Examinations in the Final Year of Transition to Mathematical Methods Computer Algebra System (CAS)

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

2009 was the final year of parallel implementation for Mathematical Methods Units 3 and 4 and Mathematical Methods (CAS) Units 3 and 4. From 2006-2009 there was a common technology-free short answer examination that covered the same function, algebra, calculus and probability content for both studies with corresponding expectations for key knowledge and key skills. There was also separate technology-assumed examinations comprising common and different multiple choice and extended response questions. In 2009 the two cohorts were of similar size, comprising 7000-8000 students each. This paper analyses student performance for both cohorts with respect to these common items.

From 2006 to 2009 Mathematical Methods Units 1-4 and Mathematical Methods (CAS) Units 1-4 were implemented as parallel and equivalent Victorian Certificate of Education (VCE) mainstream function, algebra, calculus and probability studies. An approved graphics calculator, and an approved CAS were the respective assumed enabling technology. Units 3 and 4 are typically studied in Year 12, and have end-of-year examinations. There was a common one-hour short answer technology-free examination and a two-hour multiple choice and extended response technology-assumed examination. The latter had many common questions and some distinctive questions. In 2009 all students undertaking the mainstream function, algebra, calculus and probability study in Victoria enrolled in Mathematical Methods (CAS) Units 1 and 2 in preparation for the final stage of transition to the to CAS enabled study at Units 3 and 4 in 2010.

The areas of study and outcomes for Mathematical Methods (CAS) encompass those of Mathematical Methods (hereafter referred to as MMCAS and MM respectively) and incorporate common specification of key knowledge and key skills in relation to mental or by hand approaches to mathematical routines and procedures. Thus the common technology-free examination is based on content from the areas of study for MM in relation to Outcome 1: “On completion of each unit the student should be able to define and explain key concepts as specified in the content from the areas of study, and apply a range of related mathematical routines and procedures”. In particular the elaborating key knowledge and key skill statements for this outcome clearly indicate expectations for knowledge, such as exact values for circular functions, and by hand skills, such as differentiation of combined functions using chain, product and quotient rules.

Researchers such as Kokol-Voljc (2000); Brown (2003); Flynn (2003) and Ball and Stacey (2007) have, over the past decade, considered various 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, as noted by, for example Evans, Jones, Leigh-Lancaster, and Norton (2008). To the best knowledge of the authors, the Victorian parallel implementation of MMCAS and MM is a unique context within which the achievement of two senior secondary

L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics education: Proceedings of the 33rd annual conference of the Mathematics Education Research Group of Australasia.* Fremantle: MERGA.

certificate student cohorts comprising like but distinct populations of substantial size can be considered with respect to common assessment instruments and items. While other systems and jurisdictions around the world have also employed a technology-free and technology-active examination structure where the use of CAS is permitted (for example, the College Board *Advanced Placement Calculus AB and BC Program*) or assumed (for example, Danish *Baccalaureat Mathematics*) these systems and jurisdictions have not, to our knowledge, undertaken collection and publication of large scale data of a like kind.

Mean Performance on Items from Examination 1 2006-2009

In 2009, the common one-hour technology-free Examination 1 (VCAA 2009a) comprised 16 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 can be made from the respective 2009 MM and MMCAS Assessment Reports (VCAA, 2009b, 2009c)

Mean score data on Examination 1 over the period 2006-2009 indicates that, in general, the MMCAS cohort performed at least as well as the MM cohort. In particular, for 2009, the distribution of student scores for each cohort across the mark range from 0 to 40 shows that at the very top end of the mark range the performance of the two cohorts is essentially identical; at the very bottom end the performance of the MMCAS cohort tends to be better (with some ‘noise’), while from the low to high mark range the MMCAS cohort consistently achieves a slightly higher score than the MM cohort. This pattern persists when the data is factored for general mathematical ability using the mathematics, science and technology (MST) component of the General Ability Test (GAT), conducted in the middle of the same year, as a control for ability. It should be noted that the MST component of the GAT has only a moderate positive association with study VCE Mathematics achievement, and is hence a partial indicator of study specific mathematical ability.

