in Mathematical Methods (CAS) Units 1 and 2, in preparation for the final stage of transition for the whole cohort to CAS enabled version of the study at Units 3 and 4 in 2010. Over the past decade or so, various researchers have considered aspects of assessing mathematical capabilities via examinations where students choose to use mental, by hand or technology assisted approaches, or a combination of such approaches, to tackle a range of questions (for example, Kokol-Voljc, 2000; Brown, 2003; Flynn, 2003; and Ball & Stacey; 2007). Particular aspects of student performance on common short answer, multiple choice and extended response questions from examinations for Mathematical Methods and Mathematical Methods (CAS) have been reported on previously by Evans, Jones, Leigh-Lancaster and Norton (2007) and Evans, Leigh-Lancaster and Norton (2003; 2004; 2005).

Other systems and jurisdictions around the world also employ a technology free and technology active examination structure, where the use of CAS is permitted (for example, the College Board *Advanced Placement Calculus Program*) or assumed (for example, Danish *Baccalaureat Mathematics* and CAS based New Zealand National Qualifications Framework Level 1 and 2 *Achievement Standards for Mathematics*) for some component(s) of examination assessment (see CAME, 2007). From 2010, the revised Western Australian mainstream (Mathematics) and advanced (Specialist) calculus based senior secondary courses will be CAS enabled, with a technology free and technology active examination structure (see Curriculum Council WA, 2008).

The areas of study and topics for Mathematical Methods (CAS) encompass those of Mathematical Methods (hereafter referred to as MMCAS and MM respectively) include common specification of key knowledge and key skills in relation to mental or by hand approaches to mathematical routines and procedures (VCAA 2005, p. 67-68, 73-74, 156-157). This paper discusses aspects of the common technology free examination, and student performance across the two cohorts. It should be noted that the group of students taking Mathematical Methods CAS in 2007 is not necessarily a representative sample of all students undertaking the Mathematical

Methods study in 2007. However, the mean scores for both cohorts on the Mathematics component of the 2007 General Achievement Test (GAT) completed in June 2007, that is, prior to the examinations: MM, μ = 24.4, σ = 6.0, and MMCAS, μ = 25.1, σ = 5.9 (from a total available marks of 35); and the Victorian Tertiary Admissions Committee (VCTAC, 2007) scaling report (which makes a post-examination comparison of the performance of all students in a given study with the rest of the student cohort across studies on a truncated normal scale of 0 - 50): MM, μ = 35.73, σ = 7.1 and MMCAS: μ = 35.84, σ = 6.7, indicate that the *overall* level of ability of the two cohorts is very similar.

Comparing Student Mean Performance on the 2007 Paper

In 2007, the common one hour technology free Examination 1 comprised 18 items receiving credit or partial credit (that is a question or part of a question allocated one or more marks), with a total of 40 available marks. A simple comparison of the mean scores is shown in Table 1.

Table 1

Comparison of the Mean Raw Scores for MM and MMCAS on Examination 1

	Group	N	Mean	Std. Deviation	Std. Error Mean
Raw Score	1 non-CAS	13705	16.9248	11.89609	0.10162
	2 CAS	1528	18.6165	11.91652	0.30485

On average, the MMCAS group scored 1.7 points higher than the MM group (Table 1). The independent samples *t*-test analysis confirmed that this difference is statistically significant (t =-5.272, p = 0.0001). A comparison of the mean scores of the MMCAS and MM groups for each item shows either no difference, or a positive difference in favor of the MMCAS group, as can be seen from Figure 1.