Performance by Item on Examination 2 in 2009

Examination 2 has the same structure for both MM and MMCAS, comprising a collection of 22 multiple choice questions worth a total of 22 marks, and four extended response questions each of several parts (with partial credit available) worth a total of 58 marks. For both components there are around 70% to 80% common items. These are either *technology-independent* (that is technology is not of assistance, for example a conceptual question) or *technology-active* but graphics calculator/CAS *functionality neutral* (for example a graphing or numerical equation solving question). In the following analysis performance on the common Examination 1 has been used as a control for ability with respect to performance on common multiple choice and extended response items for MM Examination 2 and MMCAS Examination 2.

Common Multiple Choice Questions in 2009

There were 17 out of 22 common multiple choice items the five different items in 2009 being questions 1, 4, 5, 12 and 18. Figure 1 shows the average score on 17 common items from Examination 2 for each Mathematical Methods Examination score out of 40.

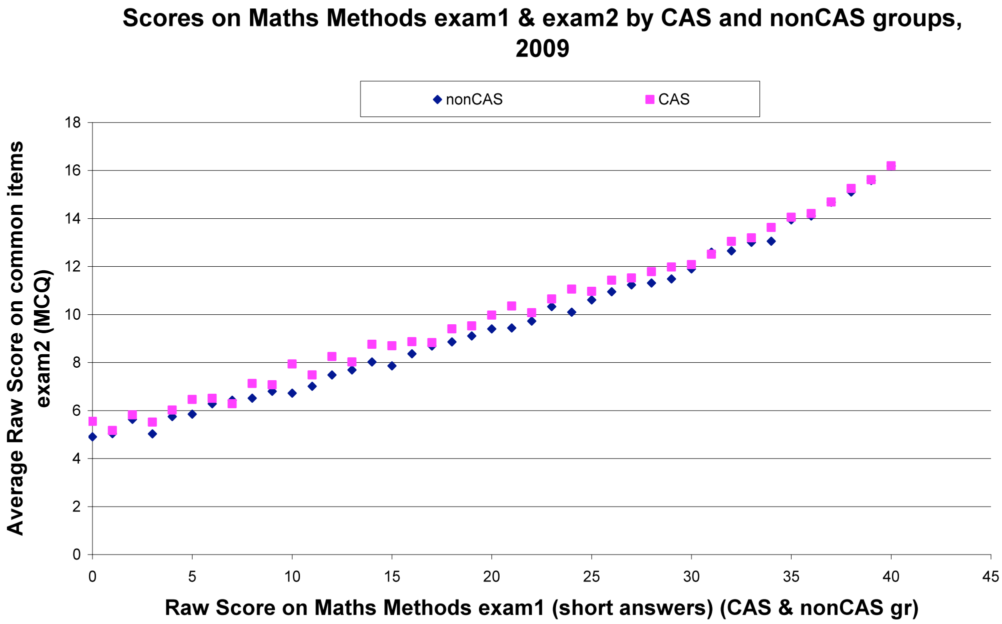

Figure 1. Average score with respect to Examination 1 (technology-free) score.

The largest difference of 1.2 score points occurs for the score of 10. The following histograms in Figure 2a and 2b provide an indication of how many students are in each data point in Figure 1 for MM and MMCAS cohorts respectively:

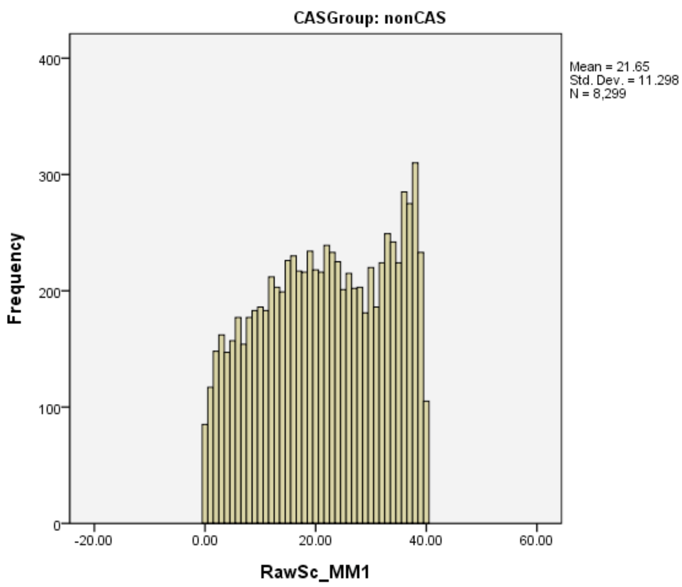

Figure 2a. Histogram for distribution of scores on MM Examination 1 in 2009.

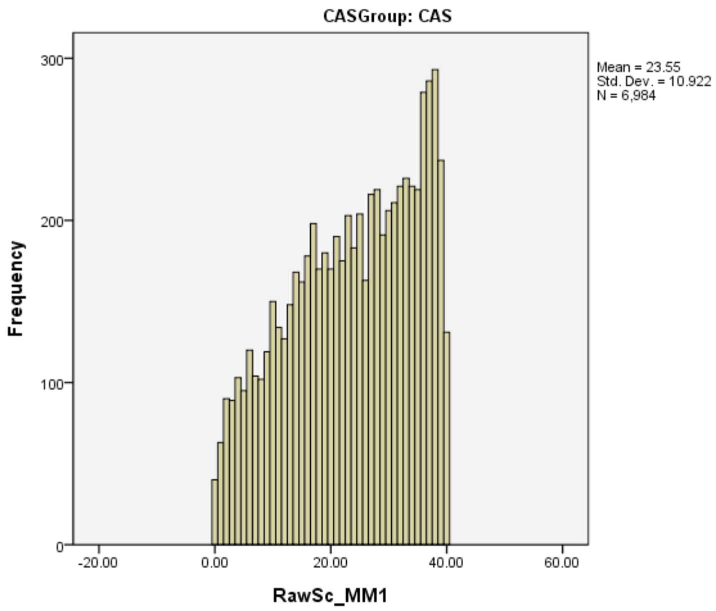

Figure 2b. Histogram for distribution of scores on MMCAS Examination 1 in 2009.

To compare MMCAS and MM cohort differences across Examination 2 multiple-choice items scores while controlling for the results on Examination 1, the Rasch two-dimensional regression model was applied, using *ConQuest* software. Figure 3 shows a comparison of mean raw scores on 17 common multiple-choice items while controlling for students' ability estimated on Examination 1.

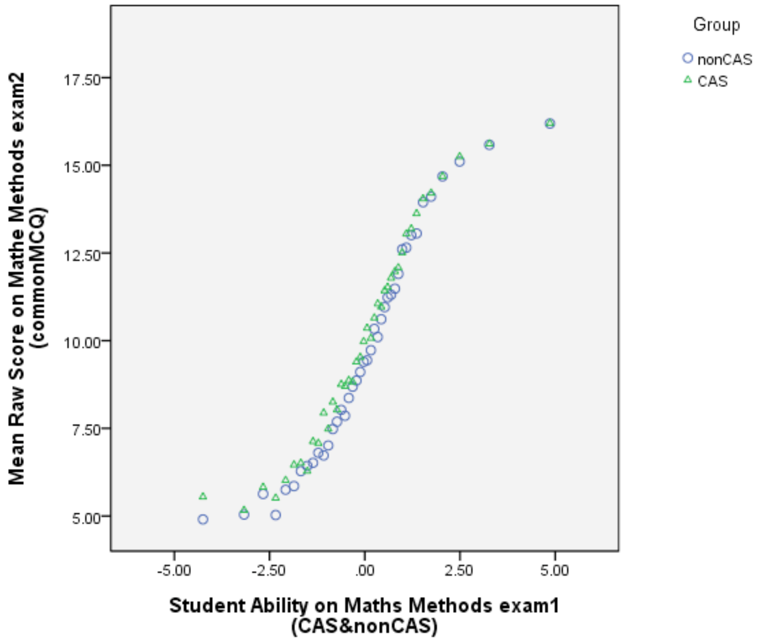

Figure 3. Comparison of student estimated ability (Examination 1) and expected score on common multiple-choice items (Examination 2).