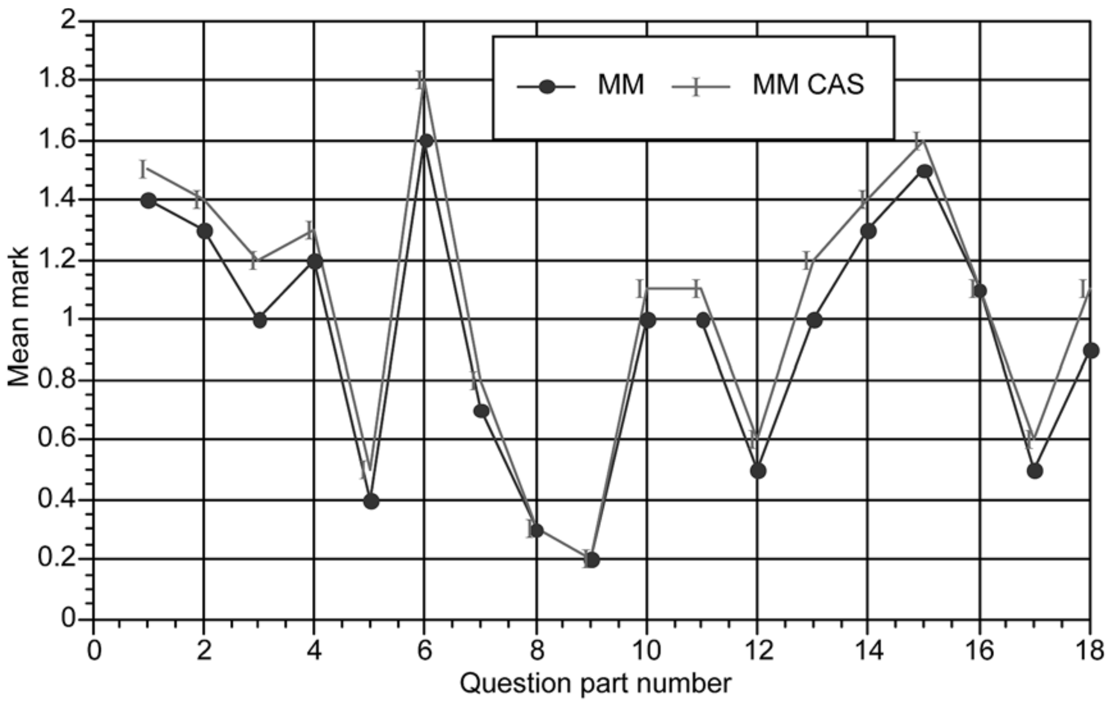

Figure 1. Mean question-part mark by population (MM CAS and MM).

These results contrast with those for Examination 1 in 2006 (see Evans, Jones, Leigh-Lancaster, & Norton, 2007), which showed essentially the same performance on all items, and a corresponding mean overall score of 22 for both groups. The lower overall mean raw score for 2007 is consistent with expert judgment and practitioner views that the Examination 1 paper was more challenging than that of the previous year - assuming that no significant variation in the overall ability of the 2007 student cohort with respect to that of 2006 and earlier years.

Several considerations may be relevant - the 2007 MMCAS group could, on average, be of slightly higher ability than the 2007 MM group. To date, 'high-achieving schools' across the sectors have generally been more inclined to continue with the MM study, where the good performance of their student cohort has been historically well established, intending to make the change to MMCAS at a later stage in the transition process. On the other hand, low-achieving schools' may also have been cautious about taking up MMCAS at an early stage, and hence also make a later transition. It may be the case that as teachers of students in the MMCAS group continue to develop their familiarity and expertise with the technology as a pedagogical tool, there is increased benefit in terms of the scaffolding it affords students in learning mental and by hand computational skills with respect to function, algebra and calculus. Such considerations, and their possible or likely impact, will be able to be more fully explored as further longitudinal data becomes available, and the number of students in the MMCAS cohort continues to increase.

Comparison of Students' Performance (CAS and non-CAS groups) on Each Individual Item

At individual item level, one can test to see if the MMCAS group performed differentially from the MM group. The detection of differential performance between the MMCAS and MM groups on individual items is carried out using item response theory, or IRT (see Lord, 1980). Item response theory was applied to control for differences in students' abilities, where ability can be regarded as a monotonic transformation of the test raw score. That is, there is a one-to-one correspondence between each raw score on the test and the estimated ability on the IRT scale. In general terms, the IRT is concerned with the probability of success when someone attempts multiple-choice or short answers questions. The basic idea behind that theory is that there exist quantities b_i – the difficulty of item *i*, and θ_j - the ability of student *j*, and that p_{ij}, the probability of student *j* responding correctly to item *i* is given by an increasing function of $(\theta_j - b_i)$. The Rasch model (see Rasch, 1960) for the probability of success for a student *j* on an item *i* can be expressed by:

$$p_{ij} = \frac{\exp(\theta_j - b_i)}{1 + \exp(\theta_j - b_i)}$$

In the case where an item has partial credit scoring (that is, the maximum score is greater than 1), the IRT partial credit model (Masters, 1982) can be applied. Using *ConQuest* software (Wu et al., 1998) graphs have been produced for all 18 items. Each of these items was scaled in the item response modelling as the same item for both cohorts. Figure 2 shows a selection of graphs for these items, the blue dotted curve (darker) shows the observed average score of students from the MM group (cas1), while the green dotted curve (lighter) shows the observed score of students from the MMCAS group (cas2), as a function of the "ability level" (labelled as Latent Trait on the horizontal axis). The latent trait level for each student can be regarded as derived from a nonlinear, monotonic, transformation of the total test score, based on the IRT model. The solid line shows the expected score of students as a function of latent trait. There is only one solid line in each graph, as each of these items was modelled as the same item whether it appeared in the MMCAS paper or the MM paper (both papers were identical). For each item the relevant mathematical content has been briefly identified.

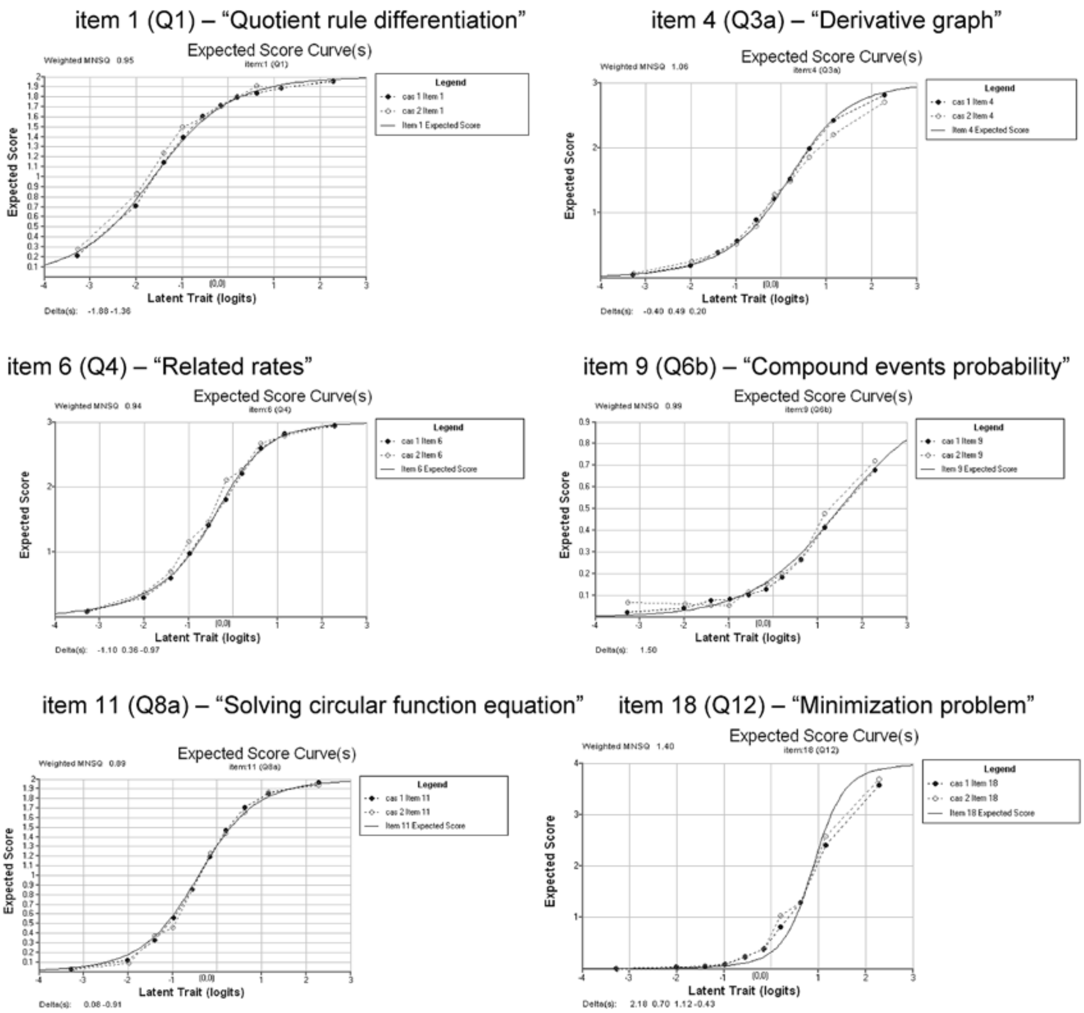

Figure 2. Expected score curves for all test items by the MMCAS and MM groups.