It can be observed again from Figure 3 that the curves for CAS and non-CAS groups are very close, although the average score on common multiple-choice items for the MMCAS cohort is slightly above that for the MM cohort. The largest difference between the expected scores for those two groups is around 1.2 score points for ability close to -1.08 logit on Examination 1.

Common Extended Response Questions in 2009

The extended response component of Examination 2 comprised 32 items on the MM paper and 33 items on the MMCAS paper, with the same total score of 58 from this component for each paper. Of these items, 21 were common for a total score of 35 marks. There were also a range of similar, but not identical items, which have not been included in this analysis. Figure 4 shows the average score on the 21 common items from Examination 2 for each Mathematical Methods Examination score out of 40.

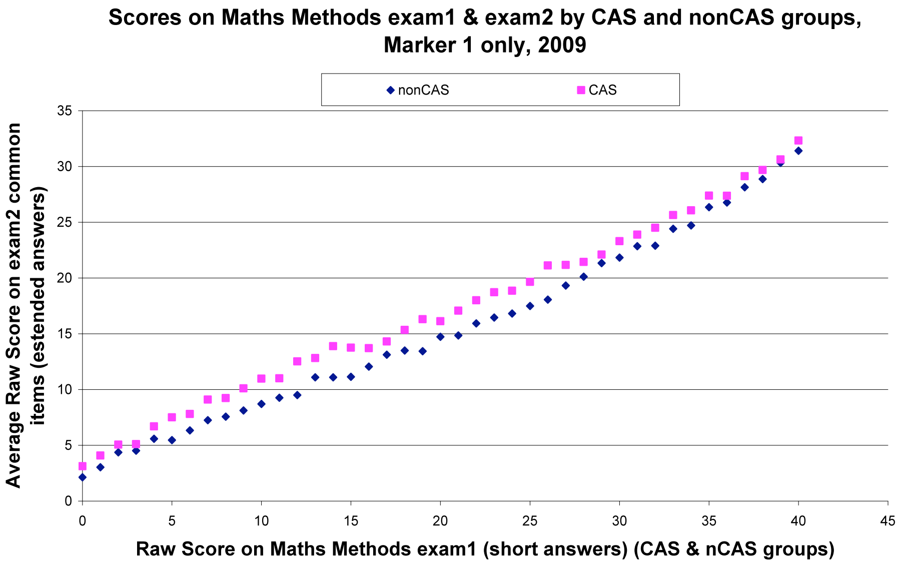

Figure 4. Average score with respect to Examination 1 (technology-free) score.

Figure 4 shows that the average score on Examination 2 common extended answer questions for the MMCAS cohort is slightly higher than that for the MM cohort, except at the very low and top end where the two curves almost overlap. The mean difference is less than 2 score points with the largest difference of 3.1 score points for the score of 26 on Examination 1. The following histograms in Figure 5a and 5b provide an indication of how many students are in each data point in Figure 1 for MM and MMCAS cohorts respectively:

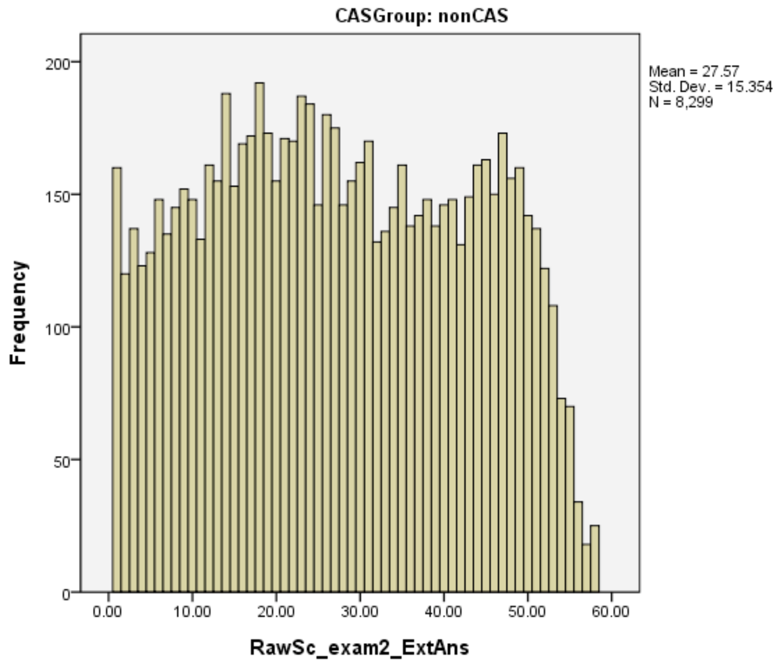

Figure 5a. Histogram for distribution of scores on MM Examination 2 in 2009 (all items).

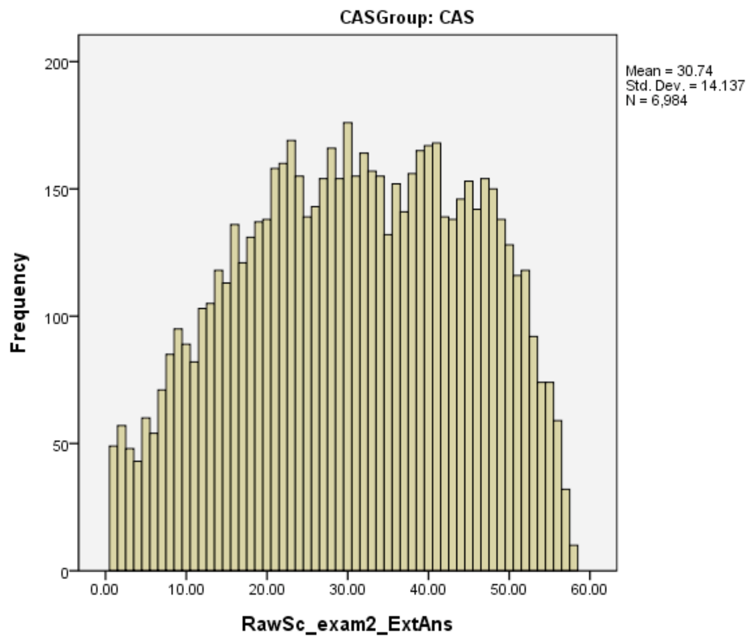

Figure 5b. Histogram for distribution of scores on MMCAS Examination 2 in 2009 (all items).

To compare MMCAS and MM cohort differences across Examination 2 multiple-choice items scores while controlling for the results on Examination 1, the Rasch two-dimensional regression model was applied, again using *ConQuest* software. Figure 6 shows a comparison of mean raw scores on 21 common extended response items while controlling for students' ability estimated on Examination 1.