The results reveal that overall, two groups performed similarly across each item, after controlling for the total score on the test, for example, for Items 1, 6 and 11 the theoretical and cohort curves almost overlap each other. Marginal differences are observed between the two groups' performance on Items 4, 9 and 18. Whereas the higher ability students from the MMCAS group have a slightly lower score than higher ability students from the MM group on Item 4, the reverse is the case for Items 9 and 18 where higher ability CAS students slightly outperformed higher ability MM students. Also, lower ability CAS students performed better on Item 9 than lower ability MM students.

In comparing the observed curves with the theoretical (expected) curves, for Item18, there is a slight but noticeable difference. The observed curves for both are not as steep as the expected curve, indicating that the item does not discriminate between lower ability and higher ability students as well as the model predicts. The other five selected show a close match between the observed scores and the expected scores (both dotted lines almost overlap with the solid line), indicating that the items fitted the IRT model well.

The application of IRT facilitates the comparison of performance of MMCAS and MM groups at individual item level. In general, the two groups of students performed very similarly on almost all items after controlling for the total score on the test. There are some marginal differences for a few items, but these differences are small. It is re-assuring that, overall, there is no strong evidence of differential performance on individual items between MMCAS and MM groups.

Conclusions

The two cohorts display essentially the same *profile* on almost all questions considered using IRT partial credit model of analysis on an item by item basis. While the mean total score on the paper for both cohorts was similar, there was a small but significant difference in favour of the CAS group. On an item by item basis, the difference in mean score MMCAS – MM was non-negative in all cases, and generally positive. This suggests that the common curriculum requirements (content and expectations) for *both* studies (in terms of key knowledge and key skills specified in the study designs) with respect to mental and by hands skills of the type tested on the common Examination 1, provide a robust basis for very developing similar overall levels of student performance for this type of assessment, where students do not have access to technology, irrespective of which cohort they come from.

The view that access to CAS as an enabling technology is likely to *diminish* student mental and by hand capabilities of the kind assessed by the common Examination 1 (where the enabling technology is not available) relative to students who do not have such access, is not supported by the data from 2006 and 2007. Similar observations have been made in other studies over the years (see, for example, Heid, 1984; Park & Travers, 1996, Dunham, 2000), including within school testing at both Unit 1 and 2 and Unit 3 and 4 levels during the VCAA Mathematical Methods (CAS) pilot study 2000 – 2005.

References

Ball, L., & Stacey, K. (2007) Using technology in high-stakes assessment: How teachers balance by-hand and automated techniques. In C-S. Lim et al., *Proceedings of the 4th East Asia Regional Conference on Mathematics Education* (pp 90 – 97). Penang, Malaysia: Universiti Sains Malaysia.

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.

Computer Algebra in Mathematics Education (2007). *Curriculum Information - International overview of CAS active curricula, assessment and qualifications*. London: CAME. Retrieved 20 March 2008 from the World Wide Web: http://www.lkl.ac.uk/research/came/curriculum.html

Curriculum Council WA (2008). *Senior secondary courses: Mathematics and Mathematics-specialist*. Retrieved 20 March 2008 from the World Wide Web: http://www.curriculum.wa.edu.au/internet/Senior_Secondary/Courses/

Dunham, P. (2000). Highlights from research on graphing calculators. *Proceedings of the Post ICME 10 Teachers Teaching with Technology Conference*, *Tokyo, Japan*. Retrieved 20 March 2008 from the World Wide Web: http://www.math.ohio-state.edu/~waitsb/papers/t3_posticme2000/dunham.pdf

Evans, M., Jones, P., Leigh-Lancaster, D., & Norton, P. (2007). Mathematical Methods and Mathematical Methods Computer Algebra System (CAS) 2006 - Concurrent implementation with a common technology free examination. In J. Watson & K. Beswick (Eds.), *Mathematics: Essential research, essential practice* (Proceedings of the 30th annual conference of the Mathematics Education Research Group of Australasia, Melbourne, pp. 223-230). Sydney: MERGA.