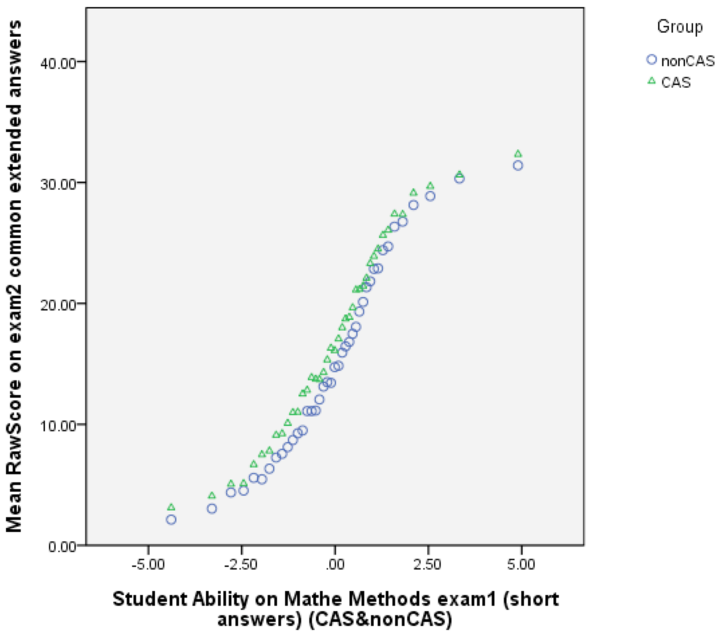

Figure 6. Comparison of student estimated ability (Examination 1) and expected score on common multiple-choice items (Examination 2).

It can be observed from Figure 6 that the curve for the MMCAS cohort is above that for the MM cohort except for the higher range of abilities where curves almost overlap. The largest difference between the expected scores on common extended responses, for those two groups, is around 3 score points for ability close to 0.56 logit on Mathematical Methods Examination 1.

Conclusions

Data from the 2009 *common technology-free* Examination 1 supports preliminary observations from earlier years of parallel implementation and indicates that, in general, the MMCAS cohort performed at least as well as the MM cohort with respect to the sorts of questions set for this examination. Such questions substantially cover expectations for traditional knowledge and mental and/or by hand skills related to the study of functions, algebra, calculus and probability. Similarly, data from *common* items (multiple choice and extended response) on the 2009 MM and MMCAS *technology active* Examination 2, also indicate that, in general, the MMCAS cohort performed at least as well as the MM cohort with respect to the sorts of questions set for this paper which substantially cover a combination of traditional knowledge and by hand skills, and the use of numerical and graphing graphics calculator/CAS functionality.

References

Ball, L., & Stacey, K. (2007). Using technology in high-stakes assessment: how teachers balance by-hand and automated techniques. In Lim, C.-S., Fatimah, S., Munirah, G., Hajar, S., Hashimah, M.Y., Gan, W.L., & Hwa, T.Y. (Eds.) ***Proceedings of EARCOME4 2007 4th East Asia Regional Conference on Mathematics Education.*** (pp 90-97) Universiti Sains Malaysia, Penang, Malaysia.

Brown, R. (2003). *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.

Evans, M., Jones, P., Leigh-Lancaster, D., Les, M., Norton, P. & Wu, M. (2008). The 2007 Common Technology Free Examination for Victorian Certificate of Education (VCE) Mathematical Methods and Mathematical Methods Computer Algebra System (CAS). In M. Goos, R. Brown & K. Makar (Eds.) *Navigating currents and charting directions* (Proceedings of the 31st annual conference of the Mathematics Education Research Group of Australasia, pp. 331-336). Brisbane: 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.

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.

Victorian Curriculum and Assessment Authority. (2009a). *Mathematical methods written examination 1*. VCAA. Melbourne. Retrieved 22 March 2009 from: http://www.vcaa.vic.edu.au/vcaa/vce/studies/mathematics/methods/pastexams/2009/2009mm1-w.pdf

Victorian Curriculum and Assessment Authority. (2009b). *Mathematical methods examination 1 Assessment Report 2009*. VCAA. Melbourne. Retrieved 22 March 2009 from: http://www.vcaa.vic.edu.au/vcaa/vce/studies/mathematics/methods/assessreports/2009/mm1_assessrep_09.pdf

Victorian Curriculum and Assessment Authority. (2009c). *Mathematical methods (CAS) examination 1 Assessment report 2009*. Retrieved 20 March 2009 from: http://www.vcaa.vic.edu.au/vcaa/vce/studies/mathematics/cas/assessreports/mmcas1_assessrep_09.pdf