Evans, M., Leigh-Lancaster, D., & Norton, P. (2003). The Victorian Curriculum and Assessment Authority (VCAA) Mathematical Methods Computer Algebra System (CAS) pilot study examinations 2003. In L. Bragg, C. Campbell, G. Herbert, & J. Mousley (Eds.), *MERINO, Mathematics Education Research: Innovation, Networking, Opportunity* (Proceedings of the 26th annual conference of the Mathematics Education Research Group of Australasia, Geelong, VIC, pp. 372-379). Sydney: MERGA.

Evans, M., Leigh-Lancaster, D., & Norton, P. (2004). *The Victorian Curriculum and Assessment Authority (VCAA) Mathematical Methods Computer Algebra System (CAS) pilot study examinations 2004*. In I. Putt, R. Faragher, & M. McLean (Eds.), Mathematics education for the third millenium: Towards 2010 (Proceedings of the 27th annual conference of the Mathematics Education Research Group of Australasia, Townsville, QLD, pp. 223-230). Sydney: MERGA.

Evans, M., Leigh-Lancaster, D., & Norton, P. (2005). (VCAA) Mathematical Methods Computer Algebra System (CAS) 2004 pilot examinations and links to a broader research agenda. In P. Clarkson, A. Downton, D. Gron, M. Horne, A. McDonough, R. Pierce and M. McLean (Eds.), *Building connections: Theory, research and practice* (Proceedings of the 28th annual conference of the Mathematics Education Research Group of Australasia, Melbourne, pp. 223-230). Sydney: MERGA.

Flynn, P. (2003). Adapting “problems to prove” for CAS-permitted examinations. *The International Journal of Computer Algebra in Mathematics Education, 10*(2), 103-121.

Heid, M. K. (1984). An exploratory study to examine the effects of resequencing skills and concepts in an applied calculus curriculum through the use of the microcomputer. Doctoral dissertation, University of Maryland. *Dissertation Abstracts International, 46*, 1548A.

Kokol-Voljc, V. (2000). Examination questions when using CAS for school mathematics teaching. *The International Journal of Computer Algebra in Mathematics Education, 7*(1), 63-75.

Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Hollsdale, NJ: Erlbaum.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149-174.

Park, K., & Travers, K. J. (1996). A comparative study of a computer-based and a standard college first-year calculus course. *CBMS Issues in Mathematics Education, 6*, 155-176.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Victorian Curriculum and Assessment Authority (2005). *Mathematics Victorian Certificate of Education study design*. VCAA. Melbourne. Retrieved 20 March 2008 from the World Wide Web: http://www.vcaa.vic.edu.au/vce/studies/index.html

Victorian Curriculum and Assessment Authority (2007a). *Mathematical Methods written examination 1*. Melbourne: VCAA. Retrieved 20 March 2008 from the World Wide Web: http://www.vcaa.vic.edu.au/vce/studies/mathematics/methods/pastexams/2007/2007mm1.pdf

Victorian Curriculum and Assessment Authority (2007b). *Mathematical Methods written examination 2*. Melbourne: VCAA. Retrieved 20 March 2008 from the World Wide Web: http://www.vcaa.vic.edu.au/vce/studies/mathematics/methods/pastexams/2007/2007mm2.pdf

Victorian Curriculum and Assessment Authority (2007c). *Mathematical Methods (CAS) written examination 2*. Melbourne: VCAA. Retrieved 20 March 2008 from the World Wide Web: http://www.vcaa.vic.edu.au/vce/studies/mathematics/cas/pastexams/2007/2007mmCAS2.pdf

Victorian Tertiary Admissions Committee (2008). 2007 *scaling report.* Melbourne: VTAC. Retrieved 20 March 2008 from the World Wide Web: http://www.vtac.edu.au/pdf/scaling_report.pdf

Wu, M. L., Adams, R. J., & Wilson, M. R. (1998). *ConQuest: Generalised item response modelling software.* Camberwell: Australian Council for Education Research.